

Research Articles | Behavioral/Cognitive

## Dissociable Roles of Prelimbic and Orbitofrontal Cortices in the Flexible Use of Model-Based Associations

<https://doi.org/10.1523/JNEUROSCI.0892-26.2026>

Received: 26 May 2026

Revised: 3 September 2026

Accepted: 23 September 2026

Copyright © 2026 the authors

---

*This Early Release article has been peer reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.*

**Alerts:** Sign up at [www.jneurosci.org/alerts](http://www.jneurosci.org/alerts) to receive customized email alerts when the fully formatted version of this article is published.

# Dissociable Roles of Prelimbic and Orbitofrontal Cortices in the Flexible Use of Model-Based Associations

Adrian Dawson<sup>1</sup>, Daniel Wilson<sup>1</sup>, Masakazu Taira<sup>1</sup>, Mirari Wilcher<sup>1</sup>, Arvie Rodriguez Abiero<sup>1</sup>, Jessica Leake<sup>1</sup>, and Melissa J. Sharpe<sup>1,2</sup>.

<sup>1</sup>Department of Psychology, University of Sydney, Camperdown, NSW Australia 2050

<sup>2</sup>Charles Perkins Centre, University of Sydney, Camperdown, NSW Australia 2050

Correspondence: [melissa.sharpe@sydney.edu.au](mailto:melissa.sharpe@sydney.edu.au) and [adaw8204@uni.sydney.edu.au](mailto:adaw8204@uni.sydney.edu.au)

**Abstract:** Our environment is complex and to navigate it effectively, we organize conceptual and spatial information into a coherent cognitive map. This allows us to be flexible and make novel inferences to maximize procurement of the things we need and minimize negative events. Decision making based on such cognitive maps is generally described as model-based behavior, which refers to an ability to simulate likely paths to reward via a transitional structure of the task. We have previously modelled how different prefrontal regions contribute to deploying cognitive maps to facilitate flexible behavior. We hypothesized that the prelimbic region- as analogous to the dorsolateral prefrontal cortex in humans- is necessary to upload the current relevant map, while orbitofrontal cortex tracks position within the map as subjects traverse it. To test this, we developed a sequential biconditional discrimination task allowing optical inhibition during either map upload or position tracking in male and female rats. We show this task engenders a detailed model-based cognitive map, fostering the ability to make novel inferences. Consistent with our framework, optical inhibition of the orbitofrontal cortex disrupts rats' ability to track their current position in the map. In contrast, optical inhibition of the prelimbic cortex reduces performance when rats must upload the current map. While rats without prelimbic activity during map upload still show discriminative performance, their discrimination is uncoupled from that seen during training, suggesting a compensatory mechanism. Together, these data reveal dissociable contributions of the orbitofrontal and prelimbic cortices to cognitive map deployment.

**Significance Statement:** Our ability to behave appropriately in new or unfamiliar situations depends on what we learn from experience. To navigate complex environments, we develop cognitive maps that comprise mental frameworks organising learned information into coherent schemas. These allow us to choose between courses of action in different contexts and to draw new inferences in entirely novel situations. The ability to form and use such maps is disrupted in psychological disorders including addiction and schizophrenia. Our work reveals new information on how the brain constructs and deploys cognitive maps, bringing us closer to understanding the cognitive disruptions that underlie these disorders.

35     **Abstract: 246 words**

36     **Introduction: 631 (excluding references)**

37     **Discussion: 1475 (excluding references)**

JNeurosci Accepted Manuscript

# Main Text

We need to develop an understanding of the environment that allows us to navigate successfully in ambiguous or novel circumstances. To do this, we develop cognitive maps that organize our learned experiences into a coherent, sophisticated framework (Behrens et al., 2018; Costa et al., 2023; Sharpe et al., 2019; Tolman, 1948; Whittington et al., 2022; Wikenheiser & Schoenbaum, 2016; Wilson et al., 2014). These maps are referred to as model-based associations and engender our ability to make novel inferences (Jones et al., 2012; Tolman, 1948). For example, when you move into a new neighborhood, you might notice a couple walking their dog together in the mornings. If you subsequently witness one of them expressing a particular belief, you're likely to assume that the other holds that view too and adapt accordingly when you next interact with them. This is important; you're more likely to have successful interpersonal relationships if you interact with people in a socially astute way. A cognitive map allows you to store separately learned information in related schema so that you can generate new predictions in unfamiliar situations and behave adaptively, even without direct experience (Behrens et al., 2018; Tolman, 1948; Wikenheiser & Schoenbaum, 2016; Wilson et al., 2014).

There are many regions of the brain that contribute to the development and use of cognitive maps, including orbitofrontal cortex, prelimbic cortex, and hippocampus (Bradfield et al., 2015; Bradfield & Hart, 2020; Costa et al., 2023; Granon & Poucet, 2000; Jones et al., 2012; Schuck et al., 2016; Schwartenbeck et al., 2023; Veselic et al., 2025; Wallis et al., 2001; Whittington et al., 2025; Wikenheiser & Schoenbaum, 2016; Wilson et al., 2014). However, it is still unclear how different brain regions contribute uniquely and causally to these processes (Sharpe et al., 2019; Veselic et al., 2025). We have put forward a hypothetical framework that helps to understand how the prelimbic and orbitofrontal cortices may differentially regulate the use of cognitive maps (Sharpe et al., 2019). In this framework, the prelimbic cortex *uploads* the relevant cognitive map that suits the current circumstance. This accounts for the role of the prelimbic cortex in control of behavior by higher-order cues or rules (Roughley & Killcross, 2019; M. Sharpe & S. Killcross, 2015; Sharpe & Killcross, 2018; Vendrell et al., 1995; Wallis et al., 2001), which allows us to adjust to different contingencies across different contexts. In contrast, the orbitofrontal cortex *tracks* the current point within the map that you're currently in (Schuck et al., 2016; Wilson et al., 2014). Here, the orbitofrontal cortex integrates all relevant information that comprises the current state of the task, which includes observable and unobservable information to determine reward probability and transition into the next state of the task. When all relevant information is observable, orbitofrontal cortex signaling is redundant, but when it is not, orbitofrontal tracking of the task state becomes essential (Sharpe et al., 2019). This considers the role of the orbitofrontal cortex in behaviours that require decision making based on information in the recent past that is no longer observable or must be inferred (Bradfield et al., 2015; Jones et al., 2012; Pickens et al., 2003; Schuck et al., 2016; Wilson et al., 2014), but explicitly not in tasks that involve simple discrimination between cues (Gardner et al., 2017; Schoenbaum et al., 2003). This model synthesizes the existing literature to propose that prelimbic and orbitofrontal regions act in a distinct but coordinated manner to facilitate the flexible use of cognitive maps. Yet this model has not been formally tested and there are currently few studies that compare function of these prefrontal regions, let alone their contribution to deployment of a cognitive map.

To understand the unique role that the prelimbic and orbitofrontal cortices play in use of cognitive maps, we developed a task that would allow us to isolate the upload and tracking of such maps in a dissociable way. We designed a sequential biconditional discrimination that would allow us to optogenetically inhibit pyramidal neurons within the orbitofrontal and prelimbic cortices at different points during the deployment of a cognitive map. This discrimination cannot be solved by simply learning the value of each independent cue. Instead, rats need to upload the current map based on early cues that dictate the contingencies in effect later in the trial, then track that information, including currently unobservable information about what was presented earlier. Thus, this task allows us to dissociate the upload and tracking of cognitive maps. Our model predicts that inhibition of the prelimbic cortex early in the trial would impair flexible responding, whereas inhibition of the orbitofrontal cortex later in the trial would abolish differential responding in the task.

## Methods

### Differential outcomes are required to solve the sequential biconditional task

#### *Subjects*

Subjects were 80, experimentally naïve, wild-type male ( $n = 51$ ) and female ( $n = 29$ ) Long-Evans rats obtained from either Ozgene (Perth, WA) or our breeding colonies (Kearns facility, University of Sydney). The rats were housed in groups of two in ventilated plastic tubs (Techniplast), measuring 40x46x40 cm (length x width x height), in a colony room that was on a reverse 12h light-dark cycle with the lights turned off at 9am. Rats had unrestricted access to water and were maintained at 85% of their free feeding body weights by daily supplemental feedings following experimental sessions. Experimental sessions occurred during the dark phase. G\*Power analysis (Faul et al., 2007) was used to compute the power achieved with 16 subjects. With this sample size, we had power of 0.8 to detect an effect size of  $\eta^2 = 0.124$  for the critical interaction between trial type (reinforcement vs non-reinforcement) and group (differential vs non-differential or NpHR vs. eGFP). This is sufficient and consistent with the previous work (Delamater et al., 2010). To accommodate the likelihood of some attrition for the optogenetic experiments, the sample size was increased to 24 for optogenetic experiments.

#### *Apparatus*

All behavioural testing was conducted in 8 identical operant chambers controlled by MED-PC V software (Med Associates, Inc., Fairfax, VT). The chambers measured 28.5x30x25 cm (height x length x depth). In each chamber, bedding lay below a floor of stainless-steel rods, 0.5cm in diameter, spaced 1.5cm apart. The ceiling and sidewalls were Plexiglas while the front and rear walls were aluminum. Each chamber was equipped with a food port on the front wall that held a receptacle for pellets and a receptacle for liquid rewards. The rewards were either two 45mg pellets (Bio-Serv, #F0021) or ~0.2mL of 20% sucrose solution for all

experiments, with the exception of experiment 2, where we used either 45mg grain pellets (Bio-Serv, #F0021) or sucrose pellets (Bio-Serv, #F0023). The port was equipped with a head entry detector to measure the number of port entries as well as the amount of time spent in the port. A sound generator and clicker were positioned on the wall opposite the food port. The four stimuli in the experiment were auditory cues. The distal cues were a white noise or clicking sound. The proximal cues were a 1400Hz tone and a siren that oscillated in frequency between 1500-1900Hz. The white noise, tone and siren were generated by a single sound module (Med Associates, product ENV-224AM-3) while the click was generated by a separate module (Med Associates, product ENV 135 M).

### **General Procedures**

All procedures were approved by the Animal Ethics Committee at the University of Sydney (project number 2023/2310).

*Food Port Training.* Rats were initially trained to retrieve food rewards from the food port for one session. During the session, 30 rewards were delivered randomly across 60 minutes with no stimulus presentations. Rats in non-differential reinforcement groups only received the reward type that would be presented in the following biconditional discrimination. Rats in differential reinforcement groups received a pseudorandom mixture of trials with one of the two rewards being delivered per trial.

*Biconditional Discrimination Training.* All subjects were trained on the same biconditional discrimination task with one session per day. Each session consisted of 40 trials with a random interval between trials that averaged around a variable 2min mean (80-160s). Each trial involved a 10s presentation of a distal cue (white noise or click), a 500ms gap, followed by a 10s presentation of a proximal cue (tone or siren). In reinforced trials, the proximal cue was followed by delivery of reward. After one distal cue (e.g., click), one proximal cue led to reward (e.g., tone) and the other did not (e.g., siren). However, this reinforcement contingency was reversed with the other distal cue (i.e., noise), such that the tone was no longer rewarded, and the siren led to reward. This setup created four distinct trial types, where reinforcement occurred in two of the combinations that did not share any common cues. The trial types that were followed by reinforcement were kept constant for each rat but counterbalanced across subjects such that different rats received reinforcement after different trial types. In our first experiment, half the rats received differential reinforcement such that one reinforced trial type was always followed by delivery of sucrose while the other reinforced trial was followed by a grain pellet (counterbalanced across subjects). The other half formed the non-differential reinforcement group that received only one type of reward for either reinforced trial type such that all reinforced trials were followed by sucrose or all were followed by a grain pellet. The reward used for each rat was kept constant but counterbalanced across subjects. In the remaining experiments, all rats received differential reinforcement.

### **Specific Procedures**

*Training with new proximal cues and the inference tests.* In Experiment 2, we tested whether rats could use the biconditional associations to make new inferences. To do this, following the differential biconditional discrimination training, rats received additional training with two new partially reinforced 10s auditory cues, termed “jingle” and “horn”. The jingle sound was an alternating tone increasing and decreasing in frequency by 200Hz every 0.2s, with a range from 1000-2000Hz. The horn consisted of an 1200Hz tone that would alternate on and off every 0.3s. One of these cues was followed by the grain pellets on 50% of trials, and the other was followed by the sucrose pellets on 50% of trials (e.g., jingle  $\square$  grain pellets; horn  $\square$  sucrose pellets; counterbalanced across rats). These cues were partially reinforced to equate them with partially reinforced proximal cues in the biconditional training phase. Rats received 4 sessions, each with 24 trials (12 reinforced), with our standard 2min variable ITI. Finally, rats received the critical inference test under extinction where four new compounds were formed (i.e., white noise  $\square$  jingle; white noise  $\square$  horn; click  $\square$  jingle; click  $\square$  horn). Cues were presented pseudo randomly with a 2min variable ITI, with a 500ms gap between distal and proximal cues presentations. Rats received 4 test sessions, each with 20 trials in total, which gave 5 trials of each trial type per test session.

*Surgeries.* Surgical followed those described previously (Hart et al., 2020; Seitz et al., 2022; M. Sharpe & S. Killcross, 2015). Rats (all >280g) were anesthetized with isoflurane (5% induction, 2-3% maintenance in 1L/m O<sub>2</sub>) and placed in a standard stereotaxic device. Burr holes were drilled bilaterally in the skull for insertion of the 33-gauge injector and optic fibres. Rats were infused with the virus at a flow rate of 0.2 $\mu$ L/minute, and injectors were subsequently left in place for an additional 2 minutes to allow for diffusion of solution. Halorhodopsin (NpHR) was used for inhibition while enhanced green fluorescent protein (eGFP) was used as a control. NpHR rats received bilateral infusions of AAV9-CamKIIa-eNpHR3.0-eYFP (0.8 $\mu$ L-1 $\mu$ L per hemisphere, titre  $\sim 10^{12}$ ). eGFP rats received bilateral infusion of AAV8-CaMKIIa-eGFP (0.8 $\mu$ L-1 $\mu$ L per hemisphere, titre  $\sim 10^{12}$ ). The coordinates used for targeting the lateral orbitofrontal cortex were: AP: +3.0 mm; ML:  $\pm 3.2$  mm; DV: -5.3 mm (male) or: AP: +2.85 mm; ML:  $\pm 3.04$  mm; DV: -5.09 mm (female) from Bregma. Optic fibres (200  $\mu$ m core diameter, Thorlabs, Newton, NJ) were implanted bilaterally at an angle 10° away from the midline at the following target: AP: +3.0 mm; ML:  $\pm 3.2$  mm; DV: -4.90 mm (male) or: AP: +2.85 mm; ML:  $\pm 3.04$  mm; DV: -4.66 mm (female) from Bregma. The coordinates for prelimbic cortex were: anteroposterior, +3.0; mediolateral,  $\pm 0.7$  and dorsoventral, -3.8 from Bregma. Optic fibres were targeted to the prelimbic cortex at the following coordinates: AP: + 3.0; ML:  $\pm 0.7$ ; DV: -3.3 from Bregma. Optic fibres in the prelimbic cortex were implanted bilaterally at an angle of 16 degrees towards the midline. Headcaps were secured with 6, 0-80 1/8” machine screws and dental acrylic. Rats were given at least four weeks to incubate viral expression before testing.

*Optogenetic Test Sessions.* Following biconditional training, rats in the optogenetic experiments experienced the proximal optogenetic test. Then subjects were exposed to 2 more training sessions before a distal optogenetic test. The test sessions involved 20 trials without delivery of reward following any trial. Conducting the test sessions under extinction (i.e., without rewards) ensures that rats must use their knowledge of the task to direct responding and cannot use the rewards to learn the structure of the task during the session. For

the proximal test, green laser light (532nm; 16-18mW) was delivered across the proximal cue, beginning 300ms before cue onset and ending 300ms after cue offset. For the distal test, the laser light began 300ms before distal cue onset and ended 300ms after cue offset. The offset in delivery of light from the cue ensures that the neurons remain inhibited across the entire cue period (Sharpe et al., 2021; Sharpe, Marchant, et al., 2017).

*Correlations for the Data with Optical Inhibition of Prelimbic Cortex:* A discrimination score was calculated for each subject at training and test, based on performance on reinforced and non-reinforced trials: (Reinforced – Non-reinforced) / (Reinforced + Non-reinforced). Discrimination scores from the final training session and the test session were then correlated using a Pearson correlation.

### ***Histology for optogenetic experiments***

Rats were anesthetized with isoflurane and underwent transcardial perfusion with 4% paraformaldehyde in phosphate buffered saline (PBS). Brains were extracted and left in the fixative solution overnight at 4 degrees and then transferred into 30% sucrose PBS solution. The samples were sectioned into 40µm thickness using a cryostat. The brain slices were mounted onto glass slides and cover-slipped with DAPI (through VectashieldDAPI, Vector Lab, Burlingame, CA). To verify the location of the optic fibre and viral vector expression, slides were scanned using a Zeiss Axioscan slide scanner (Sydney Microscopy & Microanalysis). This allowed visualisation of the fluorescent protein that was part of both viral vectors. Higher magnification images (20x and 63x) were then taken on a Zeiss LSM800 Airyscan (Sydney Microscopy & Microanalysis).

### ***Statistical Analyses for all experiments***

Behavioural data were analysed using IBM SPSS (version 29.0.1.0), with data extracted from Med-PC to Excel. The primary dependent variable across training and testing was the mean number of food port entries during the proximal cue. The one exception to this was the inference tests, where data were analysed using the time spent in the food port, consistent with data from other inferential tests (Jones et al., 2012; Sharpe et al., 2020; Sharpe, Chang, et al., 2017). This is also consistent with more recent work (Schierbeck et al., 2026). Here, the degree of time waiting for reward (i.e., duration) reflects inferred value (Schierbeck et al., 2026), and can be explicitly dissociated from number of entries in the same task that reflects observed value (Golden et al., 2025), where the former is dependent on orbitofrontal cortex (Schierbeck et al., 2026). Test data often included responses made during the expected reward period to capture expectation behaviour in the unrewarded test sessions. For training phases, sessions were combined into session blocks resulting in 5 or 6 blocks total with within-subjects variables of trial type (reinforced/non-reinforced) and session block, and between-subjects variables of group (differential/non-differential or NpHR/eGFP) where applicable. For test and inference sessions, each session was analysed separately, with the within-subjects variable being trial type (reinforced/non-reinforced, or congruent/incongruent). All port entry data were subjected to mixed-design repeated-measures ANOVAs, with follow-up simple main effects analyses when justified by a significant

interaction (Howell, 1992), and the Greenhouse-Geisser adjustment applied when sphericity was violated. A one-tailed test was applied to a significant interaction with a directional hypothesis (Howell, 1992). SEM values for within-subject comparisons were calculated (Loftus & Masson, 1994). Where relevant, predefined exclusion criteria were applied to verify virus expression and fibre placement, with rats excluded if either was clearly absent from the target region. To further characterise test data where applicable, post-hoc linear correlations were performed between discriminative responding during the final training session and during test, calculated as port entries on reinforced trials minus port entries on unreinforced trials, divided by the total number of port entries made.

## Results

### Differential outcomes are required to solve the sequential biconditional task

A biconditional discrimination typically involves setting up contingencies where a higher-order stimulus (e.g., a discrete cue or context) dictates which contingencies are currently in effect (Bradfield & Balleine, 2013; Colwill & Delamater, 1995; Delamater et al., 2010; Harris & Livesey, 2008). In our version of the task, a trial begins with presentation of one of two distal 10s auditory cues (e.g., white noise or click). Next, one of two 10s auditory proximal cues is presented (e.g., tone or siren). Whether the proximal cue is followed by food reward depends on the preceding distal cue. For example, the white noise predicts that the tone will be rewarded and the siren will not (i.e., white noise  $\square$  tone  $\square$  reward; white noise  $\square$  siren  $\square$  nothing), while the click predicts that the siren will be reward but the tone will not (click  $\square$  tone  $\square$  nothing; click  $\square$  siren  $\square$  reward; all cue identities counterbalanced across subjects). Prior research has shown that subjects solve a biconditional discrimination more effectively if each cue is paired with a distinct reward, in line with the differential outcomes effect (Delamater et al., 2010). The presence of a differential outcomes effect lends support for the notion that the associations underlying a task rely on detailed representations of the learned associations that transcend the backpropagation of general value from the food reward to the cues (Sharpe & Schoenbaum, 2018).

To test whether our specific version of the task also lends to a differential outcomes effect, we first trained two groups of rats on different versions of the task (**Fig 1A-B**). In the first group, rats learnt the biconditional discrimination with both proximal cues predicting grain pellets (e.g., tone  $\square$  pellets, siren  $\square$  pellets; termed, same group). In a second group, rats learnt the biconditional discrimination with each proximal cue predicting one of two outcomes (e.g., tone  $\square$  pellets, siren  $\square$  sucrose; termed, differential outcomes group). Across training, all rats showed an increase in responding during the proximal cues (**Fig 1C**). However, there was a considerable difference in how the same and differential group responded across training. While the differential group showed strong differentiation in responding to the proximal cues based on whether they were likely to be rewarded, the same group failed to show this differentiation. In line with this observation, a mixed-design ANOVA revealed a main effect of session block (session:  $F_{(4,56)} = 10.65, p < .001$ ), a main effect of trial type (rewarded vs. nonrewarded; ( $F_{(1,14)} = 16.27, p = .001$ ), and a main effect of group (same vs. differential;  $F_{(1,14)} = 17.57, p < .001$ ). There was also a significant interaction between trial type and

session block (rewarded vs. nonrewarded x session;  $F_{(4,56)} = 8.54, p < 0.001$ ), indicating that rats more successfully discriminated between rewarded and nonrewarded trials across time. Further, a group by trial type interaction showed that the differential group showed greater differentiation between rewarded and nonrewarded trials relative to the same group (same vs. differential x rewarded vs. nonrewarded;  $F_{(1,14)} = 7.08, p = .019$ ). Indeed, a simple main effect analysis warranted by this significant interaction showed that the differential group demonstrated significantly more responding to the proximal cue on rewarded vs. nonrewarded trials ( $F_{(1,14)} = 22.41, p < .001$ ), but the same group did not ( $F_{(1,14)} = 0.94, p = .348$ ). Finally, there was a three-way interaction between trial type, session, and group ( $F_{(4,56)} = 6.792, p < .001$ ), which was due to a significant difference in responding on rewarded vs. nonrewarded trials in the differential group on session block 3 ( $F_{(1,14)} = 12.31, p < .001$ ), session block 4 ( $F_{(1,14)} = 45.90, p < .001$ ), and session block 5 ( $F_{(1,14)} = 23.71, p < .001$ ). This difference was not present in the same group on any session block ( $F_s < 1.2$ ). Thus, our biconditional discrimination task was sensitive to the differential outcomes effect, which suggests behaviour underlying the task is dependent on development of rich sensory-specific associations that transcend general value.

## The sequential biconditional task engenders the ability to make novel inferences

The sensitivity of our task to the differential outcomes effect is consistent with the notion that our task supports the development of an associative structure that involves detailed representations of cues and rewards. This is because existing model-based frameworks that underlie cognitive maps [e.g. (Daw et al., 2005; Dayan, 1993; Gershman, 2018; McCloskey & Cohen, 1989)], predict that the ability to distinguish between predictive associations is dependent on how separable the state features are from another. This is because similarities across states cause interference between the transition structures that share these components. Thus, having differential outcomes allows these associative representations to be more independent and more discriminable from one another. At the same time, the presence of the differential outcomes effect argues against a model-free account, which does not comprise an associative chain of events akin to a cognitive map (Daw et al., 2005). This is because model-free accounts argue that learning is driven by backpropagation of scalar value to antecedent states and so the presence of the differential outcomes does not facilitate discrimination—the scalar value ascribed to the cue state is independent of reward identity, so there is no link between the cue and reward identity. Thus, the presence of the differential outcomes effect promotes the presence of a model-based associative chain and explicitly not a model-free account, which lacks such a transitional structure.

However, if our task really produces a cognitive map, then it should engender the ability of rats to make novel inferences (Jones et al., 2012; Schuck et al., 2016; Wilson et al., 2014). To test this, we trained a new group of rats on our biconditional discrimination task with differential outcomes (e.g., white noise □ tone □ grain pellets; white noise □ siren □ no reward; click □ tone □ no reward; click □ siren □ sucrose; **Fig 2A**). After rats had successfully learnt the task, we then taught them that two new proximal cues predicted the

two different outcomes (e.g., jingle  $\square$  grain pellets; horn  $\square$  sucrose; **Fig 2B**). If rats have built a cognitive map of the task, then they should be able to infer how these new contingencies fit into the associative structure that has developed across training. Specifically, we can present trials where the distal cues (i.e., white noise and click) precede delivery of the new proximal cues (i.e., jingle and horn) and test their responding (**Fig 2C**). If rats can make novel inferences about when reward is most likely to be delivered, they should predict that the jingle—paired with pellets—is more likely to lead to reward when it is preceded by the white noise. This is because the white noise previously predicted the tone would also be rewarded with pellets. In contrast, the rats should predict that the horn—paired with sucrose—is more likely to be rewarded when it is preceded by the click, which predicted that the siren would be followed by sucrose. Trials where the distal cue predicted the same reward as the new proximal cue were termed *congruent trials* and trials where the distal and proximal cues predicted different outcomes were termed *incongruent trials*. As we examined responding on congruent and incongruent trials under extinction (where no rewards are delivered), rats needed to recall the associated reward representations developed during training to make the inference that congruent trials were more likely to lead to reward relative to incongruent trials.

All rats learnt the differential biconditional task effectively, with discrimination between rewarded and nonrewarded trials emerging across training in a similar manner to our previous experiment (**Fig 2A**). Indeed, we found that there was a significant main effect of session block (session:  $F_{(5,55)} = 17.185$ ,  $p < 0.001$ ), and a significant trial type by session block interaction (rewarded vs. nonrewarded  $\times$  session:  $F_{(5,55)} = 3.717$ ,  $p = 0.006$ ). Follow up simple main effects analyses from this significant interaction revealed that this was due to a significant difference between rewarded vs. nonrewarded trials in later session blocks, particularly session block 3 (rewarded vs. nonrewarded:  $F_{(1,11)} = 10.252$ ,  $p = 0.008$ ) and session block 4 (session:  $F_{(1,11)} = 5.056$ ,  $p = 0.046$ ). This showed that rats were again able to solve the biconditional discrimination using differential outcomes.

Next, rats were trained with two new auditory cues that were paired with one of the two outcomes on a partially reinforced schedule (e.g., jingle  $\square$  pellets; horn  $\square$  sucrose; counterbalanced across rats; **Fig 2B**). Rats increased responding to all cues across sessions, with a main effect of session (session:  $F_{(3,33)} = 4.299$ ,  $p = 0.011$ ). Finally, rats received the **critical inference test** where they were presented with congruent and incongruent combinations of distal and proximal cues without reward feedback (**Fig 2C**). Here, rats showed a significant increase in responding on the congruent relative to incongruent trials (congruent vs. incongruent:  $F_{(1,11)} = 6.048$ ,  $p = 0.032$ ). This demonstrates that our biconditional discrimination task promotes development of a cognitive map that then allowed them to make novel inferences in new circumstances.

## **Optogenetic inhibition of the orbitofrontal cortex during the proximal cue, but not the distal cue, abolishes differential responding in the biconditional task**

Armed with the knowledge that our task involves the development of a rich associative structure, like a cognitive map, we set out to test our hypothesis that the orbitofrontal cortex is important for tracking the position within a cognitive map. We reasoned that if orbitofrontal cortex is necessary for tracking a position with a cognitive map, which requires integration of past unobservable information with current observable

information, then we would expect that optogenetic inhibition of orbitofrontal cortex during the proximal cue would impair differential responding on the task. This is because this is the point at which the rat needs to recall the past distal cue and integrate that knowledge with the current proximal cue to make a prediction about whether reward is likely to be delivered or not. We further predicted that optogenetic inhibition of orbitofrontal cortex during the distal cue would be without effect as all components of the map trajectory are currently observable and nothing needs to be inferred at this time point.

To test this, rats were first infused with adeno associated virus (AAV) on a CamKII $\alpha$  carrying halorhodopsin (NpHR group;  $n=9$ ; AAV9-CamKII $\alpha$ -eNpHR3.0-eYFP, titre  $\sim 10^{12}$ ) or a control AAV carrying eGFP (eGFP group;  $n=12$ ; AAV8-CamKII $\alpha$ -eGFP, titre  $\sim 10^{12}$ ) into the lateral orbitofrontal cortex (Hart et al., 2020). In this same surgery, rats received implantation of optic fibers into the lateral orbitofrontal cortex to allow for light delivery into the brain (Hart et al., 2020). Four weeks later, rats began training on the differential biconditional discrimination task. All aspects of the task were the same as prior versions and included a 500ms gap between the distal cue and proximal cue, and the proximal cue and the reward. This would allow us to optogenetically inhibit orbitofrontal cortex during either the distal or proximal cues whilst allowing the circuit to return to normal prior to the presentation of another cue (Gradinaru et al., 2010; Sharpe, Marchant, et al., 2017). All rats learnt the task well, increasing responding to the proximal cues on rewarded trials relative to nonrewarded trials, with no difference between the groups (**Fig 3B**). Accordingly, a mixed-design ANOVA revealed a main effect of session block (session:  $F_{(5,95)} = 23.22$ ,  $p < 0.001$ ), a main effect of trial type (rewarded vs. nonrewarded:  $F_{(1,19)} = 26.08$ ,  $p < 0.001$ ), and an interaction between trial type and session block (rewarded vs. nonrewarded  $\times$  session:  $F_{(5,95)} = 9.72$ ,  $p < 0.001$ ). Importantly, there was no main effect of group (NpHR vs. eGFP:  $F_{(1,19)} = 2.53$ ,  $p = 0.128$ ), no interaction between group and session block (NpHR vs. eGFP  $\times$  session:  $F_{(5,95)} = 1.06$ ,  $p = 0.376$ ), nor any interaction between group and trial type (NpHR vs. eGFP  $\times$  rewarded vs. nonrewarded:  $F_{(1,19)} = 1.66$ ,  $p = 0.214$ ). Finally, there was no three-way interaction between group, session block, and trial type (NpHR vs. eGFP  $\times$  session  $\times$  rewarded vs. nonrewarded:  $F_{(5,95)} = 0.781$ ,  $p = 0.57$ ). Together, this shows that all rats learnt to differentiate between rewarded and nonrewarded trials across sessions and this did not differ between groups.

Next, rats received two critical tests under extinction without reward feedback. Here, we inhibited the lateral orbitofrontal cortex by delivering green light into the brain (11s; 16mW; 532nm) during the proximal cues on one test (termed, proximal test) and during the distal cues on the other test (termed, distal test). Rats received reminder training with the original contingencies between sessions. During the **proximal test** we found that our eGFP control group showed greater responding on previously rewarded trials relative to nonrewarded trials, even in the absence of reward feedback (**Fig 3C**). This showed effective recall of the past trial structure. However, rats in the NpHR group failed to show differentiation between rewarded and unrewarded trials. This was confirmed with statistical analyses using a repeated measures mixed ANOVA, which showed a main effect of trial type (rewarded vs. unrewarded:  $F_{(1,19)} = 7.96$ ,  $p = 0.011$ ), an interaction between group and trial type (NpHR vs. eGFP  $\times$  rewarded vs. unrewarded:  $F_{(1,19)} = 3.49$ ,  $p = 0.039$ ). Follow up simple main effects analyses warranted by this interaction showed that there was a significant effect of trial type in the eGFP control group (rewarded vs. nonrewarded:  $F_{(1,19)} = 12.83$ ,  $p = 0.002$ ), but not in the NpHR

group (rewarded vs. nonrewarded:  $F_{(1,19)} = 0.40$ ,  $p = 0.536$ ). Further, there was no main effect of group (NpHR vs. eGFP:  $F_{(1,19)} = 0.98$ ,  $p = 0.336$ ), which demonstrates that inhibition of orbitofrontal cortex did not reduce responding *per se*. As inhibition or lesions of orbitofrontal cortex do not usually disrupt responding to a rewarded cue in simple discrimination task (Pickens et al., 2003; Schoenbaum et al., 2003), this suggests that our effect was due to the complex structure promoted by the task. This is consistent with our model, which argues that orbitofrontal cortex is important for tracking through the cognitive map in a manner that requires integration of past information that is no longer observable to make novel inferences about upcoming rewards (Sharpe et al., 2019).

Finally, during the **distal test**, we saw that rats were again able to differentiate between rewarded and nonrewarded trials (**Fig 3D**). In this case, there was no difference between our NpHR or eGFP groups. In line with this, a repeated-measure mixed-design ANOVA showed a main effect of trial type (rewarded vs. nonrewarded:  $F_{(1,19)} = 26.11$ ,  $p < 0.001$ ). Here, there was no interaction between group and trial type (NpHR vs. eGFP x rewarded vs. nonrewarded:  $F_{(1,19)} = 1.91$ ,  $p = 0.183$ ) or any main effect of group (NpHR vs. eGFP:  $F_{(1,19)} = 1.82$ ,  $p = 0.194$ ). This showed that all rats were able to distinguish between rewarded and nonrewarded trials, even in the absence of reward feedback, and this was not impacted by inhibition of the orbitofrontal cortex. This shows that the orbitofrontal cortex is not necessary to upload the cognitive map when the distal cue is presented to dictate the contingencies that are in effect on the trial.

## **Optogenetic inhibition of the prelimbic cortex during the distal cue, but not the proximal cue, attenuates responding**

Inhibition of the orbitofrontal cortex during the proximal cue— but not the distal cue— abolished differential responding in the differential biconditional discrimination task. This is consistent with the idea that the orbitofrontal cortex tracks the position within a cognitive map that requires integration of past information that is no longer observable or must be inferred. Next, we were interested in testing whether the prelimbic cortex would be necessary for the complementary process within our task, which is to upload the relevant cognitive map at the start of the trial (Sharpe et al., 2019), which occurs during presentation of the distal cue that dictates which contingencies are in effect.

To test this, rats received infusion of an AAV carrying NpHR on a CaMKII $\alpha$  promoter ( $n = 10$ ) or a control AAV carrying eGFP ( $n = 10$ ) into the prelimbic cortex, with optic fibers implanted in the prelimbic cortex (**Fig 4A**). All rats learned the task effectively, differentiating between the rewarded and nonrewarded trials as training progressed (**Fig 4B**). In line with this, a repeated-measures mixed-design ANOVA showed a main effect of trial type (rewarded vs. nonrewarded:  $F_{(1,18)} = 30.01$ ,  $p < 0.001$ ), a main effect of session block (session:  $F_{(4,72)} = 24.70$ ,  $p < 0.001$ ), and a significant interaction between trial type and session block (rewarded vs. nonrewarded x session:  $F_{(4,72)} = 10.17$ ,  $p < 0.001$ ). Importantly, there was no main effect of group (group:  $F_{(1,18)} = 0.204$ ,  $p = 0.657$ ), no interaction between group and trial type (NpHR vs. eGFP x rewarded vs. nonrewarded:  $F_{(1,19)} = 0.638$ ,  $p = 0.435$ ), nor between group and session block (NpHR vs. eGFP x session:  $F_{(4,72)} = 0.444$ ,  $p = 0.777$ ), and no three-way interaction between group, session block, and trial type (NpHR vs. eGFP x

session x rewarded vs. nonrewarded:  $F_{(4,72)} = 0.350$ ,  $p = 0.843$ ). This confirmed that all rats distinguished between rewarded and nonrewarded trials across the sessions, with no difference between groups.

Next, rats received two critical tests, the proximal test and the distal test. For the proximal test, we again delivered light into the brain (11s; 16mW; 532nm) during the proximal cue and not the distal cue. Here, rats were presented with the cue contingencies from training without reward feedback. We found that all rats were able to differentiate between the rewarded and nonrewarded trials and this was unaffected by inhibition of the prelimbic cortex in our NpHR group (**Fig 4C**). Consistent with this, statistical analyses showed a main effect of trial type (rewarded vs. nonrewarded:  $F_{(1,18)} = 32.40$ ,  $p < 0.001$ ), with no interaction between group and trial type (NpHR vs. eGFP x rewarded vs. nonrewarded:  $F_{(1,18)} = 1.16$ ,  $p = 0.229$ ), nor any main effect of group (NpHR vs. eGFP:  $F_{(1,18)} = 0.31$ ,  $p = 0.587$ ). This confirmed that all rats showed discrimination between previously rewarded and nonrewarded trials in the absence of reward feedback, demonstrating that they could recall the accurate associative structure of the task.

During the distal test, we instead delivered light into the brain during the distal cues and not the proximal cues. Rats again received the original cue contingencies in the absence of reward feedback. Here, we found that both groups showed discrimination of the rewarded and unrewarded trials (**Fig 4D**). However, the NpHR group showed reduced responding. This was confirmed with statistical analyses, which returned a main effect of trial type (rewarded vs. nonrewarded:  $F_{(1,18)} = 17.85$ ,  $p < 0.001$ ), no interaction between group and trial type (NpHR vs. eGFP x rewarded vs. nonrewarded:  $F_{(1,18)} = 0.40$ ,  $p = 0.54$ ), but a significant main effect of group (NpHR vs. eGFP:  $F_{(1,18)} = 4.88$ ,  $p = 0.04$ ). This suggested that rats without prelimbic activity during the distal cue could still accurately discriminate between the rewarded and nonrewarded trials, but the vigour with which they executed this responding was significantly attenuated. Further, we also saw evidence that the cognitive map was disrupted when examining responding to the proximal cues depending on their predicted outcome (Figure S2). Specifically, our eGFP control group showed greater responding to the proximal cue predicting pellets, relative to the proximal cue predicting sucrose. This indicates that the response is dependent on a specific outcome representation (as backpropagation of different scalar values could not solve the discrimination). However, optical inhibition of the prelimbic cortex removed the outcome specificity of this responding, where rats without prelimbic function during the distal cue showed similar responding to the proximal cue predicting sucrose and pellets. This provides further evidence that the cognitive map was distorted when we inhibited prelimbic cortex during the distal cues.

We next tested whether there was a relationship between the way that rats responded during the distal test relative to training. That is, we were interested in examining whether rats without prelimbic neuronal activity during the distal cue were exhibiting differential responding at test via a different mechanism than during training, consistent with the lack of outcome specificity seen in the NpHR group's responding to the proximal cues (Figure S2). If responding between the training and test are not positively correlated with one another, this would suggest that inhibition of prelimbic cortex led rats to rely on a different mechanism to show flexible behaviour. However, if performance during training and test are positive correlated, this would suggest that the prelimbic cortex is merely boosting performance to increase responding in a linear manner, rather than facilitating flexible behaviour *per se*. Here, we found that our eGFP group showed a strong positive

correlation in responding between training and test, indicating they were similar system at training and test (**Fig 4E**). However, rats in our NpHR group did not show this positive correlation, suggesting they relied on different systems during training and test. This demonstrates that prelimbic cortex contributes to the ability to pull up the current contingencies dictated by the distal cues, but that other systems are also capable of supporting discriminative responding once the task has been learnt. In line with the latter hypothesis, we found that while there was a strong positive correlation between discriminative performance during training and test in the eGFP control group ( $R^2=0.607$ ,  $F_{(1,8)}=12.340$ ,  $p=.008$ ), this was not the case for our NpHR group ( $R^2=0.002$ ,  $F_{(1,8)}=0.018$ ,  $p=0.895$ ). This demonstrated that while our NpHR group showed a reduction in responding but were able to discriminate between the rewarded and nonrewarded trials without PL activity during the distal cues, they were relying on a different neural and/or associative mechanism to do so.

## Discussion

We hypothesised that prelimbic cortex would be important for uploading a current cognitive map, while the orbitofrontal cortex would be important for tracking current movement through it (Sharpe et al., 2019). To test this, we used a sequential differential biconditional discrimination (Bradfield & Balleine, 2013; Colwill & Delamater, 1995; Delamater et al., 2010) to engender a cognitive map. Our task was sensitive to the differential outcomes effect, which showed that performance on the task relied on detailed representations of associative links between cue and reward identities (Daw et al., 2005; Dayan, 1993; Gershman, 2018; McCloskey & Cohen, 1989), and cannot be accounted for by a model-free mechanisms that lacks such an associative link between these events and the ability to represent subsequent states (Daw et al., 2005). Next, we demonstrated that this task supported model-based inference when new information was introduced. Thus, our task supported the development of a cognitive map, which engendered rats with the ability to make novel inferences in new situations.

We then tested the role of the prelimbic and orbitofrontal cortex in use of this cognitive map. We hypothesized that optogenetic inhibition of the prelimbic cortex during the distal cue—when the current contingencies are signaled (i.e., the current map is revealed)—would impair behavior, but inhibition during the proximal cues would be without effect. In contrast, we predicted that inhibition of the orbitofrontal cortex during the proximal cues would impair discrimination as this is when rats must recall the prior distal cue and integrate that information with the current proximal cue to predict reward. Indeed, we found that inhibition of orbitofrontal cortex during the proximal cue, but not the distal cue, impaired discrimination. Taken with data showing that inhibition does not affect discrimination between rewarded and unrewarded cues *per se* (Pickens et al., 2003; Schoenbaum et al., 2003) or even complex choices between cues predicting different identities or magnitudes of rewards (Gardner et al., 2017), this was consistent with our hypothesis that orbitofrontal cortex is necessary when rats are required to use information that is no longer available or must be inferred to predict reward (Sharpe et al., 2019). That is, the signal from orbitofrontal cortex would not be necessary to perform a simple discrimination where all the information is observable, unless something about the task had changed that requires integration of unobservable information [e.g., a devaluation procedure; (Gardner et al., 2017; Pickens et al., 2003)].

The effects with prelimbic cortex were more complex. While we found that inhibition of the prelimbic cortex during the distal cue significantly reduced responding, rats were still capable of showing discrimination between previously rewarded and nonrewarded trials. Yet when we examined behaviour more closely, we found that our NpHR group did not show a correlation between discrimination performance during training and at test, unlike our eGFP control group. Further, we saw that inhibition of prelimbic cortex during the distal cue prevented rats from demonstrating outcome specificity in their responding to the proximal cues (Figure S2), indicating disruption of the specific components of the cognitive map. This suggested that rats without prelimbic function during the distal cue were forced to rely on a different associative and/or neural mechanism to show accurate discrimination. This sheds light on data that reveal that the prelimbic cortex is necessary for development of complex or hierarchical associations but no longer necessary once the task is well learned (Ostlund & Balleine, 2005). That is, while other regions might be able to compensate when the

task is well learned, the prelimbic cortex still plays a role in uploading the relevant information to facilitate correct performance and amplify specific motivated behaviour.

The effects found with inhibition of prelimbic cortex were also not consistent with disruption of a simple attentional process, whereby prelimbic inhibition stops the rats from processing the cues. If attention to the cues *per se* was a key factor driving performance, inhibition during the proximal cue should disrupt performance just as much as inhibition during the distal cue. Instead, the more complex role of the prelimbic cortex in attentional modulation based on higher-order information (Birrell & Brown, 2000; Floresco et al., 2008; Sharpe & Killcross, 2014; M. J. Sharpe & S. Killcross, 2015; Sharpe & Killcross, 2018), can be accommodated within the current framework being proposed (Sharpe et al., 2019). Specifically, we would view the role of the prelimbic cortex in using the distal cues to upload the relevant cognitive map as analogous to its role in modulating the degree to which a cue evokes attention as dictated by higher-order information [e.g., how well a cue predicts an outcome in the current circumstance; (Sharpe & Killcross, 2018; Sharpe et al., 2019)]. Thus, the relationship between attention and prelimbic function is not straightforward, but readily accommodated within our current framework.

Of course, it is likely that rats learn a variety of associations and behaviour involves a blend of such learning. For example, biconditional discriminations support hierarchical and configural associations (Bradfield & Balleine, 2013; Delamater et al., 2017). However, if rats were using a configural association, it is not clear why we would find a difference in the contribution of the neural regions at the distal and proximal timepoints. If sequential trials were treated as one configural stimulus, then why would the prelimbic and orbitofrontal frontal cortices not contribute at both timepoints, if at all? Indeed, the orbitofrontal and prelimbic cortices are not important for the simple discriminations (Pickens et al., 2003; Schoenbaum et al., 2003; Sharpe & Killcross, 2018). If rats were treating the compounds as configural stimuli and simply need to discriminate between them, inhibition of the orbitofrontal and prelimbic cortices should be without effect. Further, if our task was solved through the development of configural associations it should not require differential outcomes as configural associations can be learnt through the backpropagation of scalar value (Delamater et al., 2017; Delamater et al., 2010). Finally, we found that rats were able to integrate new information into the previously learnt associative structure developed during the original training. If the distal cues were functioning purely as an occasion setter that exerts hierarchical control over the association between the proximal cue and the specific outcomes, then this associative structure should show very little or no transfer to a novel cue-outcome pairing (Bonardi, 2007; Fraser & Holland, 2019; Roughley & Killcross, 2019). Thus, our results are inconsistent with either a configural or hierarchical account.

Another line of theories could explain our results via associative summation [(e.g., (Harris, 2006; McLaren & Mackintosh, 2002)]. Here, cues activate representations of associated elements. For example, presentation of one distal cue can evoke a representation of the associated reward, as can the associated proximal cue that leads to that same reward. Thus, these two representations can summate to produce greater responding when they predict the same outcome. This theory explains the differential outcomes effect as the patterns of activation would become more distinct with different outcome representations (Harris, 2006). However, associative summation would not explain the dissociation we find with the prelimbic and

orbitofrontal cortices, as there should be no change in the representation of proximal and distal cues that would necessitate their being encoded differently. Indeed, such accounts are generally premised on this idea that activation of related units occur via neighboring nodes (McLaren & Mackintosh, 2002). Secondly, our analyses showed that distal and proximal cues influenced behaviour differently; rats showed more responding to the proximal cue predicting pellets relative to sucrose (Figures S1 and S2), which shows sensory-specific responding to the proximal cue. However, we did not see this with the distal cue. This is inconsistent with an associative summation account that would predict that both the distal and proximal cues should evoke a representation of their associated outcome and, thus, differential responding. Taken together, we would advocate for our task developing an associative chain that allows rats to flexibly navigate through the elemental cognitive space to predict reward.

Several questions remain. In our original model, we argued that the prelimbic and orbitofrontal cortices exerted control over responding via the dorsomedial striatum (Sharpe et al., 2019), which contains the state transition matrix of the task (Daw et al., 2005). When the distal cue appears, the prelimbic cortex exerts top-down modulation over striatum to determine which state transition matrix in effect, so that it dominates behaviour on that trial. The orbitofrontal cortex then indicates to the striatum which state within the matrix that rat currently occupies, allowing the rat's behaviour to reflect the position within the trial. This suggests that we should recapitulate the dissociable effects found with prelimbic and orbitofrontal cortices via inhibition of their respective projections to the dorsomedial striatum. Another open question is whether the prelimbic cortex continues to provide this information to the dorsomedial striatum once the task is well learned. The present data, which reveal a significant reduction in responding to the cues—and loss of outcome selectivity—when inhibition of prelimbic cortex occurs at the distal timepoint, suggest at the very least that the prelimbic cortex continues to provide input to the striatum to boost correct responding. It would also be of interest to test whether inhibiting these regions during model-based inference produces a similar dissociation. We would predict this to produce a greater reliance on prelimbic and orbitofrontal cortex as the associations are less certain and require strong modulation by prefrontal regions that cannot be compensated for by striatal mechanisms. These are all future questions of interest to test the validity of our wider model, which have important implications for theories of decision making and our understanding of what happens when these systems have gone awry, such as in disorders like addiction (Wied et al., 2013) and schizophrenia (Culbreth et al., 2016).

## Acknowledgements

This work was supported by NHMRC Investigator Grant (2024/GNT2034225, M.J.S), NIH Research Project grant (R01DA054967, M.J.S.), NIH Research Project grant (R01DA057084, M.J.S), and a UEHARA Memorial Overseas Research Fellowship (M.T.)

## Financial Disclosures

The authors do not have any financial disclosures or conflicts to declare.

## References

- Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, 100(2), 490-509.
- Birrell, J. M., & Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *The Journal of neuroscience*, 20(11), 4320-4324.
- Bonardi, C. (2007). Occasion setting is specific to the CS-US association. *Learning and Motivation*, 38(3), 208-228.
- Bradfield, L. A., & Balleine, B. W. (2013). Hierarchical and binary associations compete for behavioral control during instrumental biconditional discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 39(1), 2.
- 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.
- 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.
- Colwill, R. M., & Delamater, B. A. (1995). An associative analysis of instrumental biconditional discrimination learning. *Animal Learning & Behavior*, 23(2), 218-233.
- Costa, K. M., Scholz, R., Lloyd, K., Moreno-Castilla, P., Gardner, M. P., Dayan, P., & Schoenbaum, G. (2023). The role of the lateral orbitofrontal cortex in creating cognitive maps. *Nature Neuroscience*, 26(1), 107-115.
- Culbreth, A. J., Westbrook, A., Daw, N. D., Botvinick, M., & Barch, D. M. (2016). Reduced model-based decision-making in schizophrenia. *Journal of abnormal psychology*, 125(6), 777.

- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*, 8(12), 1704-1711.
- Dayan, P. (1993). Improving generalization for temporal difference learning: The successor representation. *Neural computation*, 5(4), 613-624.
- Delamater, A. R., Garr, E., Lawrence, S., & Whitlow Jr, J. W. (2017). Elemental, configural, and occasion setting mechanisms in biconditional and patterning discriminations. *Behavioural processes*, 137, 40-52.
- Delamater, A. R., Kranjec, A., & Fein, M. I. (2010). Differential outcome effects in pavlovian biconditional and ambiguous occasion setting tasks. *Journal of Experimental Psychology: Animal Behavior Processes*, 36(4), 471.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175-191.
- Floresco, S. B., Block, A. E., & Maric, T. (2008). Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behavioural brain research*, 190(1), 85-96.
- Fraser, K. M., & Holland, P. C. (2019). Occasion setting. *Behavioral neuroscience*, 133(2), 145.
- Gardner, M. P., Conroy, J. S., Shaham, M. H., Styer, C. V., & Schoenbaum, G. (2017). Lateral orbitofrontal inactivation dissociates devaluation-sensitive behavior and economic choice. *Neuron*, 96(5), 1192-1203. e1194.
- Gershman, S. J. (2018). The successor representation: its computational logic and neural substrates. *Journal of neuroscience*, 38(33), 7193-7200.
- Golden, C. E., Martin, A. C., Kaur, D., Mah, A., Levy, D. H., Yamaguchi, T., Lasek, A. W., Lin, D., Aoki, C., & Constantinople, C. M. (2025). Estrogen modulates reward prediction errors and reinforcement learning. *Nature neuroscience*, 1-13.
- Gradinaru, V., Zhang, F., Ramakrishnan, C., Mattis, J., Prakash, R., Diester, I., Goshen, I., Thompson, K. R., & Deisseroth, K. (2010). Molecular and cellular approaches for diversifying and extending optogenetics. *Cell*, 141(1), 154-165.

- Granon, S., & Poucet, B. (2000). Involvement of the rat prefrontal cortex in cognitive functions: a central role for the prelimbic area. *Psychobiology*, 28(2), 229-237.
- Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological review*, 113(3), 584.
- Harris, J. A., & Livesey, E. J. (2008). Comparing patterning and biconditional discriminations in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, 34(1), 144.
- Hart, E. E., Sharpe, M. J., Gardner, M. P., & Schoenbaum, G. (2020). Responding to preconditioned cues is devaluation sensitive and requires orbitofrontal cortex during cue-cue learning. *elife*, 9, e59998.
- Howell, D. C. (1992). *Statistical methods for psychology*. PWS-Kent Publishing Co.
- Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., & Schoenbaum, G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science*, 338(6109), 953-956.
- Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic bulletin & review*, 1(4), 476-490.
- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation* (Vol. 24, pp. 109-165). Elsevier.
- McLaren, I., & Mackintosh, N. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal learning & behavior*, 30(3), 177-200.
- Ostlund, S. B., & Balleine, B. W. (2005). Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *Journal of neuroscience*, 25(34), 7763-7770.
- 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.
- Roughley, S., & Killcross, S. (2019). Loss of hierarchical control by occasion setters following lesions of the prelimbic and infralimbic medial prefrontal cortex in rats. *brain sciences*, 9(3), 48.
- Schiereck, S. S., Pérez-Rivera, D. T., Mah, A., DeMaegd, M. L., Hocker, D., Ward, R. M., Savin, C., & Constantinople, C. M. (2026). The orbitofrontal cortex updates beliefs for state inference. *Neuron*, 114(3), 507-520. e508.

- 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.
- Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91(6), 1402-1412.
- Schwartenbeck, P., Baram, A., Liu, Y., Mark, S., Muller, T., Dolan, R., Botvinick, M., Kurth-Nelson, Z., & Behrens, T. (2023). Generative replay underlies compositional inference in the hippocampal-prefrontal circuit. *Cell*, 186(22), 4885-4897. e4814.
- Seitz, B. M., Hoang, I. B., DiFazio, L. E., Blaisdell, A. P., & Sharpe, M. J. (2022). Dopamine errors drive excitatory and inhibitory components of backward conditioning in an outcome-specific manner. *Current Biology*, 32(14), 3210-3218. e3213.
- Sharpe, M., & Killcross, S. (2015). The prelimbic cortex uses contextual cues to modulate responding towards predictive stimuli during fear renewal. *Neurobiology of learning and memory*, 118, 20-29.
- Sharpe, M. J., Batchelor, H. M., Mueller, L. E., Gardner, M. P., & Schoenbaum, G. (2021). Past experience shapes the neural circuits recruited for future learning. *Nature Neuroscience*, 24(3), 391-400.
- Sharpe, M. J., Batchelor, H. M., Mueller, L. E., Yun Chang, C., Maes, E. J., Niv, Y., & Schoenbaum, G. (2020). Dopamine transients do not act as model-free prediction errors during associative learning. *Nature communications*, 11(1), 106.
- Sharpe, M. J., Chang, C. Y., Liu, M. A., Batchelor, H. M., Mueller, L. E., Jones, J. L., Niv, Y., & Schoenbaum, G. (2017). Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience*, 20(5), 735-742.
- Sharpe, M. J., & Killcross, S. (2014). The prelimbic cortex contributes to the down-regulation of attention toward redundant cues. *Cerebral cortex*, 24(4), 1066-1074.
- Sharpe, M. J., & Killcross, S. (2015). The prelimbic cortex directs attention toward predictive cues during fear learning. *Learning & Memory*, 22(6), 289-293.
- Sharpe, M. J., & Killcross, S. (2018). Modulation of attention and action in the medial prefrontal cortex of rats. *Psychological review*, 125(5), 822.
- Sharpe, M. J., Marchant, N. J., Whitaker, L. R., Richie, C. T., Zhang, Y. J., Campbell, E. J., Koivula, P. P., Necarsulmer, J. C., Mejias-Aponte, C., & Morales, M. (2017). Lateral hypothalamic GABAergic neurons

encode reward predictions that are relayed to the ventral tegmental area to regulate learning. *Current Biology*, 27(14), 2089-2100. e2085.

Sharpe, M. J., & Schoenbaum, G. (2018). Evaluation of the hypothesis that phasic dopamine constitutes a cached-value signal. *Neurobiology of learning and memory*, 153, 131-136.

Sharpe, M. J., Stalnaker, T., Schuck, N. W., Killcross, S., Schoenbaum, G., & Niv, Y. (2019). An integrated model of action selection: distinct modes of cortical control of striatal decision making. *Annual review of psychology*, 70(1), 53-76.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, 55(4), 189.

Vendrell, P., Junqué, C., Pujol, J., Jurado, M. A., Molet, J., & Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia*, 33(3), 341-352.

Veselic, S., Gutierrez, E., El-Gaby, M., Reinert, S., & Sablé-Meyer, M. (2025). Cognitive maps in the prefrontal cortex. *Journal of neuroscience*, 45(46).

Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, 411(6840), 953-956.

Whittington, J. C., Dorrell, W., Behrens, T. E., Ganguli, S., & El-Gaby, M. (2025). A tale of two algorithms: Structured slots explain prefrontal sequence memory and are unified with hippocampal cognitive maps. *Neuron*, 113(2), 321-333. e326.

Whittington, J. C., McCaffary, D., Bakermans, J. J., & Behrens, T. E. (2022). How to build a cognitive map. *Nature Neuroscience*, 25(10), 1257-1272.

Wied, H. M., Jones, J. L., Cooch, N. K., Berg, B. A., & Schoenbaum, G. (2013). Disruption of model-based behavior and learning by cocaine self-administration in rats. *Psychopharmacology*, 229(3), 493-501.

Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Reviews Neuroscience*, 17(8), 513-523.

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.

## Figure Legends

**Figure 1. Our sequential biconditional discrimination task can only be learnt with differential outcomes.** In order to test whether our task involved the development of a rich associative structure that involves sensory-specific representation of outcomes (i.e., the differential outcomes effect), we tested if rats would learn the task more effectively when each proximal cue is paired with different outcomes. **A)** One group of rats ( $n=8$ ) received the biconditional discrimination with the same outcomes (termed, same group). Here, a trial preceded with one of two distal auditory cues (i.e., white noise or click; counterbalanced). These distal cues signal to the rat the contingencies that are in effect. Then, one of two proximal cues is presented (i.e., tone or siren). Whether the proximal cue leads to reward depends on the distal cue. For example, the white noise signals that the tone will lead to reward, while the siren will not (white noise  $\rightarrow$  tone  $\rightarrow$  reward; white noise  $\rightarrow$  siren  $\rightarrow$  no reward). In contrast, the click signals that the siren will be rewarded, but the tone will not (click  $\rightarrow$  siren  $\rightarrow$  reward; click  $\rightarrow$  tone  $\rightarrow$  no reward). All rewarded trials in the same outcomes group resulted in delivery of grain pellets. **B)** In the differential group ( $n=8$ ), rats received the same contingencies with the exception that one proximal cue was rewarded with grain pellets and the other was rewarded with sucrose. **C)** Rats in the same outcomes group failed to learn the discrimination, while rats in the differential outcomes group learnt to discriminate between the rewarded and unrewarded trials. Data are shown using the mean number of food port entries during the 10s proximal cue ( $\pm$ SEM), with 4 sessions per session block. \*\*indicate  $p<0.01$ .

**Figure 2. Our sequential biconditional task supports an associative structure that supports inference.** If rats develop a rich associative structure that resembles a cognitive map during our biconditional discrimination task, then they should be able to use this associative structure to make novel inferences. To test this, we asked whether rats would be able to integrate a new set of proximal cues into their cognitive map and infer the relative likelihood of reward. **A)** A new cohort of rats ( $n=24$ ) learnt the differential biconditional task. All rats could develop the discrimination, showing more responding on rewarded relative to unrewarded proximal cues across session blocks. **B)** After rats successfully learnt this task, we taught them that two new partially-reinforced proximal cues led to one of the two outcomes (e.g., jingle  $\rightarrow$  pellets; horn  $\rightarrow$  sucrose; counterbalanced). Rats responded more during the 10s proximal cues relative to the 10s pre-cue period (i.e., baseline). **C)** In the critical inference test, rats were presented with four different trial types. On half of the trials, the novel distal and proximal cues combinations predicted the same reward (termed congruent; pink), while on the other trials the novel distal and proximal cue combinations predicted different rewards (termed, incongruent; blue). This demonstrated that rats could recall the past associative structure from the original contingencies and integrate this with the new cue-outcome associations to make the novel inference that the cues predicting the same outcomes were more likely to lead to reward. Left: Group means of responding during the incongruent and congruent cues. Right: Individual difference scores between responding to congruent and incongruent cues (i.e., congruent – incongruent). Positive scores indicate more responding to the congruent cues (denoted in pink circles), while negative scores indicate more responding to the incongruent cues (denoted in blue circles). Almost all rats showed more responding to the congruent cues.

This is consistent with the idea that the task supports development of a cognitive map. Data from panel A and B are represented as the mean entries into the food port across the proximal cues ( $\pm$ SEM). Data from panel C are represented as the percent time spent in the food port during the transition between the proximal cues and expected reward periods ( $\pm$ SEM) as is standard for inference tests. There were 4 sessions per session block. \*\* indicates  $p < 0.01$ , \* indicates  $p < 0.05$ .

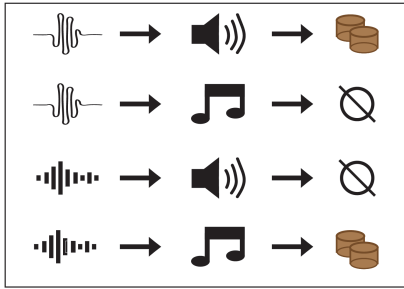
**Figure 3. Optogenetic inhibition of lateral orbitofrontal cortex during the proximal cue, but not the distal cue, disrupts performance on the sequential biconditional discrimination task.** **A)** Rats were infused with an AAV carrying NpHR on a CaMKII $\alpha$  promoter ( $n=9$ ), or a control AAV carrying eGFP ( $n=12$ ), into the lateral orbitofrontal cortex and implanted with fiber above the lateral orbitofrontal cortex. Top illustrates example eGFP expression and fiber placement in lateral orbitofrontal cortex. Bottom left indicates unilateral schematic of bilateral minimal (light) or maximal (dark) expression of eGFP (black) and NpHR (green) in lateral orbitofrontal cortex. **B)** Rats were trained on the sequential differential biconditional task where one of two distal cues were presented to dictate the current contingencies in effect (i.e., white noise or click; counterbalanced). Then, one of two proximal cues were presented (i.e., tone or siren, counterbalanced). Whether the proximal cues were rewarded depended on the preceding distal cue. For example, the white noise dictates that the tone will be rewarded with pellets, and the siren will not (white noise  $\square$  tone  $\square$  pellets; white noise  $\square$  siren  $\square$  nothing). In contrast, the click dictates that the siren will be rewarded by sucrose and the tone will not (click  $\square$  siren  $\square$  sucrose; click  $\square$  tone  $\square$  nothing). Top: Rats in the eGFP group learnt to respond more to the rewarded cues across sessions (left), and individual rats showed effective discrimination in the final session block of training (right). Bottom: Rats in the NpHR group learnt to respond more to the reward cues across sessions (left), and individual rats showed effective discrimination in the final block of training (right). **C)** During the proximal test, we presented the cues as before without reward feedback. We delivered light into the brain of rats during the proximal cue to inhibit pyramidal neurons in the orbitofrontal cortex in our NpHR group, which would be without effect in our eGFP control group. Here, we found that rats in our eGFP control group showed differentiation between previously rewarded and nonrewarded trials. This showed that our eGFP group could recall the prior contingencies. However, our NpHR group failed to differentiate between previously rewarded and nonrewarded trials. **D)** During the distal test, we again presented cue as during training without reward feedback. Here, we delivered light into the brain during the distal cues. In this test, we found that all rats could discriminate between rewarded and nonrewarded trials. This demonstrated that inhibition of lateral orbitofrontal cortex during the distal cue did not impair differential responding. Rates of responding are represented as the mean number of food port entries during the 10s proximal cue ( $\pm$ SEM), with 3 sessions per session block. \*\* indicates  $p < 0.01$ , \* indicates  $p < 0.05$ . See Figure S1 for more information.

**Figure 4. Optogenetic inhibition of prelimbic cortex during the distal cue, but not the proximal cue, disrupts performance on the sequential biconditional discrimination task.** **A)** Rats were infused with an AAV carrying NpHR on a CaMKII $\alpha$  promoter ( $n=10$ ), or a control AAV carrying eGFP ( $n=10$ ), into the

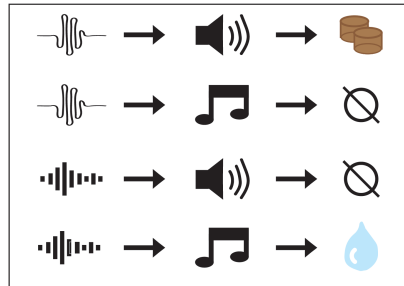
prelimbic cortex, with optic fibers implanted over the prelimbic cortex. Top illustrates example expression of eGFP expression and fiber placement in prelimbic cortex. Bottom left indicates unilateral schematic of bilateral minimal (light) or maximal (dark) expression of eGFP (black) and NpHR (green) in prelimbic cortex.

**B)** Rats were trained on the sequential differential biconditional task where one of two distal cues were presented to dictate the current contingencies in effect (i.e., white noise or click; counterbalanced). Then, one of two proximal cues were presented (i.e., tone or siren, counterbalanced). Whether the proximal cues were rewarded depended on the preceding distal cue. For example, the white noise dictates that the tone will be rewarded with pellets, and the siren will not (white noise  $\square$  tone  $\square$  pellets; white noise  $\square$  siren  $\square$  nothing). In contrast, the click dictates that the siren will be rewarded by sucrose and the tone will not (click  $\square$  siren  $\square$  sucrose; click  $\square$  tone  $\square$  nothing). Left: Rats in the eGFP group learnt to respond more to the rewarded cues across sessions (left insert), and individual rats showed effective discrimination in the final session block of training (right insert). Right: Rats in the NpHR group learnt to respond more to the reward cues across sessions (left insert), and individual rats showed effective discrimination in the final block of training (right insert). **C)** During the proximal test, we presented the cues as before without reward feedback. We delivered light into the brain of rats during the proximal cue to inhibit pyramidal neurons in the prelimbic cortex of our NpHR group, which would be without effect in our eGFP control group. Here, we found that rats in both our NpHR and eGFP control groups showed differentiation between previously rewarded and nonrewarded trials. This showed that all rats could recall the prior contingencies, despite inhibition of the prelimbic cortex in our NpHR rats during the proximal cue. **D)** During the distal test, we again presented cue as during training without reward feedback. Here, we delivered light into the brain during the distal cues. Rats in our eGFP group again showed strong discrimination. However, rats in our NpHR groups showed a significant attenuation in responding. Despite this general reduction in responding, the NpHR group were still able to discriminate between previously rewarded and nonrewarded trials. **E)** We correlated discrimination performance in training and test. Here, we found a strong positive correlation in how eGFP rats discriminated during training and test. However, we found that there was no correlation between discrimination performance in our NpHR, which demonstrates that these rats were using a different associative and/or neural mechanism to demonstrate successful discrimination when the prelimbic cortex was inhibited during the proximal cues. Rates of responding are represented as the mean number of food port entries during the 10s proximal cue ( $\pm$ SEM), with 5 sessions per session block. \*\*indicates  $p < 0.01$ , \*indicates  $p < 0.05$ . See Figure S2 for more information.

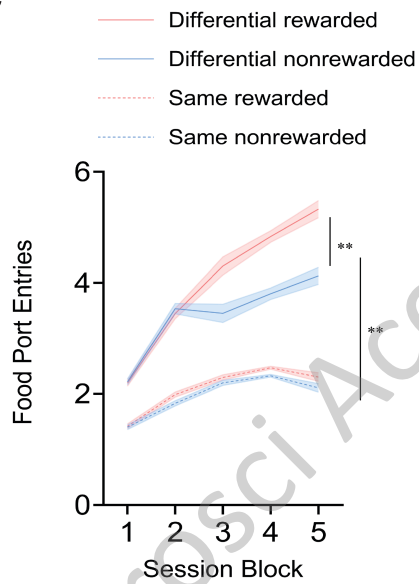
## A Same outcomes



## B Differential outcomes

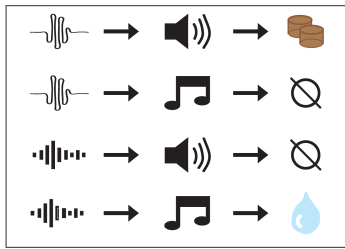


## C



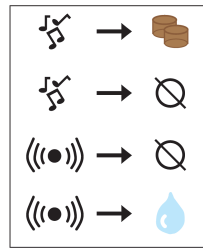
**A**

Original Contingencies



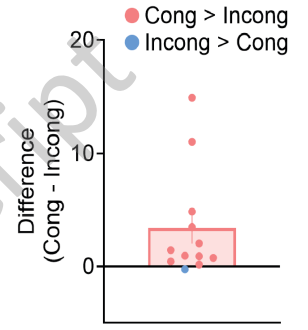
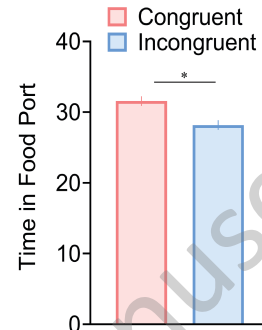
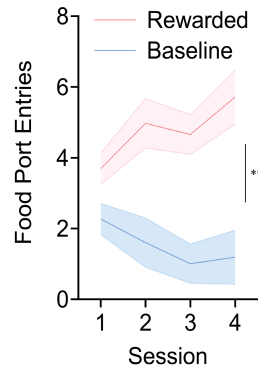
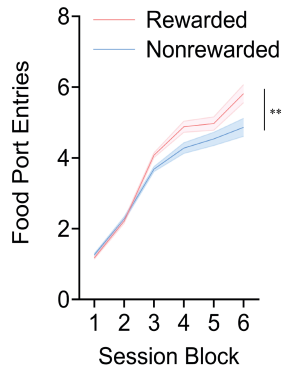
**B**

New Proximal Cues



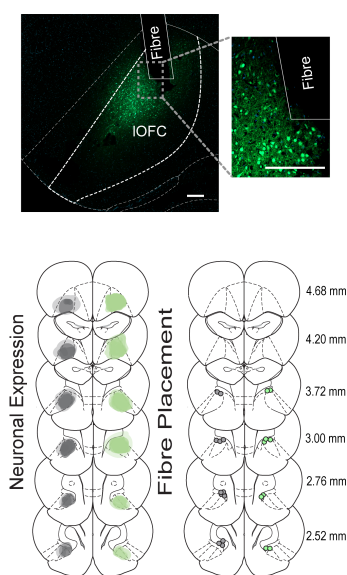
**C**

Inference Test

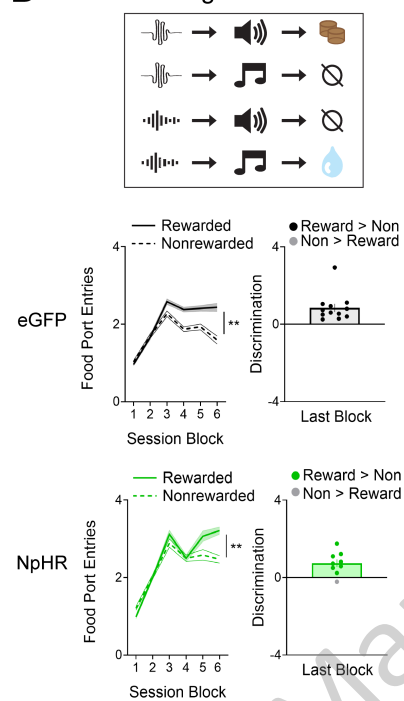


JNeurosci Accepted Manuscript

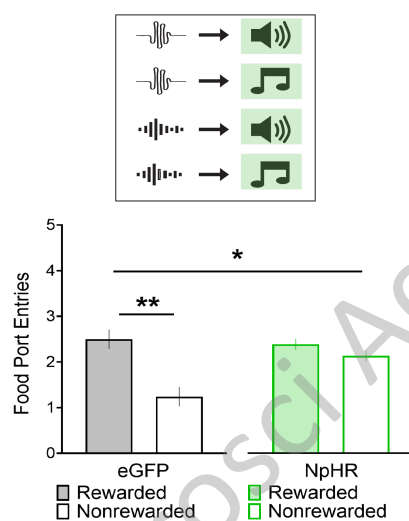
## A Histological Verification



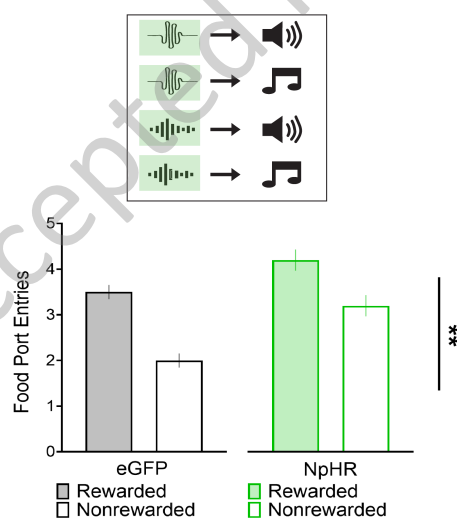
## B Training Performance



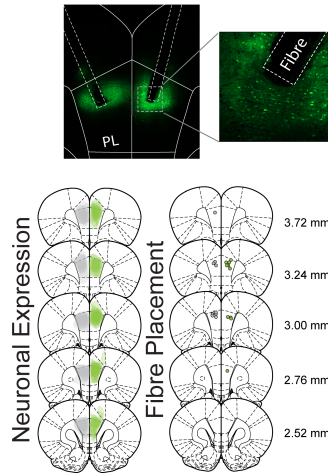
## C Proximal Test



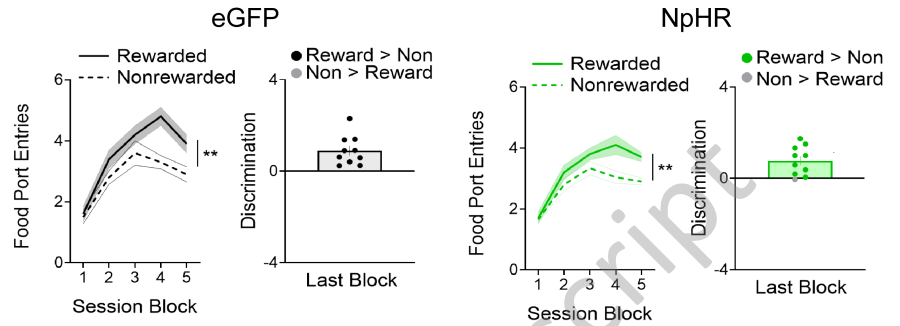
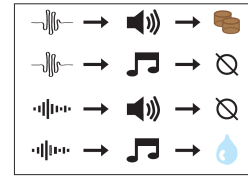
## D Distal Test



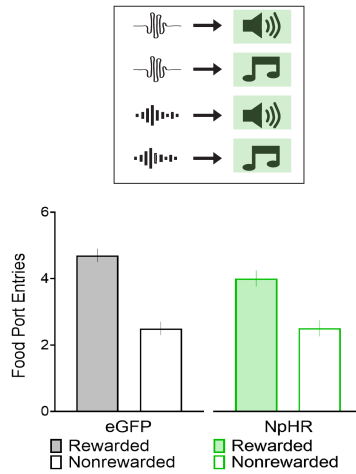
## A Histological Verification



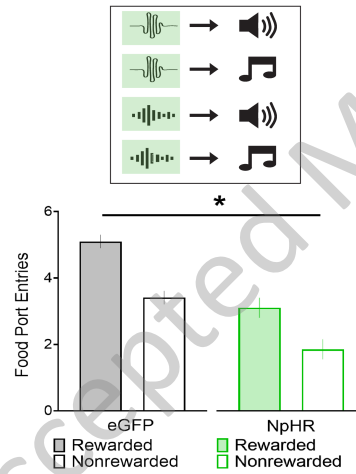
## B Training Performance



## C Proximal Test



## D Distal Test



## E Correlation: Training x Distal Test

