**Dissociable roles for the lateral orbitofrontal cortex in Pavlovian acquisition**

**Abstract**

The orbitofrontal cortex (OFC) is a critical structure in the flexible control of value based behaviours. OFC dysfunction is often only detected when task or environmental contingencies change, and initial acquisition learning and behaviour is thought to be intact. While intact acquisition following OFC lesions in simple Pavlovian cue-outcome is fundamental to models of OFC function, this predicted null effect has not been thoroughly investigated. Here we test the effects of lesions and temporary muscimol inactivation of the rodent lateral OFC on the acquisition of a simple single cue-outcome relationship. Surprisingly, pre-training lesions significantly enhance acquisition after over-training whereas post-training lesions and inactivation significantly impair acquisition. This impaired acquisition to the cue reflects a disruption of behavioural control and not learning since the cue could also act as an effective blocking stimulus in an associative blocking procedure. Finally, we demonstrate that the impaired behavioural control to the Pavlovian cue reflects a specific inability to flexibly update Pavlovian behaviours based on the current value of alternative behavioural options. Contrary to most theories of OFC function, these findings suggest that while OFC function is often associated with flexible behavioural control in complex environments, it is also involved in the simple Pavlovian acquisition.

The orbitofrontal cortex (OFC) is critical to behavioural flexibility when learning and behaviour need to be updated to reflect a change in the environment (Gardner et al., 2019; Klein-Flugge et al., 2013; Kringelbach, 2005; Murray & Rudebeck, 2018; Rudebeck & Murray, 2014). In particular, the OFC is necessary for appropriately updating behavior when the contingencies between predictive cues and outcomes change, or when outcomes change in value (Panayi & Killcross, 2018; Pickens et al., 2005; Walton et al., 2011). Information encoded in the OFC about the expected value and identity of predicted outcomes is necessary for flexibly updating behaviour when these outcome features change. Population and single-unit neuronal firing in the OFC encodes many features of reward outcomes (e.g. size, preference, identity, time, location, probability, certainty, salience (Delamater, 2007; Ogawa et al., 2013; Padoa-Schioppa, 2009; Sadacca et al., 2018; Stalnaker et al., 2014; Takahashi et al., 2013; Zhou et al., 2019)), furthermore the coding of these features develops over the course of learning to predictive cues in anticipation of the expected outcome (Schoenbaum et al., 2009). There is also substantial evidence to suggest that this outcome expectancy information in the OFC is incorporated into mid-brain dopaminergic reward prediction errors (Takahashi et al., 2011), which are critical for learning (Schultz, 1998; Steinberg et al., 2013).

However, despite these close ties to the learning process, the OFC is not necessary for the initial acquisition learning (comprehensively documented by Delamater, 2007; Izquierdo, 2017; Murray et al., 2007; Rudebeck & Murray, 2014; Stalnaker et al., 2015) in all but the most complex of circumstances (e.g. Walton et al., 2011). Lesions and functional inactivation of the OFC do not disturb initial learning about Pavlovian cue-outcome relationships in a range of tasks, and instead only reveal their effects when the cue-outcome relationships change, or when the value of expected outcomes change such as in reversal learning and outcome devaluation procedures (Butter, 1969; Dias et al., 1996; Gallagher et al., 1999; Iversen & Mishkin, 1970; Schoenbaum et al., 2003; West et al., 2011). To account for these effects, one class of OFC theories suggests that the OFC is necessary for representing information about the sensory-specific properties or identity of expected outcomes (Burke et al., 2008; Delamater, 2007; Schoenbaum et al., 2009, 2011). A second, but complementary class of theories using a reinforcement learning framework suggests that the OFC is necessary for the representation of latent state information (Wilson et al., 2014). In reinforcement learning models, tasks such as Pavlovian conditioning can be divided into discrete physically observable states, such as “cue on”, “cue off”, and “reward”, and underlying latent states signaled by partially observable information recalled into working memory such as reinforcement history.

Both theories, while couched in different computational and theoretical frameworks, suggest similar roles for the OFC. Latent states encompass specific outcome expectancies and include a broader category of potential stimuli (e.g. internal context (Niv, 2019)). Implicit in these theories is that initial acquisition should be affected by OFC dysfunction if performance depends on specific outcome expectancy or latent states (e.g. the differential outcomes effect (Boulougouris et al., 2007; Boulougouris & Robbins, 2009; McDannald et al., 2005); complex multiple-choice probabilistic learning tasks (Walton et al., 2011)), but not in putatively “simple” single CS-US learning tasks (Gallagher et al., 1999) where the outcome identity and value stays constant and is reliably predicted by the CS. While this null effect is often reported in procedures involving learning about multiple CSs and/or USs (Burke et al., 2008; Panayi & Killcross, 2018; Schoenbaum et al., 2009), there is little evidence from tasks involving only a single CS-US relationship where a null result is clearly predicted. For example, Gallagher et al (1999) found no effect of complete OFC lesions on single CS-US acquisition but stopped training before behaviour reached asymptote (Schoenbaum et al., 2003).

Both latent state and sensory-specific outcome expectancy theories of OFC function predict a null effect of OFC lesions on initial acquisition learning, particularly in situations involving only a single CS-US relationship. Indeed, this null effect is often reported as an important feature of OFC dysfunction (Murray et al., 2007; Schoenbaum et al., 2009; Stalnaker et al., 2015; Wilson et al., 2014). Here we directly tested this prediction in rats trained on a single CS-US Pavlovian task following lesions targeting the lateral OFC. Surprisingly, OFC lesions significantly increased Pavlovian acquisition behaviour after extended training. In contrast, post-training lesions and intra-OFC infusions of muscimol impaired Pavlovian acquisition behaviour. Using an associative blocking design, we confirmed that even though behaviour was impaired, the underlying learning about the CS-US contingency was left intact. Finally, we confirmed that impaired Pavlovian acquisition behaviour following post-training OFC inactivation might reflect an inability to modulate Pavlovian behaviours relative to the value of alternative behavioural options.

**Results**

**Experiment 1: Pre-training OFC lesions**

***Acquisition***

Pre-training OFC lesions significantly increased responding to the Pavlovian cue relative to sham control animals (Figure 1A; lesions depicted in Figure 1-figure supplement 1). Analysis of conditioned responding was conducted as a CS-PreCS difference score such that levels of responding reflected discriminative performance to the cue (CS) above baseline (PreCS). Acquisition of responding to the CS was significantly greater in the lesion group than the sham group (main effect of Group , , Block , , and Group x Block interaction , ). Follow up comparisons on each block revealed that responding in the lesion group was significantly higher than the sham group during the last 4 blocks (Block 1 , , Block 2 , , Block 3 , , Block 4 , , Block 5 , , Block 6 , , Block 7 , ). Given the ubiquity of non-significant effects of OFC lesions on acquisition learning in the literature, two independent replications of this novel effect were conducted (combined here; same pattern of statistical significance in both independent replications) which confirmed the effect was robust.

***Locomotor activity***

The enhanced responding observed during acquisition in the OFC lesion group could simply reflect an enhancement of general locomotor activity. However locomotor activity (Figure 1C) did not differ between groups (main effect of TimeBin , , but no significant effect of Group , , or Group x TimeBin interaction , ). Therefore, the enhanced responding during acquisition was not simply due to OFC lesions inducing hyperactivity.

***Satiety***

To test whether the enhanced responding following OFC lesions was sensitive to levels of hunger or shifts in general motivation, a subgroup of animals (subgroup 1) was tested when sated, i.e. following 24 hours *ad libitum* access to home-cage food (Figure 1B). General satiety, did not affect the rate of responding in the sham group (Sham: Satiety vs Hungry , ) but significantly suppressed responding in the lesion group (Lesion: Satiety vs Hungry , ) compared to subsequent testing 24 hours later when hungry again (no significant main effect of Group , , but a significant main effect of Hunger , , and Group x Hunger interaction , ). Since the satiety test session was rewarded, it is possible that OFC lesioned animals could learn that the reward was less valuable by direct experience with the reward, similar to incentive learning effects normally observed in instrumental conditioning (Dickinson & Balleine, 2002). However, this possibility is unlikely as responding was comparable between groups on the first trial of the satiety test (, , Figure 1-figure supplement 2), before the first reward was delivered. This suggests that animals with OFC lesions are sensitive to shifts in hunger and general motivation.

***Devaluation Test***

OFC lesions have been shown to cause characteristic deficits in Pavlovian outcome devaluation (Gallagher et al., 1999; Panayi & Killcross, 2018; Pickens et al., 2003, 2005). Therefore, to test whether the present lesion manipulation was comparable to other reports we tested a subgroup of animals (subgroup 2) on Pavlovian outcome devaluation. First the sham and lesion animals were given novel acquisition training of two novel and unique cue-outcome relationship (Figure 1-figure supplement 3A). A specific taste aversion was then established by pairing consumption of one of the outcomes with illness (i.p. injection of lithium chloride; Devalued), and the value of the other outcome was left intact (i.p. injection of saline; Non-Devalued). Both groups learned the novel cue-outcome associations and acquired the specific taste aversion (Figure 1-figure supplement 3B).

Finally, during a devaluation test (Figure 1D), the two cues were presented in extinction. The sham group showed a significant devaluation effect, i.e. responding was lower to the devalued than non-devalued cue (, ). In contrast, the devaluation effect was abolished in the lesion group, and responding remained high to both the devalued and non-devalued cue (, ; Significant Group x Cue interaction , , but no main effect of Group , , or Cue , ). This finding successfully replicates the finding that both complete OFC and focal lateral OFC lesions abolish the outcome devaluation effect in rodents (Gallagher et al., 1999; Panayi & Killcross, 2018; Pickens et al., 2003, 2005).

**Experiment 2: Post-training muscimol inactivation**

The enhanced Pavlovian responding observed following OFC lesions (Figure 1A) may be due to enhanced learning of the cue-outcome relationship in the OFC lesion group (Figure 2-figure supplement 1). This is consistent with a role for the OFC in representing outcome expectancy information. For example, incremental learning about a cue-outcome relationship is thought to depend upon prediction errors (Esber & Haselgrove, 2011; LePelley, 2004; Mackintosh, 1975; Nasser et al., 2017; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Sutton & Barto, 1998), i.e. the difference between the experience outcome value and the expected outcome value. The expected outcome value of a cue is incrementally updated until this prediction error discrepancy is minimised. If the OFC carries some aspect of outcome expectancy information (Baxter et al., 2000; Pears et al., 2003; Schoenbaum et al., 2009; Takahashi et al., 2009, 2011), then OFC lesions might consistently reduce/underestimate the expected value of a cue which in turn would result in abnormally persistent prediction errors and enhanced learning. Therefore, disruption of OFC function should temporarily lower expected value, and enhance prediction errors and learning (for modelling of this prediction see Figure 2-figure supplement 1). We tested this hypothesis by inactivating the OFC after first successfully acquiring cue-outcome learning i.e. when expected value is high and prediction errors are low. If the OFC carries some aspect of the learned expected value, then inactivation of the OFC should enhance prediction errors, and responding should increase to reflect new learning. Following this, returning function to the OFC should result in an over-expectation of the value of the outcome, and performance should decrease to reflect the extinction of this over-expectation. Importantly, while this account is couched in terms prediction-error learning mechanisms, the prediction remains true for any account of OFC lesions enhancing learning (Figure 2-figure supplement 1).

We tested this hypothesis by first training a new group of animals on the same simple Pavlovian task for 9 days, before implantation of bilateral cannulae targeting the OFC (Figure 2A, Days 1-9; significant main effect of Day , , but no main effect of Group , , or Group x Day interaction , ). Following post-operative recovery (histology depicted in Figure 2-figure supplement 2), and prior to infusion, response levels were similar in both groups (Figure 2A, Post; no significant differences between Groups , ).

Contrary to our prediction, intra-OFC muscimol infusions disrupted rather than enhanced further acquisition of responding relative to the saline group (Figure 2A, Infusion - Days 12-15; Significant Group x Day interaction , , but no main effect of Group , , or Day , ). Simple effects revealed significantly greater responding in the saline group on the last 2 days of infusions (Muscimol vs Saline: Day 12 , , Day 13 , , Day 14 , , Day 15 , ). Furthermore, the saline group increased responding across infusion days 12-15 (Saline: significant positive linear trend , ), whereas the muscimol group did not (Muscimol: no significant linear trend , ). Therefore, post-training inactivation of the OFC impaired acquisition.

Post-infusion, with function returned to the OFC, the group differences observed under drug infusion were no longer apparent, and both groups continued to acquire responding at similar levels (Figure 2A, Days 16-17; significant main effect of Day , , but no main effect of Group , , or Group x Day interaction , ). Therefore, the effect of OFC inactivation did not persist, which suggests that the disruption in acquisition following OFC inactivation might not have impaired learning *per se*.

**Experiment 3: Post-Training OFC lesions**

Next, we tested post-training lesions to rule out the possibility that the differences between pre- and post-training OFC manipulations were simply due to differences in the method of manipulation i.e. excitotoxic lesions vs inactivation using a GABA-A agonist. We trained a new cohort of animals on this simple Pavlovian cue-outcome task for 9 days, and then performed post-training excitotoxic or sham OFC lesions before continuing with acquisition (lesion extent depicted in Figure 2-figure supplement 3). Prior to surgery, animals acquired responding to the cue (Figure 2B, Pre-Surgery; significant main effect of Block , , but no main effect of Group , , or Group x Block interaction , ). After surgery, the sham group continued to acquire responding, but the lesion group did not (Figure 2B, Post-Surgery; significant Group x Block interaction , , but no main effect of Group , , or Day , ). Responding in the sham control group was significantly higher than the lesion group in the final block of 3 days (Block 4 , , Block 5 , , Block 6 , ). Furthermore, further acquisition post-surgery was completely abolished in the lesion group (Lesion: no linear trend over Blocks 4-6 , ), but continued in the sham control group (Sham: significant positive linear trend over Blocks 4-6 , . Therefore, both post-training lesions and inactivation of OFC function disrupted Pavlovian acquisition.

To facilitate comparisons between experiments, CS-PreCS response rates on Block 6 in the present experiment were Sham: M = 9.61, SD = 3.88, Lesion: M = 7.18, SD = 1.74 (Figure 2B). The terminal levels of responding in the sham group are similar to those of the saline group in Figure 2A, and the sham group in Figure 1A which used identical session parameters. This suggests that the present findings are also unlikely to be due to abnormally elevated levels of responding in the control groups in any one of these experiments.

**Experiment 4: OFC inactivation early in acquisition**

The findings presented thus far suggest that OFC inactivation temporarily suppressed acquisition performance, but not learning. However, it is also possible that the protocol is not sensitive enough to observe a learning deficit. For example, rates of responding were still quite high during muscimol inactivation (Figure 2A, days 12-15) and the subsequent recovery of responding (Figure 2A, days 16-17) could reflect rapid within-session learning in the muscimol group. Therefore, we tested the effect of OFC inactivation much earlier in the learning process, after only 4 days of acquisition (Figure 2C) when differences in learning should have greater impact. A new set of animals was implanted with bilateral cannulae (Figure 2-figure supplement 4) and then trained on a simple Pavlovian cue-outcome task (CS was a 10s house light).

Prior to drug infusions, all animals acquired responding to the cue (Figure 2C, Days 1-4; Significant main effect of Day , , but no main effect of Group , or Group x Day interaction , ). However, OFC inactivation during the next 5 days of conditioning significantly impaired acquisition in the muscimol group (Figure 2C, Days 5-9; Significant main effect of Group , , Day , , and Group x Day interaction , ). Responding in the muscimol group was significantly lower than the saline group on days 7-9 (Muscimol vs Saline: Day 5 , , Day 6 , , Day 7 , , Day 8 , , Day 9 , ). Again, this deficit was characterised by significant acquisition over days in the saline group that was abolished in the muscimol group (positive linear trend over days 5-9; Saline , , Muscimol , ). Finally, this reduction in responding persisted on day 10 when all rats were tested without infusion (Figure 2C, Day 10; , ). In contrast to OFC inactivation later in acquisition (Figure 2A), disrupting OFC activity early in learning suppressed performance which persisted when the OFC was active again. These findings suggest that OFC inactivation early in training disrupted acquisition learning rather than just behavioural performance.

**Experiment 5: OFC inactivation prior to associative blocking**

OFC inactivation during acquisition suppressed cue responding, but it is unclear if this reduction in behaviour is due to suppression of learning (Figure 2C) or behavioural performance (Figure 2A). This ambiguity is predominantly driven by the assumption that an animal’s response levels represent some monotonic function of acquired learning (Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Sutton & Barto, 1998; Wagner, 1981). To disambiguate learning from performance effects we employed an associative blocking design (Figure 3A). In a blocking experiment, first an animal is trained such that a cue (cue A) predicts an outcome (pellet). Next, A is presented in compound with a novel cue (cue B) which also leads to the same pellet outcome. If the animal has learned that cue A sufficiently predicts the pellet outcome already, then very little is learned about cue B i.e. learning about cue A blocks subsequent learning about cue B (Kamin, 1969). However, if learning about cue A is insufficient, then learning about cue B should not be blocked. We predicted that if OFC inactivation is disrupting learning, then OFC inactivation during initial learning about cue A should disrupt the blocking effect.

To test this prediction, a new set of animals was implanted with bilateral cannulae targeting the OFC and tested in a blocking procedure. During stage 1 of blocking (Figure 3B), all animals were given 10 days of acquisition training to cue A. OFC function was intact during the first 4 days of acquisition, and all animals began to acquire the cue A-outcome relationship (Days 1-4: significant main effect of Day , , but no effect of Group , or Group x Day interaction , ). All animals then received an additional 6 days of acquisition to cue A (Figure 3B, Days 5-10) following either intra-OFC infusions of muscimol or saline. Infusions of muscimol depressed overall responding relative to saline infusions (significant main effect of Group , , and Day , , but no Group x Day interaction , ). Importantly, on the final day (Day 10), responding in the muscimol group was significantly lower than the saline group (, ).

Next, animals were trained such that compounds AB and CD also predicted reward (Figure 3C, Stage 2), importantly OFC function was intact in all animals i.e. no infusions. Responding in both the saline and muscimol groups was initially lower to the novel compound CD than to AB (Significant Cue x Day interaction , , and main effect of Day , , but no other main effects or interactions with Group were significant, all remaining effects *F* < 1.91, *p* > .160; Cue AB vs CD: Day 12 , , Day 13 , , Day 14 , ). However, the pattern of means suggests that responding to compound AB in the muscimol group was similar to the novel compound CD on Day 12 (Figure 3C, Right - Day 12, Muscimol: AB vs CD , ), and lower than compound AB in the saline group (Figure 3C, Left - Day 12; Day 12, Saline: AB vs CD , ). Furthermore, Within-session changes over trials on Day 12 revealed rapid within-session acquisition to both compounds in both groups, but responding was significantly lower in the muscimol group at the start of the session (Figure 3 - figure supplement 2; First 2 trials, significant main effect of Group , , and Cue , , but no Group x Cue interaction , ). The lower responding to cue AB in the muscimol group suggests that acquisition to cue A was impaired following infusions in Stage 1 and this impairment persisted (albeit transiently) when test drug free in stage 2. Indeed, the levels of responding to compound AB in the muscimol group at the start of Day 12 (Figure 3 - figure supplement 2) are similar to levels of responding to the novel compound CD in the saline group. This would suggest that learning about cue A in the muscimol group was impaired in stage 1, and therefore cue A should not effectively block learning to cue B in stage 2.

At test both groups showed significant blocking of learning to cue B relative to the control cue D (Figure 3D; Significant main effect of Cue , , but no main effect of Group , , or Group x Cue interaction , ). This suggests that inactivation of the OFC significantly reduced behavioural performance but not learning to cue A in Stage 1, and this impairment transiently affected compound AB on Day 12 in the absence of OFC inactivation. Therefore, the impairments observed in our earlier findings (Figure 2A & C, post infusion) are unlikely to be due to impairments in learning. In addition to this, we rule out the possibility that the two groups used different attentional solutions to achieve a similar blocking result (Figure 3- figure supplement 3).

**Experiment 6: Competing response values**

One possible account of the impaired performance following OFC inactivation in the present study is an inability to potentiate learned behaviour based on the current value of the expected outcome. Specifically, OFC inactivation may disrupt the ability to potentiate performance based on the current motivational value of the outcome but leave intact knowledge about the predictive cue-outcome relationship (as suggested by intact associative blocking; Figure 3). While there are no overt alternative rewards in the simple Pavlovian task employed here, there is always an array of potential competing alternative behaviours available to the animal in the chamber e.g. exploration locomotion, grooming, rearing/orienting etc… These behaviours normally compete with the target magazine approach behaviour, and indeed the relative balance of these behaviours has been shown to develop over the course of Pavlovian acquisition (Boakes, 1977; Holland, 1977; Mackintosh, 1974; Shettleworth, 1994; Timberlake, 1994). Therefore, judgements about the current/relative value of an expected outcome are likely to incorporate the relative value of these competing behavioural options with factors such as the current motivation (e.g. hunger), reward magnitude (e.g. volume, concentration/number), reward probability.

Given this possibility, we hypothesized that post-training OFC inactivation impaired acquisition behaviour by disrupting the ability to modulate behaviour based on the relative value of these competing responses. Since the nature and value of these alternative behaviours is hard to quantify (and are likely subject-specific), we designed a task with a clearly defined alternative behaviour whose value could be experimentally manipulated. Specifically, we compared the strength of the standard magazine approach during a Pavlovian CS to the relative value of a competing unsignalled background rate of reward on the opposite side of the experimental chamber.

The task involved a Pavlovian cue-outcome procedure similar to those described above i.e. a 15s white-noise auditory stimulus predicted the delivery of a food pellet into a reward magazine at the front of the chamber (Figure 4- figure supplement 1). Independently of this CS-US contingency, the background rate of reward in the environment was also manipulated. Liquid sucrose reward was made available randomly (i.e. unsignalled delivery) throughout the session in a second reward magazine located at the back of the chamber. The probability of sucrose availability remained constant for blocks of 8 trials but changed within each session in a randomized order. Sucrose was presented in a dipper cup for 5s and then retracted so that this background reinforcement rate could only be determined by sampling the magazine location. This task provided a measure of a reward guided exploratory behaviour in the sucrose magazine, and Pavlovian behaviour to the pellet magazine driven by the expected value of the predicted outcome. Normally, animals will engage in a range of unmeasured and uncontrolled alternative behaviours in a testing chamber (e.g. exploration, orienting, grooming, etc…) that may compete with Pavlovian magazine approach. Here we provide a means to guide and control these alternative behaviours towards the sucrose magazine, and explicitly measure the integration of un-cued and cued expected value. We also confirmed that behaviour in this task was sensitive to changes in reward value/size and ruled out the influence of thirst on the valuation of the liquid sucrose reward.

An analysis of the separate magazine approach data (Figure 4A-B, D-E,G-H) was performed using a Shift (Acquisition, Thirst, 4xSucrose) x Period (PreCS, CS) x Magazine (Sucrose, Pellet) x Probability (Low, Medium, High) repeated measures ANOVA. First, we confirmed that the CS increased behaviour above PreCS levels at the Pellet but not the Sucrose magazine (Pellet magazine: PreCS vs CS , , Sucrose magazine: PreCS vs CS , ; Significant Period x Magazine interaction , , and main effect of Period , ). Therefore, Pavlovian behaviour directed at the Pellet magazine could be effectively dissociated from exploratory sampling behaviour directed at the Sucrose magazine.

Next, we confirmed that animals were sensitive to the background rate of sucrose delivery. The probability of sucrose modulated activity in the Sucrose and Pellet magazines in opposing directions (significant Magazine x Probability interaction , ). As the probability of sucrose availability increased, activity at the sucrose magazine increased whereas activity at the pellet magazine decreased (Sucrose magazine: Low vs Medium , , Low vs High , , Medium vs High , , Pellet magazine: Low vs Medium , , Low vs High , , Medium vs High , ). This shows that (1) Sucrose magazine behaviour was sensitive to the unsignalled changes in the background rate of sucrose availability, and (2) Pellet magazine behaviour reflects the trade-off between, and integration of, the expected Pavlovian pellet reward and the background rate of sucrose availability.

Finally, we confirmed that the motivation for the sucrose reward (Figure 4A-C) was not significantly driven by a motivational state of thirst (Figure 4D-F), and that the behaviours in this task were sensitive to changes in reward value (i.e. a 4-fold increase in sucrose volume; Figure 4G-I). Increasing sucrose reward size selectively increased behaviour at the Sucrose but not the Pellet magazine, whereas a motivational state of thirst did not significantly affect behaviour at either magazine (significant Magazine x Shift interaction , ; Sucrose magazine: Acquisition vs Thirst , 4xSucrose vs Acquisition , 4xSucrose vs Thirst , ; Pellet magazine: Acquisition vs Thirst , 4xSucrose vs Acquisition , 4xSucrose vs Thirst , ). No other meaningful effects were significant (significant Period x Probability interaction , ; all remaining effects *F* < 3.421, *p* > .06).

***The effect of OFC inactivation on updating relative expected value***

Next, these animals were implanted with bilateral cannulae to assess the role of the OFC in updating relative expected value. Animals were tested following muscimol or saline infusions (within-subjects, counterbalanced order) on a modified task that went from low to high probability only. This provided a shorter session to ensure the efficacy of muscimol throughout the session, and the fixed probability order minimized any potential for satiety to confound behaviour later in the session.

We hypothesized that OFC function was necessary for flexibly controlling the strength of Pavlovian anticipatory behaviour relative to the current value of the outcome and alternative behavioural options in the environment. Specifically, in the present task, following muscimol inactivation, we predicted that behaviour at the Sucrose magazine would remain sensitive to the probability of sucrose whereas behaviour at the Pellet magazine during the CS would no longer be modulated by probability of sucrose i.e. behaviour controlled by Pavlovian expected value would become inflexible.

First, during the PreCS period, all animals were able to detect changes in the probability of the unsignalled sucrose reward, and appropriately update Sucrose magazine approach behaviour (Figure 5A-B; PreCS period: Significant Probability x Magazine interaction , , and main effects of Probability , , and Magazine , ; No significant main effect or interaction with Drug). Specifically, during the PreCS period, anticipatory approach to the Sucrose magazine in both drug conditions increased with the probability of sucrose (Sucrose magazine: Low vs High probability , ), whereas approach to the Pellet magazine did not change (Pellet magazine: Low vs High probability , ).

Next, during the CS period, OFC inactivation significantly disrupted activity at the Pellet magazine but left activity at the Sucrose magazine intact (Figure 5C-D; CS period: significant Drug x Magazine x Probability 3-way interaction , ). Activity at the Sucrose magazine was not affected by muscimol infusions (Sucrose magazine: no significant main effect of Drug , , or Drug x Probability interaction , ), whereas activity at the Pellet magazine was significantly disrupted by muscimol infusions (Pellet magazine: significant Drug x Probability interaction , ). Simple effects revealed that Pellet magazine responding was lower after muscimol than saline infusions for the low probability of dipper reward (, ) but did not differ between infusions for the high probability of dipper reward (, ).

Finally, comparing the analysis of the PreCS and CS periods using a full Drug (Saline, Muscimol) x Period (PreCS, CS) x Magazine (Sucrose, Pellet) x Probability (Low, High) repeated measures ANOVA confirmed our prediction that the muscimol deficit was specific to the Pellet magazine in the CS period (Figure 5; Significant Drug x Period x Magazine x Probability 4-way interaction , ). OFC inactivation selectively disrupted the ability to modulate behaviour based on integrating Pavlovian expected values with the current rate of alternative rewards i.e. the current subjective value of the predicted reward.

**Discussion**

The present studies tested the hypothesis that the rodent lateral OFC is not necessary for Pavlovian acquisition. Here we show that OFC lesions and inactivation significantly affects Pavlovian acquisition in a simple single CS-US procedure. Furthermore, we found a dissociation between pre- and post-training OFC manipulations on Pavlovian acquisition such that pre-training OFC lesions enhance, whereas post-training lesions and inactivation impairs acquisition behaviour. Next, using an associative blocking design, we tested whether impaired behaviour following post-training OFC inactivation reflects a disruption of learning or behavioural control. OFC inactivation did not disrupt the underlying learning about the predictive CS-US relationship, and instead disrupted the appropriate control of anticipatory behaviour to the CS. Finally, we assessed whether this impaired behavioural control reflects an inability to update the current value of the Pavlovian CS relative to the value of alternative behavioural options. Indeed, OFC inactivation selectively disrupted the flexible control of Pavlovian CS approach behaviour when its relative value changed but did not disrupt sensitivity to the value of alternative/non-Pavlovian behaviours in the environment.

**Lateral OFC is necessary for simple Pavlovian acquisition**

The significant role of the OFC in Pavlovian acquisition in the present studies is surprising since OFC lesions and inactivation have consistently been reported to have no effect on acquisition in rats (e.g. Burke et al., 2008; Gallagher et al., 1999; Schoenbaum et al., 2002; Stalnaker et al., 2007), unless there are complex cue or outcome specific task demands (e.g. Ramirez & Savage, 2007). For example, in tasks involving simple single Pavlovian CS-US procedures and pre-training OFC lesions, performance often does not reach asymptote (e.g. after 9 days, (Gallagher et al., 1999)) before proceeding to a new stage of the experiment. In Experiment 1, we did not observe any significant effects of OFC lesions on acquisition until around 15-21 days of acquisition. However, after extended training Schoenbaum et al (2003) have reported significant effects of OFC lesions on acquisition in a simple cue-outcome go-nogo task when looking at response latencies, but not on trials-to-criterion. Therefore, the effects of pretraining lesions may not have been observed previously due to task specific parameters such as the length of training and the sensitivity of the response measures.

Pretraining OFC lesions have been shown to disrupt Pavlovian acquisition in sign-tracking procedures in which lever insertion is used as the CS (Chudasama & Robbins, 2003). Focal lateral OFC lesions significantly impair sign-tracking behaviour (i.e. engaging with the lever cue), and bias behaviour towards goal-tacking (i.e. approaching the magazine) (Panayi & Killcross, 2018). This is consistent with the findings of Experiment 1 that OFC lesions enhanced behaviour focused toward the magazine, and suggest that this focus may come at the expense of some alternative behaviour that sham lesion animals are engaging in.

In contrast to pre-training lesions, post-training OFC inactivation/lesions normally coincide with changes in experimental phase and continued acquisition is not assessed. In tasks in which OFC inactivation coincides with a change in experimental stage, the effects of OFC inactivation are consistent with an impairment in subsequent acquisition (2009). For example, Burke et al (2009) found that post-training OFC inactivation impaired acquisition to a Pavlovian CS in reversal task. Similarly, Takahashi et al (2009) found that OFC inactivation during a Pavlovian over-expectation task disrupted new learning. Therefore, the robust effect of impaired acquisition following post-training OFC inactivation that we report is consistent with impaired subsequent acquisition in tasks with more complex manipulations.

**Lateral OFC is not necessary for learning the predictive CS-US relationship**

Post-training OFC inactivation significantly impaired acquisition behaviour (Experiment 2), and this disruption was more profound when inactivation occurred earlier in training and persisted even after OFC function returned (Experiment 4). This strongly suggests that learning about the CS-US relationship was disrupted. The idea that the OFC is involved in learning is also consistent with a role for the OFC in the representation of expected values (Burke et al., 2008; Schoenbaum et al., 2011; Stalnaker et al., 2018) which influence mid-brain dopaminergic prediction errors (Takahashi et al., 2009, 2011), known to be necessary for Pavlovian learning (Sharpe et al., 2017a; Steinberg et al., 2013).

Unexpectedly, the impaired acquisition we observed (Experiment 2, 3, & 4) did not disrupt the ability of the CS to block learning about a novel cue (Experiment 5, Figure 3), despite significantly impaired performance post-inactivation (Figure 3-figure supplement 2; Muscimol AB is as low as Saline CD which does not show evidence of blocking). This is surprising given that in some Pavlovian learning contexts, levels of behavioural expression can dictate the extent to which learning occurs (Delamater, 2004). This finding highlights the importance of using multiple tests of learning (Rescorla, 2002a, 2002b) to assess disrupted acquisition effects.

Intact blocking despite impaired acquisition behaviour suggests that OFC inactivation did not disrupt the underlying learning about the associative strength of the CS-US relationship. Associative blocking is often used to assess the role of prediction-error based learning (Sharpe et al., 2017b; e.g. Steinberg et al., 2013), suggesting that the OFC is not necessary for this aspect of Pavlovian learning. This distinction suggests that the learned value of a Pavlovian CS-US association might be independent of the current expected or subjective value of expected reward. Informally, learning whether an outcome will be delivered might reasonably be separate from learning the subjective value or identity of that outcome (Delamater, 2007; Delamater & Oakeshott, 2007; McDannald et al., 2011; Zhou et al., 2019).

**Lateral OFC is necessary for flexible value-based Pavlovian behavioural control**

One challenge raised by the present findings is, if the OFC is not necessary for CS-US contingency learning, but is necessary for learning about the specific identity of expected outcomes and their value, why do we observe an effect in a simple single CS-US learning procedure? In this simple task, the value and identity of the US stays constant, and there is only a single unambiguous CS. We reasoned that, even in Putatively simple tasks, there are a number of unconstrained alternative behaviours that a rat can engage in which can compete with the target magazine approach behaviour. These behaviours are likely to be under the control of different behavioural systems (e.g. grooming), and the relative value of a behaviour (presumably determining its eligibility for dominating behavioural performance at any given moment) is likely to be highly variable within- and between-subjects (Shettleworth, 1994; Timberlake, 1994). Indeed, the competition between these alternative behaviours might be most prominent in simple tasks with low attentional/cognitive demands.

We hypothesized that OFC inactivation disrupted a value-based decision process involved in comparing the relative value of these alternative behaviours with the expected value of the US. We created a task (Experiment 6) in which we could direct and measure these normally unconstrained alternative behaviours and manipulate their relative value by providing unsignalled probabilistic reward similar to patch foraging tasks (Krebs, 1977; Stephens & Krebs, 1986). Behaviour to this reward site rapidly tracked the unsignalled reward rate changes, and effectively competed for Pavlovian magazine approach to a CS in a value dependent manner. As predicted, OFC inactivation disrupted the integration of relative value into the Pavlovian approach behaviour but left the valuation and control of the alternative behaviour intact.

This suggests that the lateral OFC is necessary for value based behavioural flexibility of Pavlovian behaviours. One important alternative account that cannot be ruled out is that this reflects an inability to integrate across USs of different sensory properties i.e. grain pellets and sucrose liquid (Burke et al., 2008; Delamater, 2007; Gardner et al., 2017; McDannald et al., 2011; Schoenbaum et al., 2009). However, the effect of OFC inactivation was specific to integrating either value or identity information to flexibly control Pavlovian behaviours. Lateral OFC lesions have also previously been found not to affect the acquisition or value-based control of instrumental action-outcome behaviours (Ostlund & Balleine, 2007; Panayi & Killcross, 2018). This suggests that the lateral OFC is specifically necessary for the control of Pavlovian behaviours, or tasks in which Pavlovian CS-US contingencies dominate performance (e.g. West et al., 2013).

**Pre- vs post-training effects**

The dissociable and opposite effects of pre- and post-training OFC lesions/inactivation on acquisition were surprising and rule out a simple account of OFC dysfunction in terms of prediction-error based learning impairments (Figure 2-figure supplement 1). One possibility is that pretraining lesions result in compensatory function such that learning is supported by other neural systems. In contrast, post-training lesions and inactivation disrupts learning/behaviour that has been acquired in an OFC dependent manner. This argument has been proposed when only pre-training OFC lesions (Boulougouris et al., 2007; Boulougouris & Robbins, 2009), or only post-training OFC lesions disrupt behaviour (Balleine et al., 2011; Ostlund & Balleine, 2007). We will also consider two alternative accounts of pre- vs post-training OFC lesion differences based on theoretical accounts of OFC function, sensory-specific outcome expectancy and latent state theories. Note that these theories do not predict an effect of OFC lesions on simple Pavlovian acquisition *a priori*, and therefore require additional assumptions to account for the present data.

From an associative learning framework, even putatively “simple” single cue-outcome Pavlovian learning can involve a number of different psychological/behavioural processes (Dickinson, 1980; Hall, 2002; Holland, 1977; Konorski, 1967; Mackintosh, 1974; Rescorla, 1988). Take for example a 10s light cue that reliably predicts the delivery of a pellet reward. A rat can learn that the cue predicts the sensory-specific properties of the outcome (e.g. taste, texture, sweetness, colour, size, location etc...), or the general motivational value of that reward, or simply develop a stimulus-response habit to approach the reward location when the cue is presented. Indeed, there is experimental evidence for these multiple aspects of learning occurring during Pavlovian conditioning (for review see Delamater & Oakeshott, 2007). It is possible that pretraining OFC lesions disrupt the balance of these different aspects of Pavlovian learning and behavior (Burke et al., 2007; Delamater, 2007).

If the OFC is necessary for the representation of the sensory specific properties of expected outcomes, then OFC lesions might allow a stimulus-response habit system to dominate behavioural control. Following pre-training lesions, this may lead to an unconstrained habit learning system (Coutureau & Killcross, 2003; Dickinson, 1985; Dolan & Dayan, 2013; Killcross & Coutureau, 2003) that is not necessarily bounded by the current value of the outcome, and overly sensitive to current general motivational states (e.g. overall hunger levels; Figure 1B) of the organism. However, once initial learning occurs with an intact OFC, the encoding of the identity of the expected outcome is likely to have occurred (e.g. Delamater & Holland, 2008). Subsequently, a post-training lesion or inactivation of the OFC is likely to affect the subsequent updating of this information. Here we propose that the impaired acquisition behaviour we observed following post-training inactivation reflects an inability to update the current motivational value of the specific outcome that is expected.

The latent state representation account of the OFC might also be able to account for the differences observed dissociation between pre- and post-training OFC lesions on acquisition. Computational models (e.g. Wilson et al., 2014) often assume, for simplicity, that in a simple single cue-outcome procedure, the cue state (e.g. “light on”) is stable throughout acquisition. Given that the same cue is presented, and it always leads to the pellet outcome, this stable representation is a reasonable assumption. However, it is also likely that early in acquisition this state representation is not yet stable in healthy control animals (Niv, 2019). How can the animal be certain that the light cue, the testing chamber context, or the reward pellet that they see on each trial is identical to the trials they have already experienced within the session, and from previous days? The subjective experience of these states and their physical features is very likely to be different within- and between-sessions e.g. the ambient noises, odours, temperature of the context, the location and intensity of the light cue based on where the rat happens to be located when it turns on, and the gradual onset of sensory specific satiety to the pellet etc... Informally, how does the rat know that this light is the same light that they saw at the start of the session, or the day before? The perception and recognition of these states is therefore subject to differences in variables such as generalization, confidence, and certainty.

Paradoxically, in a simple and stable cue-outcome training procedure, pre-training OFC lesions may result in an accurate, but inflexible, representation of these simple task states quite rapidly. In this stable and simple training context this could lead to enhanced Pavlovian acquisition. However, in a task with multiple or uncertain cue-outcome contingencies pretraining OFC lesions might impair acquisition (Stolyarova & Izquierdo, 2017; Walton et al., 2010). However, post-training inactivation of the OFC would disrupt the ability to update already established state representations at whatever stage of certainty/stability that they have currently achieved. In the stable single cue-outcome learning situation employed in the present studies, this would result in disruption of further acquisition. Again, in a task with interference from multiple cue-outcome relationships, post-training lesions might improve performance.

**Conclusion**

Here we show that the rodent lateral OFC is involved in Pavlovian acquisition learning process in an experience dependent manner. Once initial learning has taken place, the lateral OFC appears to be specifically necessary for updating the current value of Pavlovian behaviours driven by expected outcome value. While the OFC has often been found not to be necessary for initial acquisition learning, recently there have been reports that simple Pavlovian acquisition is significantly impaired rather than enhanced following optogenetic inhibition of OFC function in head fixed mice (Namboodiri et al., 2019; Wang et al., 2020), in a manner that does not depend on VTA prediction error signaling. In contrast to our results, these studies target more ventral and medial OFC, which is likely to be an important anatomical distinction given the emerging evidence of functional heterogeneity within the OFC (Barreiros et al., 2020, 2021; Bradfield & Hart, 2020; Sharpe et al., 2015). Indeed, there appears to be dissociable but complementary roles of the medial and lateral OFC such that lateral OFC lesions disrupt Pavlovian whereas medial OFC lesions disrupt instrumental behavioural control (Bradfield et al., 2015, 2018; Gardner et al., 2017, 2018; McDannald et al., 2011; Ostlund & Balleine, 2007; Panayi & Killcross, 2018). This suggests that the OFC as a whole is engaged in the learning and flexible updating of value based behaviours, but within the orbital subregions this process appears to be remarkably specialized for distinct types of behaviour and learning.

**References**

Balleine, B. W., Leung, B. K., & Ostlund, S. B. (2011). The orbitofrontal cortex, predicted value, and choice. *Ann N Y Acad Sci*, *1239*, 43–50. https://doi.org/10.1111/j.1749-6632.2011.06270.x

Barreiros, I. V, Ishii, H., Walton, M. E., & Panayi, M. C. (2021). Defining an orbitofrontal compass: functional and anatomical heterogeneity across anterior-posterior and medial-lateral axes. *Behavioral Neuroscience*, *under review*.

Barreiros, I. V, Panayi, M. C., & Walton, M. E. (2020). Organisation of afferents along the anterior-posterior and medial-lateral axes of the rat OFC. *BioRxiv*, 1–44. https://doi.org/https://doi.org/10.1101/2020.08.28.272591

Baxter, M. G., Parker, A., Lindner, C. C., Izquierdo, A. D., & Murray, E. A. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *Journal of Neuroscience*, *20*(11), 4311–4319. http://www.jneurosci.org/content/20/11/4311.full.pdf

Boakes, R. A. (1977). Performance on learning to associate a stimulus with positive reinforcement. In H. Davis & H. M. B. Hurwitz (Eds.), *Operant-Pavlovian interactions* (pp. 67–97). L. Erlbaum Associates.

Boulougouris, V., Dalley, J. W., & Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research*, *179*(2), 219–228. https://doi.org/10.1016/j.bbr.2007.02.005

Boulougouris, V., & Robbins, T. W. (2009). Pre-surgical training ameliorates orbitofrontal-mediated impairments in spatial reversal learning. *Behavioural Brain Research*, *197*(2), 469–475. https://doi.org/10.1016/j.bbr.2008.10.005

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*(6), 1268–1280. https://doi.org/10.1016/j.neuron.2015.10.044

Bradfield, L. A., & Hart, G. (2020). Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space. *Neuroscience & Biobehavioral Reviews*, *108*, 287–294. https://doi.org/10.1016/J.NEUBIOREV.2019.11.009

Bradfield, L. A., Hart, G., & Balleine, B. W. (2018). Inferring action-dependent outcome representations depends on anterior but not posterior medial orbitofrontal cortex. *Neurobiology of Learning and Memory*, *155*(May), 463–473. https://doi.org/10.1016/j.nlm.2018.09.008

Burke, K. A., Franz, T. M., Miller, D. N., & Schoenbaum, G. (2007). Conditioned reinforcement can be mediated by either outcome-specific or general affective representations. *Frontiers in Integrative Neuroscience*, *1*, 2. https://doi.org/10.3389/neuro.07.002.2007

Burke, K. A., Franz, T. M., Miller, D. N., & Schoenbaum, G. (2008). The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature*, *454*(7202), 340-U45. https://doi.org/Doi 10.1038/Nature06993

Burke, K. A., Takahashi, Y. K., Correll, J., Brown, P. L., & Schoenbaum, G. (2009). Orbitofrontal inactivation impairs reversal of Pavlovian learning by interfering with “disinhibition” of responding for previously unrewarded cues. *European Journal of Neuroscience*, *30*(10), 1941–1946. https://doi.org/DOI 10.1111/j.1460-9568.2009.06992.x

Butter, C. M. (1969). Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in Macaca mulatta. *Physiol. Behav*, *4*, 163–171.

Chudasama, Y., & Robbins, T. W. (2003). Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *Journal of Neuroscience*, *23*(25), 8771–8780. https://doi.org/23/25/8771 [pii]

Coutureau, E., & Killcross, A. S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behavioural Brain Research*, *146*(1–2), 167–174. http://www.sciencedirect.com/science/article/pii/S0166432803003498

Delamater, A. R. (2004). Experimental extinction in Pavlovian conditioning: Behavioural and neuroscience perspectives. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*, *57*(2), 97–132. https://doi.org/Doi 10.1080/02724990344000097

Delamater, A. R. (2007). The role of the orbitofrontal cortex in sensory-specific encoding of associations in Pavlovian and instrumental conditioning. In G. Schoenbaum, J. A. Gottfried, E. A. Murray, & S. J. Ramus (Eds.), *Linking Affect to Action: Critical Contributions of the Orbitofrontal Cortex* (Vol. 1121, pp. 152–173). Blackwell Publishing. https://doi.org/10.1196/annals.1401.030

Delamater, A. R., & Holland, P. C. (2008). The influence of CS-US interval on several different indices of learning in appetitive conditioning. *Journal of Experimental Psychology-Animal Behavior Processes*, *34*(2), 202–222. https://doi.org/Doi 10.1037/0097-7403.34.2.202

Delamater, A. R., & Oakeshott, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Annals of the New York Academy of Sciences*. https://doi.org/10.1196/annals.1390.008

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*, *380*(6569), 69–72. https://doi.org/10.1038/380069a0

Dickinson, A. (1980). Contemporary animal learning theory. In J. Gray (Ed.), *Problems in the behavioural sciences*. Cambridge University Press.

Dickinson, A. (1985). Actions and Habits : The Development of Behavioural Autonomy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *308*(1135), 67–78. http://www.jstor.org/stable/2396284

Dickinson, A., & Balleine, B. W. (2002). The Role of Learning in the Operation of Motivational Systems. In *Stevens’ Handbook of Experimental Psychology*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471214426.pas0312

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society B-Biological Sciences*, *278*(1718), 2553–2561. https://doi.org/DOI 10.1098/rspb.2011.0836

Gallagher, M., McMahan, R. W., & Schoenbaum, G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *Journal of Neuroscience*, *19*(15), 6610–6614. http://www.jneurosci.org/cgi/reprint/19/15/6610.pdf

Gardner, M. P. H., Conroy, J. C., Sanchez, D. C., Zhou, J., & Schoenbaum, G. (2019). Real-Time Value Integration during Economic Choice Is Regulated by Orbitofrontal Cortex. *Current Biology*. https://doi.org/10.1016/j.cub.2019.10.058

Gardner, M. P. H., Conroy, J. C., Styer, C. V, Huynh, T., Whitaker, L. R., & Schoenbaum, G. (2018). Medial orbitofrontal inactivation does not affect economic choice. *ELife*, *7*. https://doi.org/10.7554/eLife.38963

Gardner, M. P. H., Conroy, J. S., Shaham, M. H., Styer, C. V., & Schoenbaum, G. (2017). Lateral Orbitofrontal Inactivation Dissociates Devaluation-Sensitive Behavior and Economic Choice. *Neuron*, *0*(0). https://doi.org/10.1016/j.neuron.2017.10.026

Hall, G. (2002). Associative structures in Pavlovian and instrumental conditioning. In C. R. Gallistel (Ed.), *Steven’s handbook of experimental psychology* (Vol. 3, pp. 1–45). John Wiley & Sons.

Holland, P. C. (1977). Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *J Exp Psychol Anim Behav Process*, *3*(1), 77–104. http://www.ncbi.nlm.nih.gov/pubmed/845545

Iversen, S. D., & Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research*, *11*(4), 376–386. http://www.ncbi.nlm.nih.gov/pubmed/4993199

Izquierdo, A. (2017). Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *37*(44), 10529–10540. https://doi.org/10.1523/JNEUROSCI.1678-17.2017

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–96). Appleton-Century-Crofts.

Killcross, A. S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, *13*(4), 400–408. http://cercor.oxfordjournals.org/cgi/reprint/13/4/400.pdf

Klein-Flugge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J., & Behrens, T. E. (2013). Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *Journal of Neuroscience*, *33*, 3202–3211. http://www.jneurosci.org/content/33/7/3202.full.pdf

Konorski, J. (1967). *Integrative activity of the brain; an interdisciplinary approach*. University of Chicago Press.

Krebs, J. (1977). Optimal foraging: theory and experiment. *Nature*, *268*(5621), 583–584. https://doi.org/10.1038/268583a0

Kringelbach, M. L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience*, *6*(9), 691–702. https://doi.org/Doi 10.1038/Nrn1747

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). emmeans: estimated marginal means. R package version 1.4. 4. In *The American Statistician*. https://doi.org/10.1080/00031305.1980.10483031

LePelley, M. E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *Quarterly Journal of Experimental Psychology*, *57B*, 192–243.

Mackintosh, N. J. (1974). *The psychology of animal learning*. Academic Press.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol Rev*, *82*(4), 279–298. https://doi.org/10.1037/h0076778

McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *Journal of Neuroscience*, *31*, 2700–2705. http://www.jneurosci.org/content/31/7/2700.full.pdf

McDannald, M. A., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Lesions of orbitofrontal cortex impair rats’ differential outcome expectancy learning but not conditioned stimulus-potentiated feeding. *Journal of Neuroscience*, *25*(18), 4626–4632. https://doi.org/25/18/4626 [pii] 10.1523/JNEUROSCI.5301-04.2005

Murray, E. A., O’Doherty, J. P., & Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *Journal of Neuroscience*, *27*(31), 8166–8169. https://doi.org/10.1523/JNEUROSCI.1556-07.2007

Murray, E. A., & Rudebeck, P. H. (2018). Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. *Nature Reviews Neuroscience*, *19*(7), 404–417. https://doi.org/10.1038/s41583-018-0013-4

Namboodiri, V. M. K., Otis, J. M., van Heeswijk, K., Voets, E. S., Alghorazi, R. A., Rodriguez-Romaguera, J., Mihalas, S., & Stuber, G. D. (2019). Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. *Nature Neuroscience*, 1. https://doi.org/10.1038/s41593-019-0408-1

Nasser, H. M., Calu, D. J., Schoenbaum, G., & Sharpe, M. J. (2017). The Dopamine Prediction Error: Contributions to Associative Models of Reward Learning. *Frontiers in Psychology*, *8*, 244. https://doi.org/10.3389/fpsyg.2017.00244

Niv, Y. (2019). Learning task-state representations. *Nature Neuroscience*, *22*(10), 1544–1553. https://doi.org/10.1038/s41593-019-0470-8

Ogawa, M., van der Meer, M. A. A., Esber, G. R., Cerri, D. H., Stalnaker, T. A., & Schoenbaum, G. (2013). Risk-responsive orbitofrontal neurons track acquired salience. *Neuron*, *77*(2), 251–258. https://doi.org/10.1016/j.neuron.2012.11.006

Ostlund, S. B., & Balleine, B. W. (2007). Orbitofrontal cortex mediates outcome encoding in pavlovian but not instrumental conditioning. *Journal of Neuroscience*, *27*(18), 4819–4825. https://doi.org/Doi 10.1523/Jneurosci.5443-06.2007

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience*, *29*, 14004–14014. http://www.jneurosci.org/content/29/44/14004.full.pdf

Panayi, M. C., & Killcross, S. (2018). Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. *ELife*, *7*. https://doi.org/10.7554/eLife.37357.001

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev*, *87*(6), 532–552. http://www.ncbi.nlm.nih.gov/pubmed/7443916

Pears, A., Parkinson, J. A., Hopewell, L., Everitt, B. J., & Roberts, A. C. (2003). Lesions of the orbitofrontal but not medial prefrontal cortex disrupt conditioned reinforcement in primates. *Journal of Neuroscience*, *23*(35), 11189–11201. https://doi.org/23/35/11189 [pii]

Pickens, C. L., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behav Neurosci*, *119*(1), 317–322. https://doi.org/2005-01705-030 [pii]10.1037/0735-7044.119.1.317

Pickens, C. L., Saddoris, M. P., Setlow, B., Gallagher, M., Holland, P. C., & Schoenbaum, G. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *Journal of Neuroscience*, *23*(35), 11078–11084. http://www.jneurosci.org/content/23/35/11078.full.pdf

R Core Team (2020). (2020). R: A language and environment for statistical computing. In *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*.

Ramirez, D. R., & Savage, L. M. (2007). Differential involvement of the basolateral amygdala, orbitofrontal cortex, and nucleus accumbens core in the acquisition and use of reward expectancies. *Behav Neurosci*, *121*(5), 896–906. https://doi.org/10.1037/0735-7044.121.5.896

Rescorla, R. A. (1988). Pavlovian Conditioning - Its Not What You Think It Is. *American Psychologist*, *43*(3), 151–160. https://doi.org/Doi 10.1037/0003-066x.43.3.151

Rescorla, R. A. (2002a). Comparison of the rates of associative change during acquisition and extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(4), 406–415. https://doi.org/10.1037/0097-7403.28.4.406

Rescorla, R. A. (2002b). Savings tests: Separating differences in rate of learning from differences in initial levels. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(4), 369–377. https://doi.org/10.1037/0097-7403.28.4.369

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditiong: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokesy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64–99). Appleton Century Crofts.

Rudebeck, P. H., & Murray, E. A. (2014). The Orbitofrontal Oracle: Cortical Mechanisms for the Prediction and Evaluation of Specific Behavioral Outcomes. *Neuron*, *84*(6), 1143–1156. https://doi.org/10.1016/j.neuron.2014.10.049

Sadacca, B. F., Wied, H. M., Lopatina, N., Saini, G. K., Nemirovsky, D., & Schoenbaum, G. (2018). Orbitofrontal neurons signal sensory associations underlying model-based inference in a sensory preconditioning task. *ELife*, *7*, e30373. https://doi.org/10.7554/eLife.30373

Schoenbaum, G., Nugent, S. L., Saddoris, M. P., & Setlow, B. (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport*, *13*(6), 885–890. http://graphics.tx.ovid.com/ovftpdfs/FPDDMCOKLECKGD00/fs047/ovft/live/gv024/00001756/00001756-200205070-00030.pdf

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*, *10*(12), 885–892. https://doi.org/Doi 10.1038/Nrn2753

Schoenbaum, G., Setlow, B., Nugent, S. L., Saddoris, M. P., & Gallagher, M. (2003). Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learning & Memory*, *10*(2), 129–140. https://doi.org/Doi 10.1101/Lm.55203

Schoenbaum, G., Takahashi, Y. K., Liu, T., & McDannald, M. A. (2011). Does the orbitofrontal cortex signal value? *Ann N Y Acad Sci*, *1239*(1), 87–99. https://doi.org/10.1111/j.1749-6632.2011.06210.x

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27.

Sharpe, M. J., Chang, C. Y., Liu, M. A., Batchelor, H. M., Mueller, L. E., Jones, J. L., Niv, Y., & Schoenbaum, G. (2017a). Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience*. https://doi.org/10.1038/nn.4538

Sharpe, M. J., Chang, C. Y., Liu, M. A., Batchelor, H. M., Mueller, L. E., Jones, J. L., Niv, Y., & Schoenbaum, G. (2017b). Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience*, *20*(5), 735–742. https://doi.org/10.1038/nn.4538

Sharpe, M. J., Wikenheiser, A. M., Niv, Y., & Schoenbaum, G. (2015). The State of the Orbitofrontal Cortex. *Neuron*, *88*(6), 1075–1077. https://doi.org/10.1016/j.neuron.2015.12.004

Shettleworth, S. J. (1994). Commentary: What are behavior systems and what use are they? *Psychonomic Bulletin & Review*, *1*(4), 451–456. https://doi.org/10.3758/BF03210949

Singmann, H., Bolker, B., Westfall, J., Aust, F., Højsgaard, S., Fox, J., Lawrence, M. A., Mertens, U., & Love, J. (2020). afex: analysis of factorial experiments. *R Package Version 0.27-2*. https://cran.r-project.org/package=afex

Stalnaker, T. A., Cooch, N. K., McDannald, M. A., Liu, T. L., Wied, H., Schoenbaum, G., Tzu-Lan, L., Wied, H., & Schoenbaum, G. (2014). Orbitofrontal neurons infer the value and identity of predicted outcomes. *Nat Commun*, *5*, 3926. https://doi.org/10.1038/ncomms4926

Stalnaker, T. A., Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. In *Nature Neuroscience*. https://doi.org/10.1038/nn.3982

Stalnaker, T. A., Franz, T. M., Singh, T., & Schoenbaum, G. (2007). Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. *Neuron*, *54*(1), 51–58. https://doi.org/DOI 10.1016/j.neuron.2007.02.014

Stalnaker, T. A., Liu, T.-L., Takahashi, Y. K., & Schoenbaum, G. (2018). Orbitofrontal neurons signal reward predictions, not reward prediction errors. *Neurobiology of Learning and Memory*, *153*, 137–143. https://doi.org/10.1016/J.NLM.2018.01.013

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973. https://doi.org/10.1038/nn.3413

Stephens, D. N., & Krebs, J. (1986). *Foraging Theory*. Princeton University Press.

Stolyarova, A., & Izquierdo, A. (2017). Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. *ELife*, *6*, e27483. https://doi.org/10.7554/eLife.27483

Sutton, R. S., & Barto, A. G. (1998). *Reinforccement Learning: An introduction*. The MIT Press.

Takahashi, Y. K., Chang, C. Y., Lucantonio, F., Haney, R. Z., Berg, B. A., Yau, H.-J., Bonci, A., & Schoenbaum, G. (2013). Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron*, *80*, 507–518. http://ac.els-cdn.com/S0896627313007198/1-s2.0-S0896627313007198-main.pdf?\_tid=97ea45dc-cbcc-11e4-9c0e-00000aacb362&acdnat=1426504211\_80b43da207445d70382dd3274f8c4f11

Takahashi, Y. K., Roesch, M. R., Stalnaker, T. A., Haney, R. Z., Caiu, D. J., Taylor, A. R., Burke, K. A., Schoenbaum, G., Calu, D. J., Taylor, A. R., Burke, K. A., & Schoenbaum, G. (2009). The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for Learning from Unexpected Outcomes. *Neuron*, *62*(2), 269–280. https://doi.org/DOI 10.1016/j.neuron.2009.03.005

Takahashi, Y. K., Roesch, M. R., Wilson, R. C., Toreson, K., O’Donnell, P., Niv, Y., & Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature Neuroscience*, *14*(12), 1590–1597. https://doi.org/10.1038/nn.2957

Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychon Bull Rev*, *1*(4), 405–420. https://doi.org/10.3758/BF03210945

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear (Ed.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Erlbaum.

Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*, *65*(6), 927–939. https://doi.org/10.1016/j.neuron.2010.02.027

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. *Ann N Y Acad Sci*, *1239*, 14–24. http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.2011.06257.x/abstract

Wang, P. Y., Boboila, C., Chin, M., Stein, N. P., Abbott, L. F., Axel, R., Wang, P. Y., Boboila, C., Chin, M., Higashi-howard, A., Shamash, P., & Wu, Z. (2020). *Article Transient and Persistent Representations of Odor Value in Prefrontal Cortex Article Transient and Persistent Representations of Odor Value in Prefrontal Cortex*. 1–16.

West, E. A., DesJardin, J. T., Gale, K., & Malkova, L. (2011). Transient Inactivation of Orbitofrontal Cortex Blocks Reinforcer Devaluation in Macaques. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(42), 15128–15135. https://doi.org/10.1523/JNEUROSCI.3295-11.2011

West, E. A., Forcelli, P. A., McCue, D. L., & Malkova, L. (2013). Differential effects of serotonin-specific and excitotoxic lesions of OFC on conditioned reinforcer devaluation and extinction in rats. *Behavioural Brain Research*, *246*, 10–14. https://doi.org/10.1016/j.bbr.2013.02.027

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, *81*(2), 267–279. https://doi.org/10.1016/j.neuron.2013.11.005

Zhou, J., Gardner, M. P. H., Stalnaker, T. A., Ramus, S. J., Wikenheiser, A. M., Niv, Y., & Schoenbaum, G. (2019). Rat Orbitofrontal Ensemble Activity Contains Multiplexed but Dissociable Representations of Value and Task Structure in an Odor Sequence Task. *Current Biology*. https://doi.org/10.1016/j.cub.2019.01.048

**Methods and materials**

*Animals.* Subjects were male Long Evans rats (Monash Animal Services, Gippsland, Victoria, Australia) approximately 4 months old. Rats were housed four per cage in ventilated Plexiglass cages in a temperature regulated (22 ± 1­°C) and light regulated (12h light/dark cycle, lights on at 7:00 AM) colony room. At least one week prior to behavioural testing, feeding was restricted to ensure that weight was approximately 95% of ad libitum feeding weight, and never dropped below 85%. All animal research was carried out in accordance with the National Institute of Health Guide for the Care and Use of Laboratories Animals (NIH publications No. 80-23, revised 1996) and approved by the University of New South Wales Animal Care and Ethics Committee.

*Apparatus.* Behavioural testing was conducted in eight identical operant chambers (30.5 x 32.5 x 29.5 cm; Med Associates) individually housed within ventilated sound attenuating cabinets. Each chamber was fitted with a 3-W house light that was centrally located at the top of the left-hand wall. Food pellets could be delivered into a recessed magazine, centrally located at the bottom of the right-hand wall. Delivery of up to two separate liquid rewards via rubber tubing into the magazine was achieved using peristaltic pumps located above the testing chamber. The top of the magazine contained a white LED light that could serve as a visual stimulus. Access to the magazine was measured by infrared detectors at the mouth of the recess. Two retractable levers were located on either side of the magazine on the right-hand wall. A speaker located to the right of the house light could provide auditory stimuli to the chamber. In addition, a 5-Hz train of clicks produced by a heavy-duty relay placed outside the chamber at the back-right corner of the cabinet was used as an auditory stimulus. The chambers were wiped down with ethanol (80% v/v) between each session. A computer equipped with Med-PC software (Med Associates Inc., St. Albans, VT, USA) was used to control the experimental procedures and record data.

*Consumption chambers.* To provide individual access to reinforcers during the satiety and devaluation procedures, rats were individually placed into an individual cage (33 x 18 x 14 cm clear Perspex cage with a wireframe top). Pellet reinforcers were presented in small glass ramekins inside the box and liquid reinforcers were presented in water bottles with a sipper tube. 1 day prior to the target procedure, all rats were exposed to the individual cages and given 30 mins of free access to home cage food and water to reduce novelty to the context and consuming from the ramekin and water bottles.

*Locomotor activity.* Locomotor activity was assessed in eight identical boxes measuring 50 x 36x 18 cm (length x width x height), housed in a sound attenuated room. Each box consisted of 4 opaque white polyurethane walls and floor and a removable roof. In the center of the roof was an 18x40 cm grid of 3x3 mm ventilation holes. Two custom pairs of infrared beam detectors spanned the width of the box to detect locomotor activity and were located 15 cm from each end of the box. Beam breaks, corresponding to activity within the box, were recorded on a computer equipped with Med-PC software (Med Associates Inc.).

*Surgery.* Excitotoxic lesions targeting the lateral OFC were performed in Experiment 1 and 3. Rats were anesthetized with isoflurane, their heads shaved, and placed in a stereotaxic frame (World Precision Instruments, Inc., Sarasota, FL, USA). The scalp was incised, and the skull exposed and adjusted to flat skull position. Two small holes were drilled into the skull and the dura mater was severed to reveal the underlying cortical parenchyma. A 1-µL Hamilton needle (Hamilton Company, Reno, NV, USA) was lowered through the two holes targeting the lateral OFC (co-ordinates specified below). Stereotaxic co-ordinates were AP: +3.5 mm; ML: ±2.2 mm; D-V: -5.0 mm from bregma. At each site the needle was first left to rest for 1 min. Then an infusion of N-methyl-D-aspartic acid (NMDA; Sigma-Aldrich, Switzerland), dissolved in phosphate buffered saline (pH 7.4) to achieve a concentration of 10μg/μL, was infused for 3 mins at a rate of 0.1 µ/min. Finally, the needle was left in situ for a further 4 mins to allow the solution to diffuse into the tissue. Following the diffusion period, the syringe was retracted, and the scalp cleaned and sutured. Sham lesions proceeded identically to excitotoxic lesions except that no drugs were infused during the infusion period. After a minimum of 1 week of postoperative recovery, rats were returned to food restriction for 2 days prior to further training.

In Experiments 2, 4, 5, and 6, bilateral guide cannulae were surgically implanted targeting the lateral OFC. Rats were anesthetized with isoflurane, their heads shaved, and placed in a stereotaxic frame (World Precision Instruments, Inc., Sarasota, FL, USA). The scalp was incised, and the skull exposed and adjusted to flat skull position. Two small holes were drilled for the cannulae using a high-speed drill, and four holes were hand drilled on different bone plates to hold fixing screws. Bilateral stainless steel guide cannulae (26 gauge, length 5mm below pedestal; Plastics One, Roanoke, VA, USA) were lowered into the lateral OFC (AP: +3.5 mm; ML: ±2.2 mm; D-V: -4.0 mm from bregma). Cannulae were held in place by dental cement and anchored to the skull with 4 fixing screws. Removable dummy cannulae were inserted into the guide cannulae to prevent them from blocking. After one week of postoperative recovery, rats were returned to food restriction for 2 days prior to further testing.

*Drugs and infusions.* The GABAA agonist muscimol (Sigma-Aldrich, Switzerland) was dissolved in 0.9% (w/v) non-pyrogenic saline to obtain a final concentration of 0.5 *μ*g/0.5 *μ*l. Non-pyrogenic saline 0.9% (w/v) was used as the saline control.During infusions, muscimol or saline was infused bilaterally into the lateral OFC by inserting a 33-gauge internal cannula into the guide cannula which extended 1 mm ventral to the guide tip. The internal cannula was connected to a 25 *μ*l glass syringe (Hamilton Company, Reno, NV, USA) attached to a microinfusion pump (World Precision Instruments, Inc., Sarasota, FL, USA). A total volume of 0.5 *μ*l was delivered to each side at a rate of 0.25 *μ*l/min. The internal cannula remained in place for an additional 1 min after the infusion and then removed. During the infusion procedure animals were allowed to move freely in a bucket to minimize stress. Dummy cannulae were removed prior to, and replaced immediately after, infusions. For the two training sessions prior to infusions, all animals received dummy infusions which were identical to the infusion procedure, except that no liquids were infused. These dummy infusions were performed to familiarize the rats with the microinfusion procedure and thereby minimize stress. Dummy infusions were also conducted on test sessions after the infusions to minimise differences in handling between experimental stages.

*Reinforcers***.** The reinforcers used were a single grain pellet (45 mg dustless precision grain-based pellets; Bio-serv, Frenchtown, NJ, USA), 20% w/v sucrose solution and 20% w/v maltodextrin solution (Myopure, Petersham, NSW, Australia). Liquid reinforcers were flavoured with either 0.4% v/v concentrated lemon juice (Berri, Melbourne, Victoria, Australia) or 0.2% v/v peppermint extract (Queen Fine Foods, Alderley, QLD, Australia) to provide unique sensory properties to each reinforcer. Liquids were delivered over a period of 0.33 s via a peristaltic pump which corresponded to a volume of 0.2 mL. The volume and concentration of liquid reinforcers was chosen to match the calorific value of the corresponding grain pellet reward and have been found to elicit similar rates of Pavlovian and instrumental responding as a pellet reward in other experiments conducted in this lab. In all experiments involving liquids, the magazine was scrubbed with warm water and thoroughly dried between sessions to remove residual traces of the liquid reinforcer. To reduce neophobia to the reinforcers, one day prior to magazine training sessions all animals were pre-exposed to the reinforcers (10 g of pellets per animal and 25 ml of liquid reinforcer per animal) in their home cage.

*Magazine training.* All animals received one session of magazine training for each experimental reinforcer with the following parameters: reward delivery was on an RT60 s schedule for 16 rewards. When necessary, sessions were separated by at least 2 hours and the order of reinforcer identity was counterbalanced between groups.

*Behaviour.* CS responding was operationalized as the number of magazine entries during the CS period. PreCS responding was operationalized as the frequency of responding during the immediately preceding the CS period and was used as a measure of baseline responding to the testing context. PreCS responding was analysed separately, and any group differences identified and reported. Data were presented as CS – PreCS difference scores, which reflect discriminative responding to the CS. All data were analysed with mixed ANOVAs using R statistical software (Lenth et al., 2020; R Core Team (2020), 2020; Singmann et al., 2020), and significant interactions of interest were followed up with ANOVAs on the relevant subset of data, and simple effects with a Tukey family-wise error rate correction. Where relevant, planned linear and quadratic orthogonal trend contrasts and their interactions between groups were analysed to assess differences in rates of responding.

**Experiment 1: Pre-training OFC lesions**

**Subjects.**

Subjects were forty-eight (N = 48) rats, tested in two cohorts. Cohort 1, n = 16 rats weighing between 280-361 g (M = 312.2 g) and cohort 2, n = 32 rats weighing between 271-328 g (M = 296.3 g).

Training

Pavlovian Acquisition

Following magazine training, all rats received 21 sessions of Pavlovian acquisition training. Each session consisted of 16 presentations of a single auditory CS (a 15 s train of clicks) presented on a VT90s schedule (ranging from 60 to 120 s). A single pellet (US) was delivered at the termination of each CS. The session duration was 28 mins and animals were left in the chamber for an additional 2 mins before being removed. Animals received either one session per day, or two sessions per day separated by at least 2 hours.

Subgroup 1: General Satiety Pre-Feeding

At the end of acquisition training on day 21, a subgroup of animals (sham n = 8, lesion n = 8) were taken off food restriction and given 24 hours free access to their home cage food before further acquisition training on day 22. This session was rewarded as per acquisition training. At the end of day 22 animals were put back on food restriction and continued acquisition training.

***Subgroup 2: Devaluation***

Following initial Pavlovian acquisition of a single CS-US association, a subgroup of animals (sham n = 8, lesion n = 8) were re-trained with two novel unique CS-US associations intended to test devaluation in a taste aversion procedure.

Novel Acquisition

Novel acquisition of two unique CS-US associations was conducted with identical parameters to initial acquisition training, 2 session per day for 14 days, each session consisting of 16 trials consisting of a 15s CS co-terminating with reward with a vITI90s. Unlike initial acquisition the two CSs were an 80dB white noise and a 2800 Hz, 80 dB tone followed by either a single pellet or 20% w/v maltodextrin liquid (CS-US identities counterbalanced between animals).

Taste Aversion

Taste aversion took place in the devaluation chambers and involved 30 mins exposure to one US every day, alternating each day for 4 days. Following fee access to a US, animals were immediately injected i.p. with either 0.15M LiCl or 0.9% saline (15 mL/Kg). The outcome paired with nausea induced by injection of LiCl was designated the devalued outcome and the outcome paired with neutral saline injections was designated the non-devalued outcome (counterbalanced between animals). Following the final day of injections all animals were given a day of rest in their home cage to allow hunger levels to return to normal after taste aversion training.

Devaluation Test

Animals were tested with a single session of CS training except that no rewards were delivered i.e. in extinction. The magazine frequency measure that was available was not as sensitive to devaluation as a measure of duration, so only data from the first trial was analysed at test.

Locomotor Activity

At the end of the experimental procedures, all animals were assessed for locomotor activity over a 1-hour period.

**Histology and Group Allocation**

Lesion damage is depicted in Figure 1-figure supplement 1. Lesion extent was judged by a trained observer blind to group allocation. A lesion was retained if there was evidence of significant bilateral damage constrained to LO or DLO. Animals were excluded if there was only unilateral LO/DLO damage, evidence of damage to the dorsal part of the anterior olfactory nucleus ventral to LO/DLO or if there was extensive damage to the white matter of the forceps minor of the corpus callosum. One lesioned animal did not recover from surgery, four lesion animals had only unilateral OFC damage, and one lesioned animal had extensive white matter damage. Forty-two animals were retained (N = 42, sham n = 24, lesion n = 18), of which subgroup 1 contained fifteen (*N* = 15; sham *n* = 8, lesion *n* = 7) and subgroup 2 contained thirteen (*N* = 13; sham *n* = 8, lesion *n* = 5).

**PreCS Analysis**

Analysis of the PreCS period using a Group (sham, lesion) x Block (1-7) mixed ANOVA revealed that responding was significantly higher in the lesion group than the sham group (main effect of Group *F*(1, 40) = 7.24, *p* = .01). Furthermore, while responding increased over blocks (main effect of Block *F*(6, 240) = 20.37, *p* < .001; positive linear trend *F*(1, 40) = 33.18, *p* < .001), this increase was greater in the lesion than the sham group (Block x Group interaction *F*(6, 240) = 2.52, *p* = .02; linear trend interaction *F*(1, 40) = 5.34, *p* = .03). During the first block PreCS responding was similar between groups (Sham M = 2.07, SD = 0.60; Lesion M = 2.13, SD = 0.90), by the final block PreCS responding was higher in the Lesion group (M = 4.30, SD = 1.95) than the sham group (M = 2.76, SD = 2.30).

**Experiment 2: Post-training muscimol inactivation**

**Subjects**

Subjects were thirty-two (total N = 32) male Long Evans rats (Monash Animal Services, Gippsland, Victoria, Australia) approximately 4 months old, weighing between 285-350 g (M = 319.7 g).

***Pavlovian Acquisition***

Animals were given 9sessions, 1 session per day, of Pavlovian acquisition training with session parameters identical to those described in Experiment 3a. This number of session was chosen because the effect of pre-training lesions appeared after around 9 session in Experiments 3a and 3b. Briefly, each session consisted of a VT90s ITI with 16 trials consisting of a 15s click CS co-terminating with a single pellet US. Following the final day of training all animals were taken off food restriction and received surgical implantation of guide cannulae.

**Post-Training**

***Pre-Infusion***

Following post-operative recovery animals were returned to food restriction for a day before receiving a further 2 days of acquisition training as per pre-training. However, immediately prior to entering the chamber all animals received a dummy infusion.

***Infusion***

Animals were assigned to one of two infusion groups such that performance there were no differences between groups on the final day of pre-infusion acquisition. For the next 4 days, all animals received an infusion of saline or Muscimol immediately prior to entering the testing chamber for a Pavlovian acquisition session.

***Post-Infusion***

On the final 2 days of training all animals received a further 2 days of acquisition training immediately preceded by a dummy infusion.

**Histology and Group Allocation**

Cannulae placements are illustrated in Figure 2-figure supplement 2. One animal did not recover from surgery and was excluded. Three animals were excluded as a result of the cannulae assembly detaching from the skull. A further 3 animals were excluded as a result of failing to consume the pellets after recovery from surgery. One animal from the muscimol group was excluded from analysis as a result of a cannula tip embedded within the white matter of the forceps minor of the corpus callosum. Therefore, a total of 8 animals were excluded leaving *N* = 24 (saline *n* = 12, muscimol *n* = 12).

**PreCS Rates**

PreCS baseline responding did not differ between infusion groups across training and justified the use of CS-preCS difference scores for analyses of discriminative responding. In particular, during the infusion period a Group x Day (4 days) mixed ANOVA on preCS responses revealed a significant effect of Day (*F*(3, 66) = 5.95, *p* = .001) but no significant effect of Group (*F*(1, 22) = 0.01, *p* = .93) or Group x Day interaction (*F*(3, 66) = 0.41, *p* = .741). PreCS response rates on these days were, saline *M* = 0.70, *SD* = .48, muscimol *M* = 0.72, *SD* = .48.

**Experiment 3: Post-Training OFC lesions**

**Methods**

**Subjects**

Subjects were twenty-four (total N = 24) male Long Evans rats (Monash Animal Services, Gippsland, Victoria, Australia) approximately 4 months old, weighing between 317-369 g (M = 338.9 g).

**Pre-lesion Training**

***Pavlovian Acquisition***

All animals received 9 days of Pavlovian acquisition training, 1 session per day. On the final day of training all animals were removed from food restriction for at least 24 hours before receiving sham or excitotoxic lesions of the OFC. Lesion conditions were pseudo-randomly assigned to animals such that group performance was matched on the final day of acquisition and an equal number of animals were assigned to each lesion condition in each homecage.

**Post-lesion Training**

***Pavlovian Acquisition***

Following post-operative recovery all animals were returned to food restriction for 24 hrs before receiving an additional 9 days of acquisition training.

**Histology and Group Allocation**

Lesion damage is depicted in Figure 2-figure supplement 3. Lesion extent was judged by a trained observer blind to group allocation. A lesion was retained if there was evidence of significant bilateral damage constrained to LO or DLO. Animals were excluded if there was only unilateral LO/DLO damage, evidence of damage to the dorsal part of the anterior olfactory nucleus ventral to LO/DLO or if there was extensive damage to the white matter of the forceps minor of the corpus callosum. Three lesion animals had only unilateral OFC damage and were excluded from analysis (final *N* = 21; sham *n* = 12, lesion *n* = 9).

**PreCS Responding**

PreCS levels of responding did not differ between groups across days of training, and on the final block of 3 days (post-operative) response rates (15s) were sham *M* = 2.55, *SD* = 2.03, lesion *M* = 2.74, *SD* = 0.94. A mixed Group x DayBlock (6 blocks of 3 days) ANOVA on preCS responding supported this observation with only a significant main effect of DayBlock (*F*(5, 95) = 11.52, *p* < .001, effect of Group and Group x DayBlock interaction *F* < 1.00, *p* > .81).

**Experiment 4: OFC inactivation early in acquisition**

**Subjects**

Subjects were thirty-two (total N = 16) male Long Evans rats (Monash Animal Services, Gippsland, Victoria, Australia) approximately 4 months old, weighing between 321-399 g (M = 357.4 g).

**Surgery**

Surgical implantation of cannulae occurred prior to any behavioural training.

***Pavlovian Acquisition***

Animals were given 10sessions, 1 session per day. Briefly, each session consisted of a VI 200s ITI with 16 trials consisting of a 10s light CS (illumination of the house light at the back of the chamber) co-terminating with a single pellet US. Subjects received mock infusions on days 3 and 4, and either Saline or Muscimol was infused prior to entering the chamber on days 5-9. On day 10 all animals received a mock infusion.

**Histology and exclusions**

Cannulae placements are illustrated in Figure 2-figure supplement 4. One rat in the Muscimol condition had a blocked guide cannulae and was excluded from experimental analysis. Final numbers N = 15 (Muscimol n = 7, Saline n = 8).

**PreCS Rates**

PreCS responding did not differ between infusion groups across the 10 days of Pavlovian conditioning (Group *F*1,13 = 2.72, *p* = .12; Day *F*9,117 = 1.49, *p* = .16; Group x Day *F*9,117 = 2.72, *p* = .25).

**Experiment 5: OFC inactivation prior to associative blocking**

**Subjects**

Subjects were thirty-two (total N = 32) male Long Evans rats (Monash Animal Services, Gippsland, Victoria, Australia) approximately 4 months old, weighing between 299-395 g (M = 331.5 g).

**Surgery**

Surgical implantation of cannulae occurred prior to any behavioural training.

**Training**

The design of the experiment was such that 4 CSs were designated as cues A, B, C and D. Cues A and C were always visual cues, either darkness caused by extinguishing the houselight or flashing panel lights (5Hz; Figure 3A). Cues B and D were always auditory cues, either an 80dB white noise or a 5Hz train of clicks. Throughout all training sessions the house light was always illuminated unless it was extinguished to act as a visual cue. All cues lasted 10s and co-terminated with the delivery of the US, 2 pellets delivered consecutively 0.25s apart. The identity of the cues was counterbalanced between subjects except that A and C were always visual cues and B and D were always auditory cues. Simultaneous audio-visual compounds were designated as AB and CD. Pavlovian training sessions were always 56 mins long such that there were 16 trials with a vITI 200s (range 100 to 300s); animals were left in the chambers for an additional 2 mins before being removed.

***Food Restriction and Magazine Training***

Magazine training sessions consisted of an RT120s reward delivery schedule for 16 rewards. Each reward consisted of 2 pellets delivered to the magazine 0.25s apart.

***Stage 1***

Stage 1 acquisition involved 10 days of acquisition to cue A, 16 trials per session. On days 1-4 of training all animals received dummy infusions to familiarise them to the infusion procedure. Animals were then split into two groups with matched performance on day 4. On days 5-10 all animals received an infusion of saline or muscimol immediately prior to entering the test chambers.

***Pre-exposure***

On day 11 all rats received pre-exposure to auditory cues B and D, 4 non-rewarded presentations of each cue vITI 200s. This was done to minimise novelty to the auditory cues during compound training in stage 2. All animals received dummy infusions prior to the session.

***Stage 2***

On days 12-14 all animals received stage 2 audio-visual compound training. Sessions involved 8 presentations of compound AB and 8 presentations of CD (pseudo randomly presented such that a compound was never repeated more than 2 times in a row). The compounds were rewarded with 2 pellets, the same US that was used in stage 1. All animals received dummy infusions prior to each session.

***Test***

On day 15 and 16 all animals were tested in extinction for responding to the target auditory cue B and the overshadowing control cue D (8 presentations of each cue, pseudorandom trial order, vITI 200s). All animals received dummy infusions prior to each session.

***Re-acquisition***

On days 17-19, all animals received re-acquisition training to cue B (16 trials per session) to test for differences in rates of re-acquisition to the blocked cue. On days 20-21 animals were tested for re-acquisition to cue A (16 trials per session) to test for differences in the rate of re-acquisition to the blocking cue.

**Results**

**Histology and Group Allocation**

Cannulae placements are illustrated in Figure 3-figure supplement 1. 1 animal failed to consume pellets throughout the experiment and was excluded from testing. One animal from the muscimol group lost its cannula assembly during the infusion period and was excluded from testing. One animal in the muscimol group was euthanized due to severe illness. A further 2 animals were excluded after histological analysis revealed that the cannulae were only unilaterally targeting DLO and LO. Therefore, a total of 6 animals were excluded leaving *N* = 26 (saline *n* = 13, muscimol *n* = 13).

**PreCS Responding**

Baseline levels of responding did not differ between groups during training, and on the final day of infusions (day 10 of stage 1) preCS response rates (10s) were saline *M* = 0.122, *SD* = 0.24, muscimol *M* = 0.67, *SD* = 0.87. These observations were supported by mixed Group x Day ANOVAs on preCS responding in stage1 suggesting that there were no group differences on days 1-4 prior to infusion (all *F* < 1.69, *p* > .21) or on days 5-10 during infusions (significant main effect of Day *F*(5, 120) = 15.21, *p* < .001, all remaining *F* < 1.00, *p* > .50).

**Experiment 6: Competing response values**

**Subjects**

Subjects were eight (total N = 8) male Long Evans rats (Monash Animal Services, Gippsland, Victoria, Australia) approximately 4 months old, weighing between 285-331 g (M = 314.9 g).

**Apparatus**

The apparatus comprised of 8 operant chambers (Med Associates Inc.) individually housed in light and sound attenuating cabinets. Each chamber comprised of a transparent Perspex back wall, roof and front door, with aluminium left and right-hand walls. The floor consisted of 19 steel bars (3.8mm diameter, spaced 1.6 cm apart), aligned perpendicular to the back of the chamber. Rewards could be independently delivered into one of two recessed magazines located centrally at the bottom of the left and right-hand walls. The magazine on the right-hand wall could be rewarded with food pellets (45 mg; Bio-Serv) whereas the magazine on the left hand wall could be rewarded with liquid rewards delivered by a dipper cup mechanism that could be retracted from the magazine. Access to the magazines was measured by infrared detectors at the mouth of the recess. Two panel lights (2 cm diameter) were located on either side of the right-hand magazine at the top of the right-hand wall. A 3-W house light was located at the top left of the left-hand wall. A speaker located to the right of the house light (on the top far right of the left-hand wall) could provide auditory stimuli to the chamber. In addition, a 5-Hz train of clicks produced by a heavy-duty relay placed outside the chamber at the back-right corner of the cabinet was used as an auditory stimulus. A computer equipped with Med-PC software (Med Associates Inc.) was used to control the experimental procedures and record data.

**Food Restriction and Magazine Training**

All animals were food restricted for at least 2 days prior to any training, and pre-exposed to sucrose (10 mL) and pellets (5g per rat) in their homecage.

Magazine training involved the unsignalled delivery of the reinforcer in the experimental chamber to familiarise the subjects with retrieving rewards from each of the two magazines. All rats received two separate magazine training sessions in one day (separated by at least 2 hours), one for each magazine, order counterbalanced. The dipper magazine was always paired with 20% w/v sucrose solution, the other magazine was always rewarded with pellets. Magazine training involved un-signalled delivery of reward on an RT60s schedule for 16 rewards.

**Acquisition**

Acquisition training lasted for 16 days. In each session rats received 3 consecutive blocks of 8 trials. Each trial consisted of a vITI 105s, a 15s CS (80 dB white noise) co-terminating in a pellet delivered into the pellet magazine. Simultaneously, there was always a probability of un-signalled 5s access to sucrose in the dipper magazine (dipper cup held 0.01cm3 fluid). The probability of sucrose availability changed randomly between each block from low (p = 2/24), medium (p = 4/24) to high (p=8/24). Each trial was defined by 24 bins of 5s (trial length = 120s). During this period the CS would occur across 3 5s bins (15s CS) and there was the possibility of un-signalled reward at the dipper magazine at the start of each 5s time bin. Notably, un-signalled reward could occur during any 5s bin, including the CS period. Unsignalled rewards occurred 2, 4, or 8 times in each trial. The pellets were therefore signalled rewards (as they were reliably preceded by the CS). The sucrose dipper was un-signalled, and if the sucrose was not collected during the 5s period the dipper would be retracted and lost.

**Water deprivation**

On days 17 and 18, all animals were water restricted for 22h prior to testing. Test sessions were identical to acquisition sessions. Animals were given 2 hours of free access to water 2 hours after test sessions. Animals were given 24 hours of ad libitum access to water after day 18 before any further testing to ensure animals were no longer thirsty in subsequent tests.

**Un-Signalled Reward Shift**

On testing days 19 and 20 received 2 sessions of acquisition with an increased magnitude of un-signalled reward delivery. Specifically, the size of the dipper cup was increased from 0.01 to 0.04 cm3 of fluid so that each un-signalled reward was increased in volume.

**Surgery and drug infusions**

Following testing day 20 all animals were taken off food restriction and underwent surgical implantation of guide cannulae targeting the lateral OFC.

**Test**

Following post-operative recovery all animals received 3 days of training immediately preceded by dummy infusions. All post-operative sessions used the larger sucrose dipper cup volume (0.04 cm3). Following this, all animals received 2 days of training in which they received 1 test day under saline infusion and 1 test day under muscimol infusion. The order of infusion days was counterbalanced. Test sessions were shortened to only 2 blocks with a fixed progression from low probability (p = 2/24) to high probability (p = 8/24) of un-signalled reward. This was done to ensure that muscimol was still active during the test session by keeping the session duration to 30 mins.

**Response Analysis**

It is important to note that all magazine responding in this procedure was analysed from periods in which the dipper reward was not physically present to eliminate the possibility that magazine responses simply reflect sucrose consumption.

**Histology**

Cannulae placements are illustrated in Figure 5-figure supplement 1. Two animals were due to the cannulae assembly losing patency during post-operatively. Therefore, a total *N*= 6 animals were tested post-operatively.