

# **Persistent coding of outcome-predictive cue features in the rat nucleus accumbens.**

**Authors:** Jimmie M. Gmaz<sup>1</sup>, James E. Carmichael<sup>1</sup>, Matthijs A. A. van der Meer<sup>1\*</sup>

<sup>1</sup>Department of Psychological and Brain Sciences, Dartmouth College, Hanover NH 03755

\*Correspondence should be addressed to MvdM, Department of Psychological and Brain Sciences, Dartmouth College, 3 Maynard St, Hanover, NH 03755. E-mail: mvdm@dartmouth.edu.

**Number of Figures:** 12

**Number of Tables:** 1

**Total Word Count:** ?

**Abstract Word Count:** ?

**Introduction Word Count:** ?

**Discussion Word Count:** ?

**Acknowledgments:** We thank Nancy Gibson, Martin Ryan and Jean Flanagan for animal care, and Min-Ching Kuo and Alyssa Carey for technical assistance. This work was supported by Dartmouth College (Dartmouth Fellowship to JMG and JEC, and start-up funds to MvdM) and the Natural Sciences and Engineering Research Council (NSERC) of Canada (Discovery Grant award to MvdM, Canada Graduate Scholarship to JMG).

**Conflict of Interest:** The authors declare no competing financial interests.



## **1 Abstract (eLife: 150 words)**

2 The nucleus accumbens (NAc) has been shown to be important for learning from feedback, and biasing and  
3 invigorating behavior in response to outcome-predictive cues. NAc encodes outcome-related cue features  
4 such as the magnitude and identity of reward. However, not much is known about how features of cues  
5 themselves are encoded. We designed a decision making task where rats learned multiple sets of outcome-  
6 predictive cues, and recorded single-unit activity in the NAc during performance. We found that coding  
7 of various cue features occurred alongside coding of expected outcome. Furthermore, this coding persisted  
8 during a delay period, after the rat made a decision and was waiting for an outcome, but not after the outcome  
9 was revealed. Encoding of cue features in the NAc may enable contextual modulation of ongoing behavior,  
10 and provide an eligibility trace of outcome-predictive stimuli for updating stimulus-outcome associations to  
11 inform future behavior.

## **12 Introduction**

13 Theories of nucleus accumbens (NAc) function generally agree that this brain structure contributes to moti-  
14 vated behavior, with some emphasizing a role in learning from RPEs (Averbeck and Costa, 2017; Joel et al.,  
15 2002; Khamassi and Humphries, 2012; Lee et al., 2012; Maia, 2009; Schultz, 2016)(see also the addiction  
16 literature on effects of drug rewards; Carelli, 2010; Hyman et al., 2006; Kalivas and Volkow, 2005) and  
17 others a role in the modulation of ongoing behavior through stimuli associated with motivationally relevant  
18 outcomes (invigorating, directing; Floresco, 2015; Nicola, 2010; Salamone and Correa, 2012). These pro-  
19 posals echo similar ideas on the functions of the neuromodulator dopamine (Berridge, 2012; Maia, 2009;  
20 Salamone and Correa, 2012; Schultz, 2016), with which the NAc is tightly linked functionally as well as  
21 anatomically (Cheer et al., 2007; du Hoffmann and Nicola, 2014; Ikemoto, 2007; Takahashi et al., 2016).

22 Much of our understanding of NAc function comes from studies of how cues that predict motivationally  
23 relevant outcomes (e.g. reward) influence behavior and neural activity in the NAc. Task designs that asso-  
24 ciate such cues with rewarding outcomes provide a convenient access point eliciting conditioned responses  
25 such as sign-tracking and goal-tracking (Hearst and Jenkins, 1974; Robinson and Flagel, 2009), pavlovian-  
26 instrumental transfer (Estes, 1943; Rescorla and Solomon, 1967) and enhanced response vigor (Nicola, 2010;  
27 Niv et al., 2007), which tend to be affected by NAc manipulations (Chang et al., 2012; Corbit and Balleine,  
28 2011; Flagel et al., 2011) (although not always straightforwardly; Giertler et al., 2004; Chang and Holland,  
29 2013). Similarly, analysis of RPEs typically proceeds by establishing an association between a cue and sub-  
30 sequent reward, with NAc responses transferring from outcome to the cue with learning (Day et al., 2007;  
31 Roitman et al., 2005; Setlow et al., 2003; Schultz et al., 1997).

32 Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues  
33 (Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Goldstein et al., 2012;  
34 Hollerman et al., 1998; Lansink et al., 2012; McGinty et al., 2013; Nicola, 2004; Roesch et al., 2009; Roit-  
35 man et al., 2005; Saddoris et al., 2011; Setlow et al., 2003; Sugam et al., 2014; West and Carelli, 2016),

36 much less is known about how outcome-predictive cues themselves are encoded in the NAc (Sleeker et al.,  
37 2016). This is an important issue for at least two reasons. First, in reinforcement learning, motivationally  
38 relevant outcomes are typically temporally delayed relative to the cues that predict them. In order to solve  
39 the problem of assigning credit (or blame) across such temporal gaps, some trace of preceding activity needs  
40 to be maintained (Lee et al., 2012; Sutton and Barto, 1998). For example, if you become ill after eating  
41 food X in restaurant A, depending on if you remember the identity of the restaurant or the food at the time  
42 of illness, you may learn to avoid all restaurants, restaurant A only, food X only, or the specific pairing of  
43 X-in-A. Therefore, a complete understanding of what is learned following feedback requires understanding  
44 what trace is maintained. Since NAc is a primary target of DA signals interpretable as reward prediction  
45 errors (RPEs), and NAc lesions impair RPEs related to timing, its activity trace will help determine what can  
46 be learned when RPEs arrive (Hamid et al., 2015; Hart et al., 2014; Ikemoto, 2007; McDannald et al., 2011;  
47 Takahashi et al., 2016).

48 Second, for ongoing behavior, the relevance of cues typically depends on context. In experimental settings,  
49 context may include the identity of a preceding cue, spatial or configural arrangements (Holland, 1992;  
50 Bouton, 1993; Honey et al., 2014), and unsignaled rules as occurs in set shifting and other cognitive control  
51 tasks (Grant and Berg, 1948; Cohen and Servan-Schreiber, 1992; Floresco et al., 2006; Sleeker et al., 2016).  
52 In such situations, the question arises how selective, context-dependent processing of outcome-predictive  
53 cues is implemented. For instance, is there a gate prior to NAc, or are all cues represented in NAc but  
54 their current values dynamically updated (Goto and Grace, 2008; FitzGerald et al., 2014; Sleeker et al.,  
55 2016). Representation of cue identity would allow for context-dependent mapping of outcomes predicted by  
56 specific cues.

57 Thus, both from a learning and a flexible performance perspective, it is of interest to determine how cue  
58 identity is represented in the brain, with NAc of particular interest given its anatomical and functional po-  
59 sition at the center of motivational systems. We sought to determine whether cue identity is represented in  
60 the NAc, if cue identity is represented alongside other motivationally relevant variables, such as cue value,

61 and if these representations are maintained after a behavioral decision has been made (Figure 1). To address  
62 these questions, we recorded the activity of NAc units as rats performed a task in which multiple, distinct  
63 sets of cues predicted the same outcome.

64 [Figure 1 about here.]

## 65 **Methods**

### 66 **Subjects:**

67 Adult male Long-Evans rats ( $n = 4$ , Charles River, Saint Constant, QC) were used as subjects. Rats were in-  
68 dividually housed with a 12/12-h light-dark cycle, and tested during the light cycle. Rats were food deprived  
69 to 85-90% of their free feeding weight (weight at time of implantation was 440 - 470 g), and water restricted  
70 4-6 hours before testing. All experimental procedures were approved by the the University of Waterloo An-  
71 imal Care Committee (protocol# 11-06) and carried out in accordance with Canadian Council for Animal  
72 Care (CCAC) guidelines.

### 73 **Overall timeline:**

74 Each rat was first handled for seven days during which they were exposed to the experiment room, the  
75 sucrose solution used as a reinforcer, and the click of the sucrose dispenser valves. Rats were then trained  
76 on the behavioral task (described in the next section) until they reached performance criterion. At this point  
77 they underwent hyperdrive implantation targeted at the NAc. Rats were allowed to recover for a minimum  
78 of five days before being retrained on the task, and recording began once performance returned to pre-  
79 surgery levels. Upon completion of recording, animals were gliosed, euthanized and recording sites were

80 histologically confirmed.

81 **Behavioral task and training:**

82 The behavioral apparatus was an elevated, square-shaped track (100 x 100 cm, track width 10 cm) containing  
83 four possible reward locations at the end of track “arms” (Figure 2). Rats initiated a *trial* by triggering a  
84 photobeam located 24 cm from the start of each arm. Upon trial initiation, one of two possible light cues  
85 (L1, L2), or one of two possible sound cues (S1, S2), was presented that signaled the presence (*reward-*  
86 *available trial*, L1+, S1+) or absence (*reward-unavailable trial*, L2-, S2-) of a 12% sucrose water reward  
87 (0.1 mL) at the upcoming reward site. A trial was classified as an *approach trial* if the rat turned left at the  
88 decision point and made a nosepoke at the reward receptacle (40 cm from the decision point), while trials  
89 were classified as a *skip trial* if the rat instead turned right at the decision point and triggered the photobeam  
90 to initiate the next trial. A trial is labeled *correct* if the rat approached (i.e. nosepoked) on reward-available  
91 trials, and skipped (i.e. did not nosepoke) on reward-unavailable trials. On reward-available trials there  
92 was a 1 second delay between a nosepoke and subsequent reward delivery. *Trial length* was determined by  
93 measuring the length of time from cue onset until nosepoke (for approach trials), or from cue onset until  
94 the start of the following trial (for skip trials). Trials could only be initiated through clockwise progression  
95 through the series of arms, and each entry into the subsequent arm on the track counted as a trial.

96 Each session consisted of both a *light block* and a *sound block* with 100 trials each. Within a block, one cue  
97 signaled reward was available on that trial (L1+ or S1+), while the other signaled reward was not available  
98 (L2- or S2-). Light block cues were a flashing white light, and a constant yellow light. Sound block cues  
99 were a 2 kHz sine wave and a 8 kHz sine wave whose amplitude was modulated from 0 to maximum by  
100 a 2 Hz sine wave. Outcome-cue associations were counterbalanced across rats, e.g. for some rats L1+ was  
101 the flashing white light, and for others L1+ was the constant yellow light. The order of cue presentation  
102 was pseudorandomized so that the same cue could not be presented more than twice in a row. Block order  
103 within each day was also pseudorandomized, such that the rat could not begin a session with the same block

104 for more than two days in a row. Each session consisted of a 5 minute pre-session period on a pedestal (a  
105 terracotta planter filled with towels), followed by the first block, then the second block, then a 5 minute post-  
106 session period on the pedestal. For approximately the first week of training, rats were restricted to running  
107 in the clockwise direction by presenting a physical barrier to running counterclockwise. Cues signaling the  
108 availability and unavailability of reward, as described above, were present from the start of training. Rats  
109 were trained for 200 trials per day (100 trials per block) until they discriminated between the reward-available  
110 and reward-unavailable cues for both light and sound blocks for three consecutive days, according to a chi-  
111 square test rejecting the null hypothesis of equal approaches for reward-available and reward-unavailable  
112 trials, at which point they underwent electrode implant surgery.

113 [Figure 2 about here.]

114 **Surgery:**

115 Surgical procedures were as described previously (Malhotra et al., 2015). Briefly, animals were administered  
116 analgesics and antibiotics, anesthetized with isoflurane, induced with 5% in medical grade oxygen and main-  
117 tained at 2% throughout the surgery ( 0.8 L/min). Rats were then chronically implanted with a “hyperdrive”  
118 consisting of 16 independently drivable tetrodes, either all 16 targeted for the right NAc (AP +1.4 mm and  
119 ML +1.6 mm relative to bregma; Paxinos and Watson 1998), or 12 in the right NAc and 4 targeted at the  
120 mPFC (AP +3.0 mm and ML +0.6 mm, relative to bregma; only data from NAc tetrodes was analyzed).  
121 Following surgery, all animals were given at least five days to recover while receiving post-operative care,  
122 and tetrodes were lowered to the target (DV -6.0 mm) before being reintroduced to the behavioral task.

123 **Data acquisition and preprocessing:**

124 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with a  
125 RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex). Sig-

126 nals were referenced against a tetrode placed in the corpus callosum above the NAc. Candidate spikes for  
127 sorting into putative single units were obtained by band-pass filtering the data between 600-9000 Hz, thresh-  
128 olding and aligning the peaks (UltraMegaSort2k, Hill et al., 2011). Spike waveforms were then clustered  
129 with KlustaKwik using energy and the first derivative of energy as features, and manually sorted into units  
130 (MClust 3.5, A.D. Redish et al., <http://redishlab.neuroscience.umn.edu/MClust/MClust.html>). Isolated units  
131 containing a minimum of 200 spikes within a session were included for subsequent analysis. Units were clas-  
132 sified as fast spiking interneurons (FSIs) by an absence of interspike intervals (ISIs) > 2 s, while medium  
133 spiny neurons (MSNs) had a combination of ISIs > 2 s and phasic activity with shorter ISIs (Atallah et al.,  
134 2014; Barnes et al., 2005).

135 **Data analysis:**

136 *Behavior.* To determine if rats distinguished behaviorally between the reward-available and reward-unavailable  
137 cues (*cue outcome*), we generated linear mixed effects models to investigate the relationships between cue  
138 type and our behavioral variables, with *cue outcome* (reward available or not) and *cue identity* (light or  
139 sound) as fixed effects, and the addition of an intercept for rat identity as a random effect. For each cue,  
140 the average proportion of trials approached and trial length for a session were used as response variables.  
141 Contribution of cue outcome to behavior was determined by comparing the full model to a model with cue  
142 outcome removed for each behavioral variable.

143 *Neural data.* To investigate the contribution of different cue features (cue identity and cue outcome) on the  
144 firing rates of NAc single units, we first determined whether firing rates for a unit were modulated by the  
145 onset of a cue by collapsing across all cues and comparing the firing rates for the 1 s preceding cue-onset  
146 with the 1 s following cue-onset. Single units were considered to be *cue-modulated* if a Wilcoxon signed-  
147 rank test comparing pre- and post-cue firing was significant at  $p < .01$ . Cue-modulated units were then  
148 classified as either increasing or decreasing if the post-cue activity was higher or lower than the pre-cue  
149 activity, respectively.

150 To determine the relative contribution of different task parameters to firing rate variance (as in Figures 5-6),  
151 a forward selection stepwise general linear model (GLM) was fit to each cue-modulated unit. Cue identity  
152 (light block, sound block), cue location (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available, reward-  
153 unavailable), behavior (approach, skip), trial length, trial number, and trial history (reward availability on the  
154 previous 2 trials) were used as predictors, and the 1 s post-cue firing rate as the response variable. Units were  
155 classified as being modulated by a given task parameter if addition of the parameter significantly improved  
156 model fit using deviance as the criterion ( $p < .01$ ). A comparison of the R-squared value between the final  
157 model and the final model minus the predictor of interest was used to determine the amount of firing rate  
158 variance explained by the addition of that predictor for a given unit. To investigate more finely the temporal  
159 dynamics of the influence of task parameters to unit activity, we then fit a sliding window GLM with the  
160 same task parameters using 500 ms bins and 100 ms steps, starting 500 ms before cue-onset, up to 500 ms  
161 after cue-onset, and measured the proportion of units and average R-squared value for a given time bin where  
162 a particular predictor contributed significantly to the final model. To control for amount of units that would  
163 be affected by a predictor by chance, we shuffled the order of firing rates used for a time bin, and took the  
164 average of this value over 100 shuffles.

165 To better visualize responses to cues and enable subsequent population level analyses (as in Figures 5, 7,  
166 and 8), spike trains were convolved with a Gaussian kernel ( $\sigma = 100$  ms), and peri-event time histograms  
167 (PETHs) were generated by taking the average of the convolved spike trains across all trials for a given  
168 task condition. For analysis of population-level responses for cue features (Figure 7), convolved spike trains  
169 for all units where cue identity, cue location, or cue outcome explained a significant portion of firing rate  
170 variance were z-scored. Within a given cue feature, normalized spike trains were then separated according  
171 to the preferred and non-preferred cue condition (e.g. light vs. sound block), and averaged across units to  
172 generate population-level averages. To account for separation that would result from any random selection  
173 of units, unit identity was shuffled and the shuffled average for preferred and non-preferred cue conditions  
174 was generated for 1000 shuffles.

175 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units  
176 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure  
177 8). To compare representations of task space across cue conditions for a cue feature, the ordering of units  
178 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other  
179 condition (e.g. sound block). For control comparisons within cue conditions, half of the trials for a condition  
180 were compared against the other half. To look at the correlation of firing rates of all units within and across  
181 various cue conditions, trials for each cue condition for a unit were shuffled and divided into two averages,  
182 and averages within and across cue conditions were correlated. A linear mixed effects model was run for  
183 each cue condition to determine if correlations of firing rates within cue conditions were more similar than  
184 correlations across cue conditions.

185 To identify the responsivity of units to different cue features at the time of a nosepoke into a reward re-  
186 ceptacle, and subsequent reward delivery, the same cue-responsive units from the cue-onset analyses were  
187 analyzed at the time of nosepoke and outcome receipt using identical analysis techniques (Figures 9, 10, 11,  
188 and 12).

189 Given that some of our analyses compare firing rates across time, particularly comparisons across blocks,  
190 we sought to exclude units with unstable firing rates that would generate spurious results reflecting a drift  
191 in firing rate over time unrelated to our task. To do this we ran a Mann-Whitney U test comparing the  
192 cue-evoked firing rates for the first and second half of trials within a block, and excluded 99 of 443 units  
193 from analysis that showed a significant change for either block, leaving 344 units for further analyses.  
194 All analyses were completed in MATLAB R2015a, the code is available on our public GitHub repository  
195 (<http://github.com/vandermeelab/papers>), and the data can be accessed through DataLad.

196 **Histology:**

197 Upon completion of the experiment, recording channels were gliosed by passing 150  $\mu A$  current for 10

198 seconds and waiting 5 days before euthanasia, except for rat R057 whose implant detached prematurely.  
199 Rats were anesthetized with 5% isoflurane, then asphyxiated with carbon dioxide. Transcardial perfusions  
200 were performed, and brains were fixed and removed. Brains were sliced in 50  $\mu$ m coronal sections and  
201 stained with thionin. Slices were visualized under light microscopy, tetrode placement was determined, and  
202 electrodes with recording locations in the NAc were analyzed (Figure 3).

203 [Figure 3 about here.]

## 204 **Results**

### 205 **Behavior**

206 Rats were trained to discriminate between cues signaling the availability and absence of reward on a square  
207 track with four identical arms for two distinct set of cues (Figure 2). During each session, rats were pre-  
208 sented sequentially with two behavioral blocks containing cues from different sensory modalities, a light and  
209 a sound block, with each block containing a cue that signalled the availability of reward (reward-available),  
210 and a cue that signalled the absence of reward (reward-unavailable). To maximize reward receipt, rats should  
211 approach reward sites on reward-available trials, and skip reward sites on reward-unavailable trials (see Fig-  
212 ure 4A for an example learning curve). All four rats learned to discriminate between the reward-available  
213 and reward-unavailable cues for both the light and sound blocks as determined by reaching significance ( $p <$   
214 .05) on a daily chi-square test comparing approach behavior for reward-available and reward-unavailable  
215 cues for each block, for at least three consecutive days (range for time to criterion: 22 - 57 days). Mainte-  
216 nance of behavioral performance during recording sessions was assessed using linear mixed effects models  
217 for both proportion of trials where the rat approached the receptacle, and trial length. Analyses revealed  
218 that the likelihood of a rat to make an approach was influenced by whether a reward-available or reward-

219 unavailable cue was presented, but was not significantly modulated by whether the rat was presented with a  
220 light or sound cue (Percentage approached: light reward-available = 97%; light reward-unavailable = 34%;  
221 sound reward-available = 91%; sound reward-unavailable 35%; cue identity p = .115; cue outcome p< .001;  
222 Figure 4B). A similar trend was seen with the length of time taken to complete a trial (Trial length: light  
223 reward-available = 1.85 s; light reward-unavailable = 1.74 s; sound reward-available = 1.91 s; sound reward-  
224 unavailable 1.78 s; cue identity p = .115; cue outcome p < .001; Figure 4C). Thus, during recording, rats  
225 successfully discriminated the cues according to whether or not they signaled the availability of reward at  
226 the reward receptacle.

227 [Figure 4 about here.]

## 228 NAc units encode behaviorally relevant and irrelevant cue features

### 229 Single unit responses discriminate cue features:

230 We sought to address which parameters of our task were encoded by NAc activity, specifically whether  
231 the NAc encodes aspects of motivationally relevant cues not directly tied to reward, such as the identity  
232 and location of the cue, and whether this coding is independent or integrated with coding of cue outcome.  
233 To do this we recorded a total of 443 units with > 200 spikes in the NAc from 4 rats over 57 sessions  
234 while they performed a cue discrimination task (Table 1). Units that exhibited a drift in firing rate over  
235 the course of either block were excluded from further analysis, leaving 344 units for further analysis. The  
236 activity of 133 (39%) of these 344 units were modulated by the cue, with more showing a decrease in firing  
237 (n = 103) than an increase (n = 30) around the time of cue-onset (Table 1). Within this group, 24 were  
238 classified as FSIs, while 109 were classified as SPNs. Upon visual inspection, we observed several patterns  
239 of firing activity, including units that discriminated firing upon cue-onset across various cue conditions,  
240 showed sustained differences in firing across cue conditions, had transient responses to the cue, showed a

241 ramping of activity starting at cue-onset, and showed elevated activity immediately preceding cue-onset, for  
242 example (Figure 5). To characterize more formally whether these cue-evoked responses were modulated by  
243 various aspects of the task, we fit a GLM to each cue-modulated unit. Fitting GLMs revealed that a variety  
244 of task parameters accounted for a significant portion of firing rate variance in NAc cue-modulated units  
245 (Figure 6, Table 1). Notably, there were units that discriminated between whether the rat was performing  
246 in the light or sound block (28% of cue-modulated units, accounting for 6% of variance on average), which  
247 arm the rat was currently on (38% of cue-modulated units, accounting for 6% of variance on average), and  
248 whether the rat was engaged in the common portion of a reward-available or reward-unavailable trial (26%  
249 of cue-modulated units, accounting for 4% of variance on average), suggesting that the NAc encodes features  
250 of reward-predictive cues separate from expected outcome (Figure 5A-F). Furthermore, overlap of coding of  
251 cue features within units was not different than expected by chance according to chi-square tests, suggesting  
252 for integrated coding across various aspects of a cue (Figure 5G,H). Additionally, a sliding window GLM  
253 centered on cue-onset revealed that cue identity and cue location contributed to the activity of a significant  
254 proportion of cue-modulated units throughout this epoch, whereas an increase in units encoding cue outcome  
255 became apparent after cue-onset (Figure 6D). Together, these findings show that various cue features are  
256 represented in the NAc, and that this coding is both integrated and separate from expected outcome (Figure  
257 1; H2,H3).

258 [Table 1 about here.]

259 [Figure 5 about here.]

260 [Figure 6 about here.]

261 **Population level averages reveal characteristic response profiles:**

262 We observed a variety of single unit response profiles around the time of cue onset (Figure 5). To investigate

whether these firing rate patterns were related to what cue features were encoded, we plotted the population level averages for units that were modulated by each feature. To do this, we normalized firing activity for each unit that was modulated by a given cue feature, such as light block, then generated the cue-onset aligned population average firing rate for each of the cue features (Figure 7). Overall, this analysis revealed that cells that showed an increase upon cue presentation had stronger responses for the preferred cue condition (Figure 7A,C,E). Interestingly, units that were classified as decreasing in response to the cue showed a biphasic response at the population level, with a small peak at a time in alignment with entry into the arm, followed by a sustained dip after cue-onset (Figure 7B,D,F). Units that were modulated by cue identity showed a stronger increase in response to the preferred task block, as well as a higher tonic firing rate to the preferred task block, most notably in units that decreased in firing rate to the cue (Figure 7A,B). Units that were modulated by cue location showed a graded response to locations of decreasing preference, with peak firing occurring around cue-onset (Figure 7C,D). Units that were modulated by cue outcome showed a ramping of activity after cue-onset for their preferred cue type. Additionally, units that exhibited a decrease in firing in response to the cue and whose activity was modulated by cue outcome, showed a sustained discriminatory response to reward-available and reward-unavailable cues that extended beyond cue-onset (Figure 7F). Together, these visualizations of the averaged population responses revealed nuanced differences in the way NAc units are modulated by cue conditions across cue features.

[Figure 7 about here.]

#### **NAc units dynamically segment the task:**

Given the varied time courses and response profiles of NAc units to various aspects of the cue, the NAc may be computing a temporally evolving state value signal (Pennartz., 2011). If this is the case, then the recruitment of NAc units should vary alongside changes in the environment. To look at the distribution of responses throughout our task space and see if this distribution is modulated by cue features, we z-scored the firing rate of each unit and plotted the normalized firing rates of all units aligned to cue-onset and sorted

them according to the time of peak firing rate (Figure 8). We did this separately for both the light and sound blocks, and found a nearly uniform distribution of firing fields in task space that was not limited to alignment to the cue (Figure 8A). Furthermore, to determine if this population level activity was similar across blocks, we also organized firing during the sound blocks according to the ordering derived from the light Blocks. This revealed that while there was some preservation of order, the overall firing was qualitatively different across the two blocks, implying that population activity distinguishes between light and sound blocks. To control for the possibility that any comparison of trials would produce this effect, we did a within block comparison, comparing half of the trials in the light block against the other half. This comparison looked similar to our test comparison of sound block trials ordered by light block trials. Additionally, given that the majority of our units showed an inhibitory response to the cue, we also plotted the firing rates according to the lowest time in firing, and again found some maintenance of order, but largely different ordering across the two blocks, and the within block comparison (Figure 8B). To further test this, we divided each block into two halves and looked at the correlation of the average smoothed firing rates across various combinations of these halves across our cue-aligned centered epoch. A linear mixed effects model revealed that within block correlations (e.g. one half of light trials vs other half of light trials) were higher and more similar than across block correlations (e.g. half of light trials vs half of sound trials) suggesting that activity in the NAc discriminates across various cue conditions (within block correlations = .383 (light), .379 (sound); across block correlations = .343, .338, .337, .348; within block vs./ within block comparison =  $p = .934$ ; within block vs./ across block comparisons =  $p < .001$ ). This process was repeated for cue location (Figure 8C-D; within block correlations = .369 (arm 1), .350 (arm 2); across block correlations = .290, .286, .285, .291; within block vs./ within block comparison =  $p = .071$ ; within block vs./ across block comparisons =  $p < .001$ ) and cue outcome (Figure 8E-F; within block correlations = .429 (reward-available), .261 (reward-unavailable); across block correlations = .258, .253, .255, .249; within block vs./ within block comparison =  $p < .001$ ; within block vs./ across block comparisons =  $p < .001$ ), showing that NAc segmentation of the task is qualitatively different even during those parts of the task not immediately associated with a specific cue, action, or outcome, although the within condition comparison of reward-unavailable trials was less correlated than reward-available trials, and more similar to the across condition comparisons, potentially due to the greater behavioral variability for the reward-unavailable trials.

316 **Encoding of cue features persists until outcome:**

317 In order to be useful for credit assignment in reinforcement learning, a trace of the cue must be maintained  
318 until the outcome, so that information about the outcome can be associated with the outcome-predictive cue.  
319 To test whether representations of cue features persisted post-approach until the outcome was revealed, we  
320 fit a GLM to the post-approach firing rates of cue-modulated units aligned to the time of nosepoke into the  
321 reward receptacle. This analysis showed that a variety of units still discriminated firing according to various  
322 cue features, but not other task parameters, showing that NAc activity discriminates various cue conditions  
323 well into a trial (Table 1, Figures 9,10). Additionally, these units were a mix between most of the units  
324 that encoded cue features at cue-onset (observed overlap greater than expected by chance according to chi-  
325 square tests), and those that did not previously have a cue feature as a predictor (29, 48, and 30 out of 133  
326 cue-modulated units encoded both time points for cue identity, cue location, and cue outcome, respectively).  
327 Population level averages for units that increased to cue-onset showed a ramping up of activity that peaked  
328 upon nosepoke, whereas units that decreased to cue-onset showed a gradual reduction of firing activity that  
329 reached a minimum upon nosepoke (Figure 11). Additionally, a peak is seen for preferred cue outcome  
330 in decreasing units at 1 second post cue-onset when reward was received, demonstrating an integration of  
331 expected and received reward (Figure 11F). Furthermore, aligning normalized peak firing rates to nosepoke  
332 onset, revealed a clustering of responses around outcome receipt for all cue conditions where the rat would  
333 have received reward (Figure 12). To determine whether coding of cue features persisted after the outcome  
334 was revealed, a GLM was fit to the firing rates of cue-modulated units at the time of outcome receipt, during  
335 which the cue was still present. Fitting a GLM revealed 10 units (8%) where cue outcome accounted for an  
336 average of 32% of firing rate variance (Table 1, data not shown). An absence of cue identity or cue location  
337 coding at this level of analysis was observed, but looking at the data more closely with a sliding window GLM  
338 revealed that cue identity, cue location, and cue outcome were encoded throughout time epochs surround cue-  
339 onset, nosepoke hold, and outcome receipt, suggesting that the NAc maintains a representation of these cue

340 features once the rat receives behavioral feedback for its decision (Figure 10E-J).

341 [Figure 9 about here.]

342 [Figure 10 about here.]

343 [Figure 11 about here.]

344 [Figure 12 about here.]

## 345 **Discussion**

346 The main result of the present study is that NAc units encode not only the expected outcome of outcome-  
347 predictive cues, but also the identity of such cues. Importantly, this identity coding was maintained on  
348 approach trials during a delay period where the rat held a nosepoke until the outcome was received (H2 in  
349 Figure 1B). Units coding for cue identity showed partial overlap with those coding for expected outcome  
350 (H3 in Figure 1A). Units that coded different cue features (identity, outcome, location) exhibited different  
351 temporal profiles as a whole, although across all recorded units a tiling of task structure was observed such  
352 that all points within our analyzed task space was accounted for by the ordered peak firing rates of all units.  
353 Furthermore, this tiling differed between various conditions with a cue feature, such as light versus sound  
354 blocks. We discuss these observations and their implications below.

### 355 **Cue identity:**

356 Our finding that NAc units can discriminate between different outcome-predictive stimuli with similar moti-  
357 vational significance (i.e. encodes cue identity) expands upon an extensive rodent literature examining NAc  
358 correlates of conditioned stimuli (Ambroggi et al., 2008; Atallah et al., 2014; Bissonette et al., 2013; Cooch  
359 et al., 2015; Day et al., 2006; Dejean et al., 2017; Goldstein et al., 2012; Ishikawa et al., 2008; Lansink et al.,  
360 2012; McGinty et al., 2013; Nicola, 2004; Roesch et al., 2009; Roitman et al., 2005; Saddoris et al., 2011;  
361 Setlow et al., 2003; Sugam et al., 2014; West and Carelli, 2016; Yun et al., 2004). Perhaps the most compara-  
362 ble work in rodents comes from a study that found distinct coding for an odor when it predicted separate but  
363 equally valued rewards (Cooch et al., 2015). The present work is complementary to such *outcome identity*  
364 coding as it shows that NAc units *cue identity*, both separately and integrated with the reward it predicts (H2  
365 and H3 in Figure 1A). Similarly, Setlow et al. 2003 paired distinct cues with appetitive or aversive outcomes,  
366 and found separate populations of units that encoded each cue. Once again, our study was different in asking  
367 how distinct cues encoding the same anticipated outcome are encoded. Such cue identity encoding suggests  
368 that even when the biological relevance of these stimuli is similar, NAc dissociates their representations at  
369 the level of the single-units. A possible interpretation of this coding of cue features alongside expected out-  
370 come is that these representations are used to associate reward with relevant features of the environment,  
371 so-called credit assignment in the reinforcement learning literature (Sutton and Barto, 1998). A burgeoning  
372 body of human and non-human primate work has started to elucidated neural correlates of credit assignment  
373 in the PFC, particularly in the lateral orbitofrontal cortex (Akaishi et al., 2016; Asaad et al., 2017; Chau et al.,  
374 2015; Noonan et al., 2017). Given the importance of cortical inputs in NAc associative representations, it is  
375 possible that information related to credit assignment is relayed from the cortex to NAc (Cooch et al., 2015;  
376 Ishikawa et al., 2008).

377 A different possible function for cue identity coding is to support contextual modulation of the motivational  
378 relevance of specific cues. A context can be understood as a particular mapping between specific cues and  
379 their outcomes: for instance, in context 1 cue A but not cue B is rewarded, whereas in context 2 cue B but not  
380 cue A is rewarded. Successfully implementing such contextual mappings requires representation of the cue  
381 identities. Indeed, (Sleeker et al., 2016) recorded NAc responses during the Wisconsin Card Sorting Task

382 (WCST), a common set-shifting task used in both the laboratory and clinic, and found units that preferred  
383 firing to stimuli when a certain rule, or rule category was currently active. Further support for a modulation  
384 of NAc responses by strategy comes from an fMRI study that examined BOLD levels during a set-shifting  
385 task (FitzGerald et al., 2014). In this task, participants learned two sets of stimulus-outcome contingencies,  
386 a visual set and auditory set. During testing they were presented with both simultaneously, and the stimulus  
387 dimension that was relevant was periodically shifted between the two. Here, they found that bilateral NAc  
388 activity reflected value representations for the currently relevant stimulus dimension, and not the irrelevant  
389 stimulus. The current finding of separate, but overlapping, populations of units encoding cue identity and  
390 expected outcome, suggests that the fMRI finding is generated by the combined activity of several different  
391 functional cell types.

392 Our analyses were designed to eliminate several potential alternative interpretations to cue identity coding.  
393 Because the different cues were separated into different blocks, units that discriminated between cue identi-  
394 ties could instead be encoding time or other slowly-changing quantities. We excluded this possible confound  
395 by excluding units that showed a drift in firing between the first and second half within a block. However, the  
396 possibility remains that instead of or in addition to stimulus identity, these units encode a preferred context,  
397 or even a macroscale representation of progress through the session. Indeed, encoding of the current strategy  
398 could be an explanation for the sustained difference in population averaged firing across stimulus blocks  
399 (Figure 7), as well as a potential explanation for the differentially tiling of task structure across blocks in the  
400 current study (Figure 8).

401 A different potential confound is that between outcome and action value coding. We discriminated between  
402 these possibilities by analyzing error trials, where the rat approached reward (left turn) after presentation of  
403 the reward-unavailable cue. Units that were modulated by the expected outcome of the cue maintained their  
404 specific firing patterns even during error trials, as expected from outcome value coding but not action value  
405 coding. Additionally, NAc signals have been shown to be modulated by response vigor (McGinty et al.,  
406 2013); to detangle this from our results we included trial length (i.e. latency to arrival at the reward site) as a

407 predictor in our GLMs, and found units with cue feature correlates independent of trial length.

408 An overall limitation of the current study is that rats were never presented with both sets of cues simultane-  
409 ously, and were not required to switch strategies between multiple sets of cues. Thus, it is unknown to what  
410 extent the cue identity encoding we observed is behaviorally relevant, although extrapolating data from other  
411 work (Sleeker et al., 2016) suggests that cue identity coding would be modulated by relevance. NAc core  
412 lesions have been shown to impair shifting between different behavioral strategies (Floresco et al., 2006),  
413 and it is possible that selectively silencing the units that prefer responding for a given modality or rule would  
414 impair performance when the animal is required to use that information, or artificial enhancement of those  
415 units would cause them to use the rule when it is the inappropriate strategy.

416 **Encoding of position:**

417 Our finding that cue-evoked activity was modulated by cue location is in alignment with several previous  
418 reports (Lavoie and Mizumori, 1994; Mulder et al., 2005; Strait et al., 2016; Wiener et al., 2003). The NAc  
419 receives inputs from the hippocampus, and the communication of place-reward information across the two  
420 structures suggests that the NAc tracks locations associated with reward (Lansink et al., 2008, 2009, 2016;  
421 Pennartz, 2004; Sjulson et al., 2017; Tabuchi et al., 2000; van der Meer and Redish, 2011). NAc units can also  
422 signal progress through a sequence of cues and/or actions (Shidara et al., 1998; Atallah et al., 2014; Berke  
423 et al., 2009; Khamassi et al., 2008; Lansink et al., 2012; Mulder et al., 2004). Given that the current task was  
424 pseudo-random, it is possible that the rats learned the structure of sequential cue presentation, and the neural  
425 activity could reflect this. However, this is unlikely as including a previous trial variable in the analysis did  
426 not explain a significant amount of firing rate variance in response to the cue for the vast majority of units.  
427 In any case, NAc units on the present task continued to distinguish between different locations, even though  
428 location, and progress through a sequence, were explicitly irrelevant in predicting reward. We speculate that  
429 this persistent coding of location in NAc may represent a bias in credit assignment, and associated tendency  
430 for rodents to associate motivationally relevant events with the locations where they occur.

431 **Implications:**

432 Maladaptive decision making, as occurs in schizophrenia, addiction, Parkinsons, among others, can result  
433 from dysfunctional RPE and value signals (Frank et al., 2004; Gradin et al., 2011; Maia and Frank, 2011).  
434 This view has been successful in explaining both positive and negative symptoms in schizophrenia, and  
435 deficits in learning from feedback in Parkinsons (Frank et al., 2004; Gradin et al., 2011). However, the  
436 effects of RPE and value updating are contingent upon encoding of preceding action and cue features, the  
437 eligibility trace (Lee et al., 2012; Sutton and Barto, 1998). Value updates can only be performed on these  
438 aspects of preceding experience that are encoded when the update occurs. Therefore, maladaptive learning  
439 and decision making can result from not only aberrant RPEs but also from altered cue feature encoding. For  
440 instance, on this task the environmental stimulus that signaled the availability of reward was conveyed by  
441 two distinct cues that were presented in four locations. While in our current study, the location and identity  
442 of the cue did not require any adjustments in the animals behavior, we found coding of these features along-  
443 side the expected outcome of the cue that could be the outcome of credit assignment computations computed  
444 upstream. Identifying neural coding related to an aspect of credit assignment is important as inappropriate  
445 credit assignment could be a contributor to conditioned fear overgeneralization seen in disorders with patho-  
446 logical anxiety such as generalized anxiety disorder, post traumatic stress disorder, and obsessive-compulsive  
447 disorder (Shmuel Lissek, 2013; Kaczkurkin et al., 2017; Lissek et al., 2014), and delusions observed in disor-  
448 ders such as schizophrenia, Alzheimers and Parkinsons (Corlett et al., 2010; Kapur, 2003). Thus, our results  
449 provide a neural window into the process of credit assignment, such that the extent and specific manner in  
450 which this process fails in e.g. syndromes such as schizophrenia, obsessive-compulsive disorder, etc. can be  
451 experimentally accessed.

452 **OUTTAKES**

453 **Tiling of task structure:**

454 We found that the population of recorded units had a relatively uniform distribution of firing fields within  
455 our task space, similar to what has been reported previously (Berke et al., 2009; Lansink et al., 2012; Shidara  
456 et al., 1998). Uniquely, we found that this representation was statistically different across conditions of a cue  
457 features, such as according to whether the rat was currently engaged in the light or sound block, suggesting  
458 that this could be a possible neural correlate for encoding the currently relevant strategy in the NAc. It has  
459 been previously shown that during progress through a predictable trial series, units represented state value of  
460 cue, and that single-unit responses allowed the monkey to know how it was progressing throughout the task  
461 (Shidara et al., 1998). Likewise, the tiling we saw could be a consequence of upstream cortical or limbic  
462 inputs informing the striatum of the current task rules. Another possibility is that the NAc not only pays  
463 attention to progress throughout a task within a trial, but also higher-order task information, like blocks.  
464 Furthermore, dopamine levels in the NAc fluctuate through a trial, and it is possible that the observed tiling  
465 could be a NAc-representation of state value related to this temporally evolving dopamine signal. Future  
466 experiments should monitor this mapping of task structure during the application of dopamine antagonists.  
467 Finally, the presence of functional correlates not evident when looking at single-unit responses time-locked  
468 to salient task events emphasizes the need to employ ensemble level analyses across all aspects of a task.

469 **Potential functional consequences of persistent coding:**

470 In the current study we found that the coding of cue features persisted after the choice point, during a delay  
471 period while the rat waited at the receptacle for the outcome (H2 in Figure 1B). Having an enriched repre-  
472 sentation that includes details about the environmental context the animal finds itself in alongside expected  
473 outcome, and is maintained online until the outcome is revealed could be useful for assigning credit of a re-  
474 ward to the appropriate elements of an environment, to develop an accurate value function and optimize long  
475 term acquisition of reward (Lee et al., 2012). A recent non-human primate experiment found evidence for  
476 simultaneous coding of outcome and outcome-predictive stimuli at the time of feedback in the dorsolateral  
477 PFC, and that this coding of stimulus information was strongest before behavioral performance stabilized  
478 (Asaad et al., 2017). Interestingly, we found coding of cue identity and location after presentation of the

479 cue, and during a delay period between nosepoke and outcome, but not after the outcome was received. The  
480 absence of coding during feedback, and that these recordings were done after behavioral performance had  
481 stabilized suggests the possibility that these NAc representations represent an integrated value estimate that  
482 is the outcome of earlier upstream credit assignment in cortical inputs. From a motivated behavior stand-  
483 point, the presence of these enriched representations can help inform action selection involving Pavlovian  
484 behaviors such as conditioned approach (Ciano et al., 2001; Parkinson et al., 2000; Saunders and Robinson,  
485 2012), and after a decision has been made the maintained representation can help to hold the response in the  
486 face of competing alternatives while waiting for an outcome (Di Ciano et al., 2008; Floresco et al., 2008;  
487 Floresco, 2015; Peters et al., 2008).

488 If these representations function as an eligibility trace, then there should be a relationship between the con-  
489 tent and robustness of these representations and the degree to which organisms form adaptive associations  
490 between reinforcers and their environment, which could contribute to the presence of prediction errors in  
491 the NAc in those that learned an experimental task in humans versus those that did not (?). In the present  
492 experiment, understanding what information about the environment the NAc is keeping track of when re-  
493 ward arrives could be useful for potentially determining what representations are being reinforced. Would be  
494 interesting to see if selectively reward an animal when certain representations are active (e.g. arm 1), would  
495 help shape their assignment of reward to that arm. Could be useful for altering a maladaptive preference  
496 back to a more adaptive one, as the loss of certain aspects of coding via BLA and PFC inputs leads to the  
497 loss of outcome-predictive activity to discrete cues in the NAc. Furthermore, non-discriminatory coding in  
498 the NAc could hypothetically be correlated to over-generalization of situations, in which the animal may  
499 not recognize it is in a different situation and perform actions that are inefficient or maladaptive for reward  
500 procurement.

501 A fundamental problem faced by all reinforcement learning agents is which features to credit and blame for  
502 particular outcomes (Sutton and Barto, 1998); adaptive behavior requires associating motivationally relevant  
503 outcomes with the cues that predict them through learning from feedback. Much work has focused on value

504 signals such as reward prediction errors (RPEs), state values and action values (Lee et al., 2012; Maia, 2009).  
505 However, successful learning requires not only a RPE but also a trace of the preceding actions and/or cues.

## 506 References

- 507 Akaishi, R., Kolling, N., Brown, J. W., and Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue  
508 Environment. *Journal of Neuroscience*, 36(4):1096–1112.
- 509 Ambroggi, F., Ishikawa, A., Fields, H. L., and Nicola, S. M. (2008). Basolateral Amygdala Neurons Facilitate Reward-Seeking  
510 Behavior by Exciting Nucleus Accumbens Neurons. *Neuron*, 59(4):648–661.
- 511 Asaad, W. F., Lauro, P. M., Perge, J. A., and Eskandar, E. N. (2017). Prefrontal Neurons Encode a Solution to the Credit-Assignment  
512 Problem. *The Journal of Neuroscience*, 37(29):6995–7007.
- 513 Atallah, H. E., McCool, A. D., Howe, M. W., and Graybiel, A. M. (2014). Neurons in the ventral striatum exhibit cell-type-specific  
514 representations of outcome during learning. *Neuron*, 82(5):1145–1156.
- 515 Averbeck, B. B. and Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*,  
516 20(4):505–512.
- 517 Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., and Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and  
518 recoding of procedural memories. *Nature*, 437(7062):1158–1161.
- 519 Berke, J. D., Breck, J. T., and Eichenbaum, H. (2009). Striatal Versus Hippocampal Representations During Win-Stay Maze  
520 Performance. *Journal of Neurophysiology*, 101(3):1575–1587.
- 521 Berridge, K. C. (2012). From prediction error to incentive salience: Mesolimbic computation of reward motivation. *European  
522 Journal of Neuroscience*, 35(7):1124–1143.
- 523 Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., Kashtelyan, V., and Roesch, M. R.  
524 (2013). Separate Populations of Neurons in Ventral Striatum Encode Value and Motivation. *PLoS ONE*, 8(5):e64673.
- 525 Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological  
526 bulletin*, 114(1):80–99.
- 527 Carelli, R. M. (2010). Drug Addiction: Behavioral Neurophysiology. In *Encyclopedia of Neuroscience*, pages 677–682. Elsevier.
- 528 Chang, S. E. and Holland, P. C. (2013). Effects of nucleus accumbens core and shell lesions on autoshaped lever-pressing. *Be-  
529 havioural Brain Research*, 256:36–42.
- 530 Chang, S. E., Wheeler, D. S., and Holland, P. C. (2012). Roles of nucleus accumbens and basolateral amygdala in autoshaped lever  
531 pressing. *Neurobiology of Learning and Memory*, 97(4):441–451.

- 532 Chau, B. K. H., Sallet, J., Papageorgiou, G. K., Noonan, M. A. P., Bell, A. H., Walton, M. E., and Rushworth, M. F. S. (2015).  
533 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5):1106–  
534 1118.
- 535 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., and Wightman, R. M. (2007). Coordinated Accumbal  
536 Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2):237–244.
- 537 Ciano, P. D., Cardinal, R. N., Cowell, R. A., Little, S. J., Everitt, B. J., Di Ciano, P., Cardinal, R. N., Cowell, R. A., Little, S. J.,  
538 and Everitt, B. J. (2001). Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens  
539 core in the acquisition and performance of pavlovian approach behavior. *The Journal of neuroscience*, 21(23):9471–9477.
- 540 Cohen, J. D. and Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in  
541 schizophrenia. *Psychological review*, 99(1):45.
- 542 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., and Schoenbaum, G. (2015).  
543 Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature Communi-*  
544 *cations*, 6:7195.
- 545 Corbit, L. H. and Balleine, B. W. (2011). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are  
546 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33):11786–11794.
- 547 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., and Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in*  
548 *Neurobiology*, 92:345–369.
- 549 Day, J. J., Roitman, M. F., Wightman, R. M., and Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine  
550 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8):1020–1028.
- 551 Day, J. J., Wheeler, R. A., Roitman, M. F., and Carelli, R. M. (2006). Nucleus accumbens neurons encode Pavlovian approach  
552 behaviors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5):1341–1351.
- 553 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., Boraud, T., and Le Moine, C. (2017). Memories of Opiate  
554 Withdrawal Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychopharmacology*,  
555 42(5):1157–1168.
- 556 Di Ciano, P., Robbins, T. W., and Everitt, B. J. (2008). Differential effects of nucleus accumbens core, shell, or dorsal striatal  
557 inactivations on the persistence, reacquisition, or reinstatement of responding for a drug-paired conditioned reinforcer. *Neu-*  
558 *ropsychopharmacology*, 33(6):1413–1425.
- 559 du Hoffmann, J. and Nicola, S. M. (2014). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the  
560 Nucleus Accumbens. *Journal of Neuroscience*, 34(43):14349–14364.
- 561 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental*  
562 *Psychology*, 32(2):150–155.
- 563 Fitzgerald, T. H. B., Schwartenbeck, P., and Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent  
564 and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4):1271–1279.

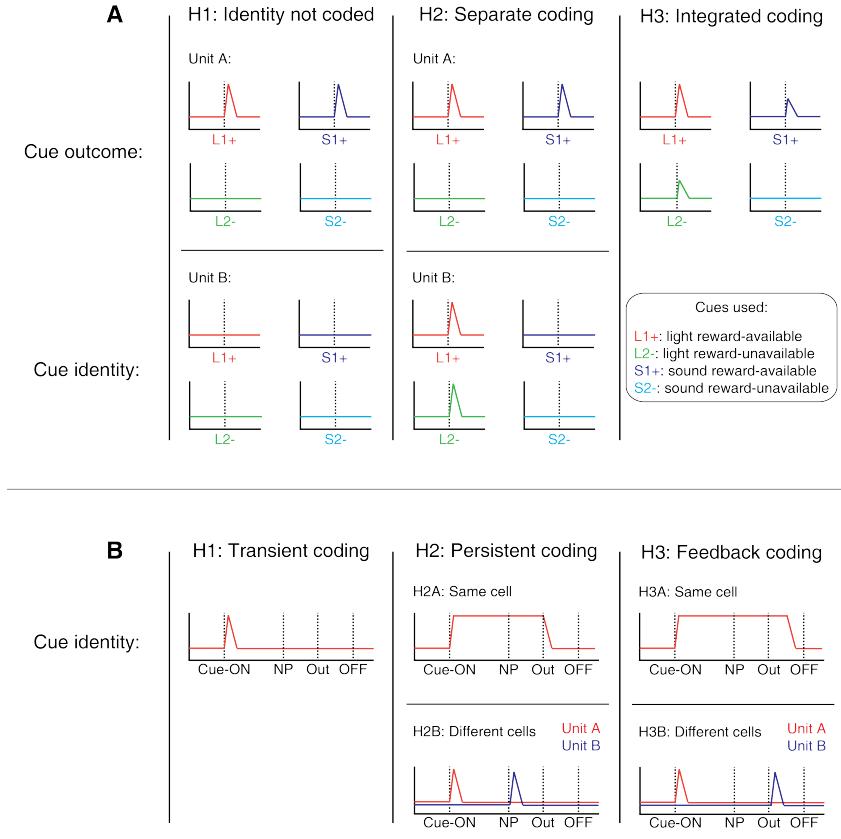
- 565 Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. M., and Akil,  
566 H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, 469(7328):53–59.
- 567 Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of  
568 Psychology*, 66(1):25–52.
- 569 Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., and Magyar, O. (2006). Dissociable roles for the nucleus accumbens core and  
570 shell in regulating set shifting. *J Neurosci*, 26(9):2449–2457.
- 571 Floresco, S. B., McLaughlin, R. J., and Haluk, D. M. (2008). Opposing roles for the nucleus accumbens core and shell in cue-induced  
572 reinstatement of food-seeking behavior. *Neuroscience*, 154(3):877–884.
- 573 Frank, M. J., Seeberger, L. C., and O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism.  
574 *Science*, 306(5703):1940–1943.
- 575 Giertler, C., Bohn, I., and Hauber, W. (2004). Transient inactivation of the rat nucleus accumbens does not impair guidance of  
576 instrumental behaviour by stimuli predicting reward magnitude. *Behavioural Pharmacology*, 15(1):55–63.
- 577 Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., Bryden, D. W., and Roesch, M. R.  
578 (2012). Ventral Striatum Encodes Past and Predicted Value Independent of Motor Contingencies. *Journal of Neuroscience*,  
579 32(6):2027–2036.
- 580 Goto, Y. and Grace, A. A. (2008). Limbic and cortical information processing in the nucleus accumbens. *Trends in Neurosciences*,  
581 31(11):552–558.
- 582 Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., and Steele, J. D. (2011). Expected value  
583 and prediction error abnormalities in depression and schizophrenia. *Brain*, 134(6):1751–1764.
- 584 Grant, D. A. and Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a  
585 Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4):404–411.
- 586 Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J.,  
587 and Berke, J. D. (2015). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1):117–126.
- 588 Hart, A. S., Rutledge, R. B., Glimcher, P. W., and Phillips, P. E. M. (2014). Phasic Dopamine Release in the Rat Nucleus Accumbens  
589 Symmetrically Encodes a Reward Prediction Error Term. *The Journal of Neuroscience*, 34(3):698–704.
- 590 Hearst, E. and Jenkins, H. M. (1974). *Sign-tracking: the stimulus-reinforcer relation and directed action*. Psychonomic Society.
- 591 Hill, D. N., Mehta, S. B., and Kleinfeld, D. (2011). Quality Metrics to Accompany Spike Sorting of Extracellular Signals. *Journal  
592 of Neuroscience*, 31(24):8699–8705.
- 593 Holland, P. C. (1992). Occasion setting in pavlovian conditioning. *Psychology of Learning and Motivation - Advances in Research  
594 and Theory*, 28(C):69–125.
- 595 Hollerman, J. R., Tremblay, L., and Schultz, W. (1998). Influence of Reward Expectation on Behavior-Related Neuronal Activity in  
596 Primate Striatum. *Journal of Neurophysiology*, 80(2):947–963.

- 597 Honey, R. C., Iordanova, M. D., and Good, M. (2014). Associative structures in animal learning: Dissociating elemental and  
598 configural processes. *Neurobiology of Learning and Memory*, 108:96–103.
- 599 Hyman, S. E., Malenka, R. C., and Nestler, E. J. (2006). NEURAL MECHANISMS OF ADDICTION: The Role of Reward-Related  
600 Learning and Memory. *Annual Review of Neuroscience*, 29(1):565–598.
- 601 Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-  
602 olfactory tubercle complex. *Brain Research Reviews*, 56(1):27–78.
- 603 Ishikawa, A., Ambroggi, F., Nicola, S. M., and Fields, H. L. (2008). Dorsomedial Prefrontal Cortex Contribution to Behavioral and  
604 Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, 28(19):5088–5098.
- 605 Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives.  
606 *Neural Networks*, 15(4-6):535–547.
- 607 Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., Sponheim, S. R., and Lissek,  
608 S. (2017). Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174(2):125–134.
- 609 Kalivas, P. W. and Volkow, N. D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal  
610 of Psychiatry*, 162(8):1403–1413.
- 611 Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in  
612 schizophrenia. *American Journal of Psychiatry*, 160(1):13–23.
- 613 Khamassi, M. and Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and  
614 model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, 6:79.
- 615 Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., and Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal  
616 neurons of behaving rats. *European Journal of Neuroscience*, 28(9):1849–1866.
- 617 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., and Pennartz, C. M. A. (2008).  
618 Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, 28(25):6372–  
619 6382.
- 620 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., and Pennartz, C. M. (2009). Hippocampus leads ventral  
621 striatum in replay of place-reward information. *PLoS Biology*, 7(8):e1000173.
- 622 Lansink, C. S., Jackson, J. C., Lankelma, J. V., Ito, R., Robbins, T. W., Everitt, B. J., and Pennartz, C. M. A. (2012). Reward  
623 Cues in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles. *Journal of  
624 Neuroscience*, 32(36):12444–12459.
- 625 Lansink, C. S., Meijer, G. T., Lankelma, J. V., Vinck, M. A., Jackson, J. C., and Pennartz, C. M. A. (2016). Reward Expectancy  
626 Strengthens CA1 Theta and Beta Band Synchronization and Hippocampal-Ventral Striatal Coupling. *Journal of Neuroscience*,  
627 36(41):10598–10610.
- 628 Lavoie, A. M. and Mizumori, S. J. (1994). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons  
629 of rats. *Brain Research*, 638(1-2):157–168.

- 630 Lee, D., Seo, H., and Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of  
631 Neuroscience*, 35(1):287–308.
- 632 Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., and Grillon, C. (2014). Generalized anxiety disorder is associated  
633 with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75(11):909–915.
- 634 Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and  
635 Behavioral Neuroscience*, 9(4):343–364.
- 636 Maia, T. V. and Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature  
637 Neuroscience*, 14(2):154–162.
- 638 Malhotra, S., Cross, R. W., Zhang, A., and Van Der Meer, M. A. A. (2015). Ventral striatal gamma oscillations are highly variable  
639 from trial to trial, and are dominated by behavioural state, and only weakly influenced by outcome value. *European Journal of  
640 Neuroscience*, 42(10):2818–2832.
- 641 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., and Schoenbaum, G. (2011). Ventral Striatum and Orbitofrontal Cortex  
642 Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, 31(7):2700–2705.
- 643 McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., and Nicola, S. M. (2013). Invigoration of reward seeking by cue and proximity  
644 encoding in the nucleus accumbens. *Neuron*, 78(5):910–922.
- 645 Mulder, A. B., Shibata, R., Trullier, O., and Wiener, S. I. (2005). Spatially selective reward site responses in tonically active neurons  
646 of the nucleus accumbens in behaving rats. *Experimental Brain Research*, 163(1):32–43.
- 647 Mulder, A. B., Tabuchi, E., and Wiener, S. I. (2004). Neurons in hippocampal afferent zones of rat striatum parse routes into  
648 multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7):1923–1932.
- 649 Nicola, S. M. (2004). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a Discrimi-  
650 native Stimulus Task. *Journal of Neurophysiology*, 91(4):1840–1865.
- 651 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of  
652 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49):16585–16600.
- 653 Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor.  
654 *Psychopharmacology*, 191(3):507–520.
- 655 Noonan, M. P., Chau, B. K., Rushworth, M. F., and Fellows, L. K. (2017). Contrasting Effects of Medial and Lateral Orbitofrontal  
656 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29):7023–7035.
- 657 Parkinson, J. A., Willoughby, P. J., Robbins, T. W., and Everitt, B. J. (2000). Disconnection of the anterior cingulate cortex  
658 and nucleus accumbens core impairs pavlovian approach behavior: Further evidence for limbic cortical-ventral striatopallidal  
659 systems. *Behavioral Neuroscience*, 114(1):42–63.
- 660 Paxinos, G. and Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates*. Academic Press, San Diego, 4 edition.
- 661 Pennartz, C. M. A. (2004). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation by

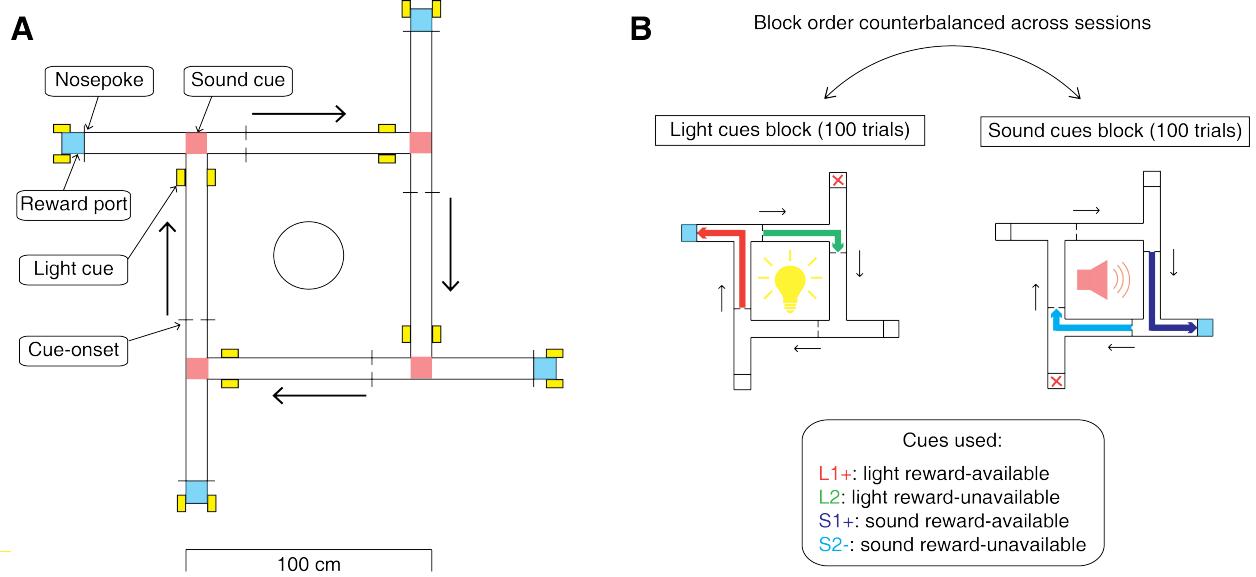
- 662 Hippocampal Ripples. *Journal of Neuroscience*, 24(29):6446–6456.
- 663 Peters, J., LaLumiere, R. T., and Kalivas, P. W. (2008). Infralimbic Prefrontal Cortex Is Responsible for Inhibiting Cocaine Seeking  
664 in Extinguished Rats. *Journal of Neuroscience*, 28(23):6046–6053.
- 665 Rescorla, R. A. and Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and  
666 Instrumental Learning. *Psychological Review*, 74(3):151–182.
- 667 Robinson, T. E. and Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues  
668 Through the Study of Individual Differences. *Biological Psychiatry*, 65(10):869–873.
- 669 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., and Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the  
670 Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42):13365–13376.
- 671 Roitman, M. F., Wheeler, R. A., and Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for rewarding and  
672 aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4):587–597.
- 673 Saddoris, M. P., Stamatakis, A., and Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nu-  
674 cleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*,  
675 33(12):2274–2287.
- 676 Salamone, J. D. and Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3):470–485.
- 677 Saunders, B. T. and Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of pavlovian-conditioned  
678 responses. *European Journal of Neuroscience*, 36(4):2521–2532.
- 679 Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1):23–32.
- 680 Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.
- 681 Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning.  
682 *Neuron*, 38(4):625–636.
- 683 Shidara, M., Aigner, T. G., and Richmond, B. J. (1998). Neuronal signals in the monkey ventral striatum related to progress through  
684 a predictable series of trials. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18(7):2613–25.
- 685 Shmuel Lissek, A. N. K. (2013). Generalization of Conditioned Fear and Obsessive-Compulsive Traits. *Journal of Psychology &*  
686 *Psychotherapy*, 7:3.
- 687 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., and Buzsáki, G. (2017). Cocaine place conditioning strengthens location-  
688 specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, pages 1–10.
- 689 Sleezer, B. J., Castagno, M. D., and Hayden, B. Y. (2016). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.  
690 *Journal of Neuroscience*, 36(44):11223–11237.
- 691 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., and Hayden, B. Y. (2016). Neuronal selectivity for spatial  
692 positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3):1098–1111.
- 693 Sugam, J. A., Saddoris, M. P., and Carelli, R. M. (2014). Nucleus accumbens neurons track behavioral preferences and reward

- 694       outcomes during risky decision making. *Biological Psychiatry*, 75(10):807–816.
- 695   Sutton, R. and Barto, A. (1998). Reinforcement Learning: An Introduction.
- 696   Tabuchi, E. T., Mulder, A. B., and Wiener, S. I. (2000). Position and behavioral modulation of synchronization of hippocampal and
- 697       accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6):717–728.
- 698   Takahashi, Y. K., Langdon, A. J., Niv, Y., and Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled
- 699       by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1):182–193.
- 700   van der Meer, M. A. A. and Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward
- 701       Information. *Journal of Neuroscience*, 31(8):2843–2854.
- 702   West, E. A. and Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after
- 703       Reinforcer Devaluation. *Journal of Neuroscience*, 36(4):1128–1139.
- 704   Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., and Mulder, A. B. (2003). Spatial and behavioral correlates
- 705       in nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,
- 706       1250(C):275–292.
- 707   Yun, I. A., Wakabayashi, K. T., Fields, H. L., and Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral
- 708       and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, 24(12):2923–2933.

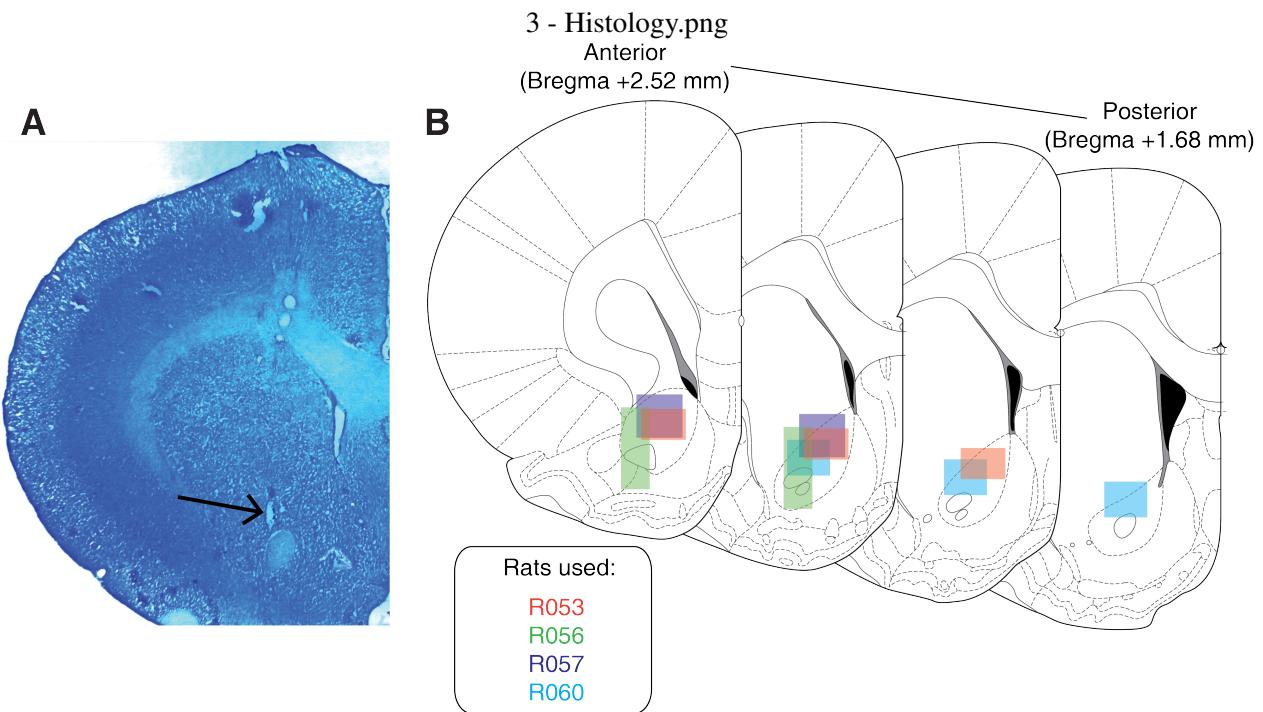


**Figure 1:** Schematic of potential coding strategies for cue identity (light, sound) and cue outcome (reward-available, reward-unavailable) employed by single units in the NAc across different units (A) and in time (B). **A:** Displayed are schematic PETHs illustrating putative responses to different cues under different hypotheses of how cue identity and outcome are coded. H1 (left panel): Coding of cue identity is absent in the NAc. Top: Unit A encodes a motivationally relevant variable, such as expected outcome, similarly across other cue features, such as cue identity or physical location. Hypothetical plot is firing rate across time. L1+ (red) signifies a reward-available light cue, S1+ (navy blue) a reward-available sound cue, L2- (green) a reward-unavailable light cue, S2- (light blue) a reward-unavailable sound cue. Dashed line indicates onset of cue. Bottom: No units within the NAc discriminate their firing according to cue identity. H2 (middle panel): Coding of cue identity occurs independently of encoding of motivationally relevant variables such as expected outcome or subsequent vigor. Top: Same as H1, with unit A discriminating between reward-available and reward-unavailable cues. Bottom: Unit B discriminates firing across stimulus modalities, depicted here as firing to light cues but not sound cues. H3 (right panel): Coding of cue identity is integrated with coding of other motivationally relevant variables. Hypothetical example demonstrating a unit that responds to outcome-predictive cues, but firing rate is also modulated by cue identity, firing most for the reward-available light cue. **B:** Displayed are schematic PETHs illustrating potential ways in which cue identity signals may persist over time. H1 (left panel): Cue-onset triggers a transient response to a unit that codes for cue identity. Dashed lines indicate time of a behavioral or environmental event. 'Cue-ON' signifies onset of cue, 'NP' signifies when the rat holds a nosepoke at a reward receptacle, 'Out' signifies when the outcome is revealed, 'OFF' signifies when the cue turns off. H2 (middle panel): Coding of cue identity persists during a nosepoke hold period until outcome is revealed. Coding can either be maintained by the same unit as during cue-onset (H2A) or by a sequence of units (H2B). H3 (right panel): Coding of cue identity persists after the outcome is received when the rat gets feedback about his decision, by either the same unit as during cue-onset (H3A) or by a sequence of units (H3B). The same hypotheses apply to other information-containing aspects of the environment when the cue is presented, such as the physical location of the cue.

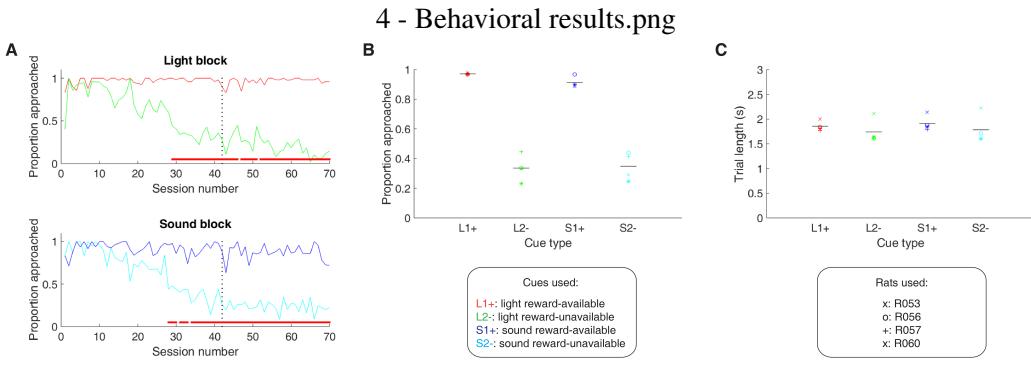
2 - Schematic task.png



**Figure 2:** Schematic of behavioral task. **A:** To scale depiction of square track consisting of multiple identical T-choice points. At each choice point, the availability of 12% sucrose reward at the nearest reward receptacle (light blue fill) was signaled by one of four possible cues, presented when the rat initiated a trial by crossing a photobeam on the track (dashed lines). Photobeams at the ends of the arms by the receptacles registered Nosepokes (solid lines). Rectangular boxes with yellow fill indicate location of LEDs used for light cues. Speakers for tone cues were placed underneath the choice points, indicated by magenta fill on track. Arrows outside of track indicate correct running direction. Circle in the center indicates location of pedestal during pre- and post-records. Scale bar is located beneath the track. **B:** Progression of a recording session. A session was started with a 5 minute recording period on a pedestal placed in the center of the apparatus. Rats then performed the light and sound blocks of the cue discrimination task in succession for 100 trials each, followed by another 5 minute recording period on the pedestal. Left in figure depicts a light block, showing an example trajectory for a correct reward-available (approach trial; red) and reward-unavailable (skip trial; green) trial. Right in figure depicts a sound block, with a reward-available (approach trial; navy blue) and reward-unavailable (skip trial; light blue) trial. Ordering of the light and sound blocks was counterbalanced across sessions. Reward-available and reward-unavailable cues were presented pseudo-randomly, such that not more than two of the same type of cue could be presented in a row. Location of the cue on the track was irrelevant for behavior, all cue locations contained an equal amount of reward-available and reward-unavailable trials.

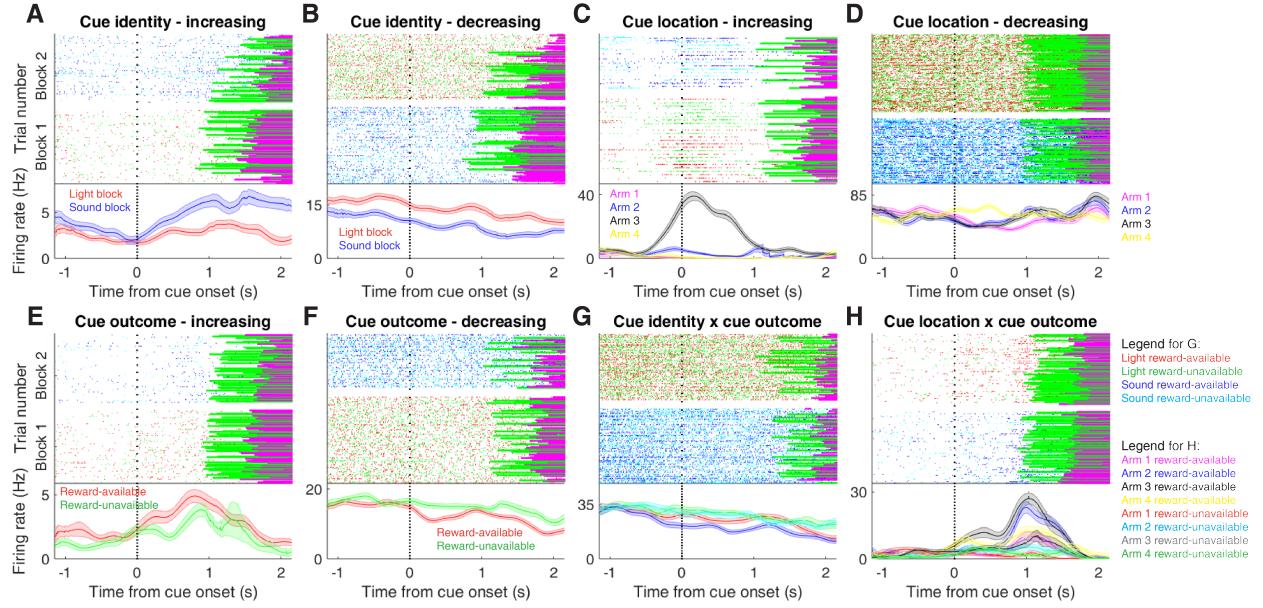


**Figure 3:** Histological verification of recording sites. Upon completion of experiments, brains were sectioned and tetrode placement was confirmed. **A:** Example section from R060 showing a recording site in the NAc core just dorsal to the anterior commissure (arrow). **B:** Schematic showing recording areas for all subjects.

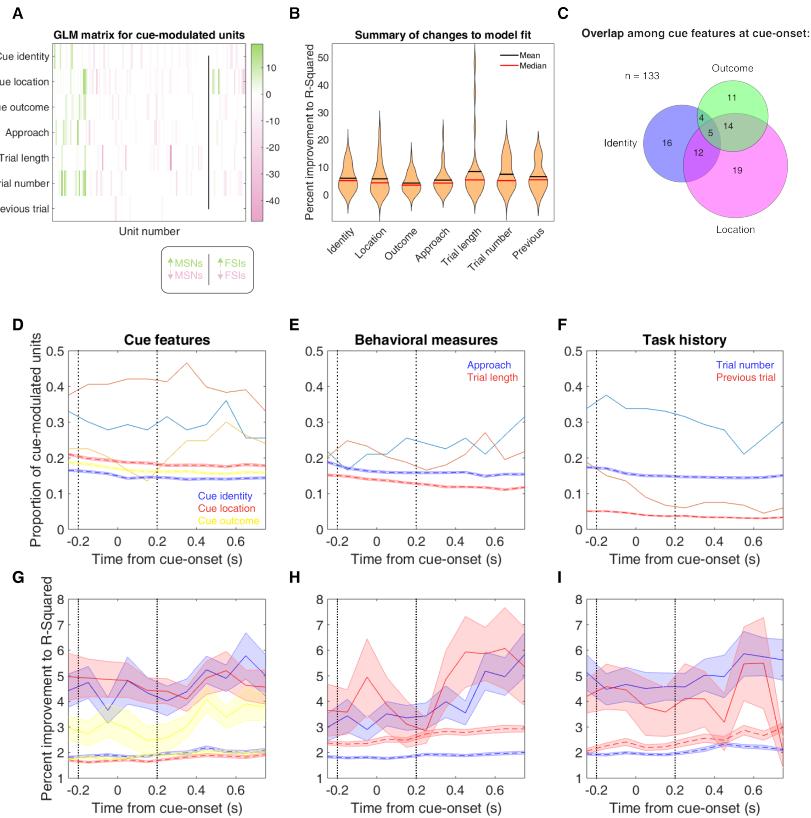


**Figure 4:** Performance on the behavioral task. **A.** Example learning curves across sessions from a single subject (R060) showing the proportion approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light block, light blue line for sound block) for light (top) and sound (bottom) blocks. Fully correct performance corresponds to an approach proportion of 1 for reward-available trials and 0 for reward-unavailable trials. Rats initially approach on both reward-available and reward-unavailable trials, and learn with experience to skip non-rewarded trials. Red bars indicate days in which a rat statistically discriminated between reward-available and reward-unavailable cues, determined by a chi square test. Dashed line indicates time of electrode implant surgery. **B-C:** Summary of performance during recording sessions for each rat. **B:** Proportion approached for all rats, averaged across all recording sessions. Different columns indicate the different cues (reward-available (red) and reward-unavailable (green) light cues, reward-available (navy blue) and reward-unavailable (light blue) sound cues). Different symbols correspond to individual subjects; horizontal black line shows the mean. All rats learned to discriminate between reward-available and reward-unavailable cues, as indicated by the clear difference of proportion approached between reward-available (~90% approached) and reward-unavailable cues (~30% approached), for both blocks (see Results for statistics). **C:** Average trial length for each cue. Note that the time to complete a trial was comparable for the different cues.

5 - Neural examples.png



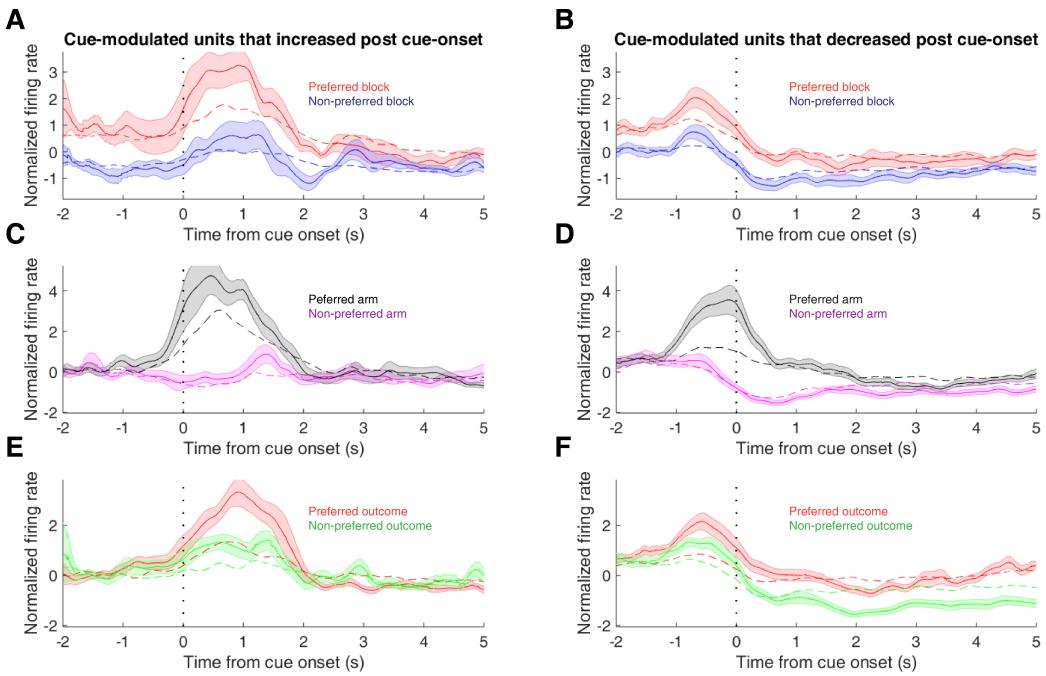
**Figure 5:** Examples of different cue-modulated NAc units influenced by various task parameters. **A:** Example of a cue-modulated NAc unit that showed an increase in firing following the cue, and encoded cue identity. Top: rasterplot showing the spiking activity across all trials aligned to cue-onset. Spikes across trials are color-coded according to cue type (red: reward-available light; green: reward-unavailable light; navy blue: reward-available sound; light blue: reward-unavailable sound). Green and magenta bars indicate trial termination when a rat initiated the next trial or made a nosepoke, respectively. White space halfway up the rasterplot indicates switching from one block to the next. Dashed line indicates cue onset. Bottom: PETHs showing the average smoothed firing rate for the unit for trials during light (red) and sound (blue) blocks, aligned to cue-onset. Lightly shaded area indicates standard error of the mean. Note this unit showed a larger increase in firing to sound cues. **B:** An example of a unit that was responsive to cue identity as in A, but for a unit that showed a decrease in firing to the cue. Note the sustained higher firing rate during the light block. **C-D:** Cue-modulated units that encoded cue location, each color in the PETHs represents average firing response for a different cue location. **C:** The firing rate of this unit only changed on arm 3 of the task. **D:** Firing decreased for this unit on all arms but arm 4. **E-F:** Cue-modulated units that encoded cue outcome, with the PETHs comparing reward-available (red) and reward-unavailable (green) trials. **E:** This unit showed a slightly higher response during presentation of reward-available cues. **F:** This unit showed a dip in firing when presented with reward-available cues. **G-H:** Examples of cue-modulated units that encoded multiple cue features. **G:** This unit integrated cue identity and outcome. **H:** An example of a unit that integrated cue identity and location.



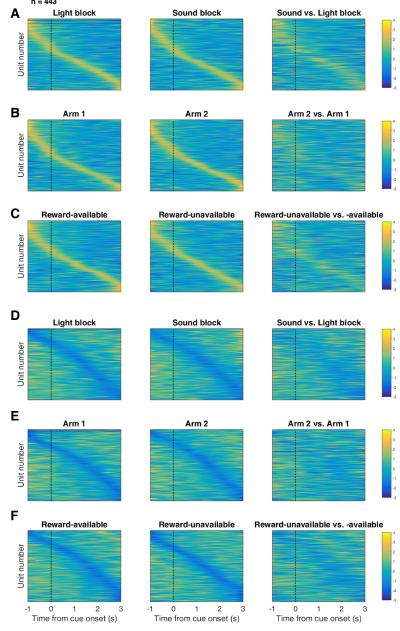
6 - GLM.png

**Figure 6:** Summary of influence of various task parameters on cue-modulated NAc units after cue-onset. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank test. Each row represents a given task parameter, and each column corresponds to a single unit. Colors indicate how much of the firing rate variance an individual predictor contributed to the model, as measured by differences in R-squared between the final model and the model minus the predictor of interest. Ordering from left to right: MSNs that increased firing in response to the cue (green, left of line), MSNs with a decreasing response (red, left of line), FSIs with an increasing response (green, right of line), FSIs with a decreasing response (red, right of line). Darker shades indicate more firing rate variance explained by a given predictor. Black line indicates separation of MSNs and FSIs. Scale bar indicates range of improvements to model fit for units with an increasing (green) and decreasing (red) response to the cue. **B:** Violin plots demonstrating changes in R-squared values with the addition of each of the individual predictors. The mean, median, and distribution of changes in R-squared values is plotted for each of the seven task parameters used in the GLM. **C:** Venn diagram illustrating the number of cue-modulated units encoding cue identity (blue circle), cue location (green circle), cue outcome (pink circle), as well as the overlap among units that encoded multiple cue features. **D-F:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of cue-onset. **D:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), cue location (red solid line), and cue outcome (yellow solid line) significantly contributed to the model at various time epochs relative to cue-onset. Dashed colored lines indicate the average of shuffling the firing rate order that went into the GLM 100 times. Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the GLM. **E:** Same as D, but for approach behavior and trial length. **F:** Same as D, but for trial number and previous trial. **G-I:** Average improvement to model fit. **G:** Average percent improvement to R-squared for units where cue identity, cue location, or cue outcome were significant contributors to the final model for time epochs surrounding cue onset. Shaded area around mean represents the standard error of the mean. **H:** Same as G, but for approach behavior and trial length. **I:** Same as G, but for trial number and previous trial.

## 7 - Population averages.png



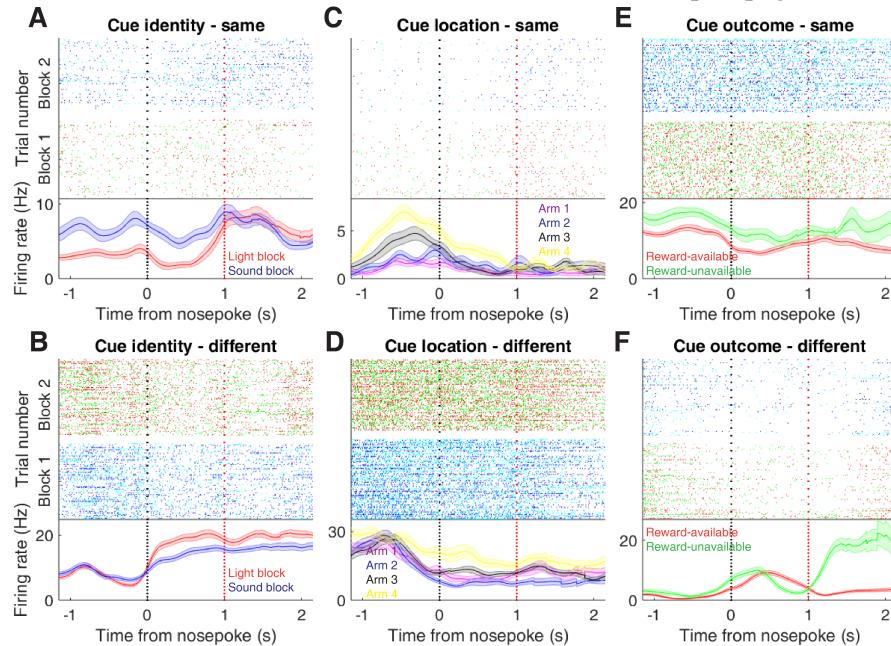
**Figure 7:** Population-level averages of cue feature sensitive NAc units. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to cue-onset. Activity is plotted for preferred stimulus block (red) and non-preferred stimulus block (blue). Dashed vertical line indicates onset of cue. Dashed color lines indicate the result of shuffling the identity of the units used for this average 1000 times. Lightly shaded area indicates standard error of the mean. Note larger increase to preferred stimulus block over nonpreferred stimulus block. Black lines indicate the average of 1000 rounds of random sampling of units from the non-drifting population for the preferred and non-preferred blocks. **B:** Same as A but for units that decreased in firing. Note population level activity reveals units classified as decreasing in response to cue show a biphasic response at the population level, with a transient increase around the time the rat starts on the arm, followed by a minimum after cue onset. Also, note the sustained difference in firing between the two blocks. **C-D:** Same as A-B for cue location. Activity is plotted for most preferred arm (black) and least preferred arm (magenta). **E-F:** Same as A-B for cue outcome. Activity is plotted for preferred expected outcome (red), and nonpreferred outcome (green). Note the larger increase to the cue representing the units preferred outcome (E), and the sustained decrease to the nonpreferred outcome (F).



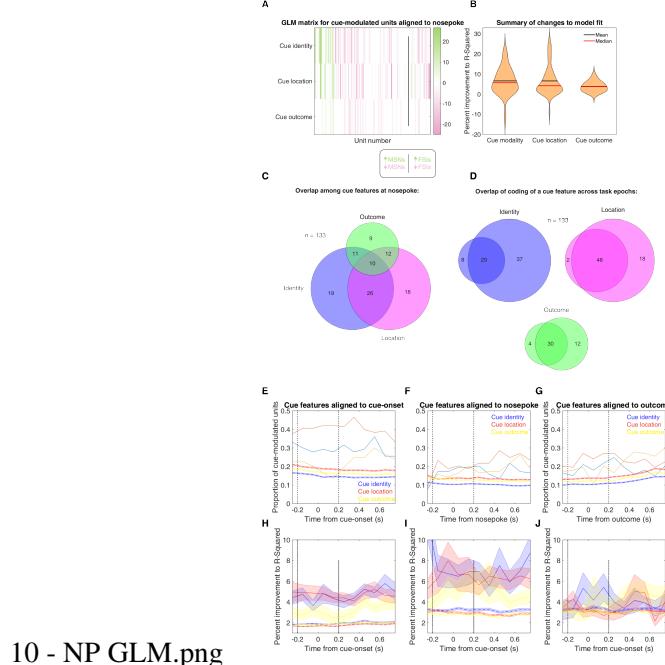
8 - Task tiling.png

**Figure 8:** Distribution of NAc firing rates across time surrounding cue onset. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue onset), averaged across all trials for a specific cue type, indicated by text labels. **A-C:** Heat plots aligned to normalized peak firing rates. **A, left:** Heat plot showing smoothed normalized firing activity of all recorded NAc units ordered according to the time of their peak firing rate during the light block. Each row is a units average activity across time to the light block. Dashed line indicates cue onset. Notice the yellow band across time, indicating all aspects of visualized task space were captured by the peak firing rates of various units. **A, middle:** Same units ordered according to the time of the peak firing rate during the sound block. Note that for both blocks, units tile time approximately uniformly with a clear diagonal of elevated firing rates. **A, right:** Unit firing rates taken from the sound block, ordered according to peak firing rate taken from the light block. Note that a weaker but still discernible diagonal persists, indicating partial similarity between firing rates in the two blocks. A similar pattern exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text). **B:** Same layout as in A, except that the panels now compare two different locations on the track instead of two cue modalities. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C:** Same layout as in A, except that panels now compare reward-available and reward-unavailable trials. **D-F:** Heat plots aligned to normalized minimum firing rates. **D:** Responses during different stimulus blocks as in A, but with units ordered according to the time of their minimum firing rate. **E:** Responses during trials on different arms as in B, but with units ordered by their minimum firing rate. **F:** Responses during cues signalling different outcomes as in C, but with units ordered by their minimum firing rate. Overall, NAc units "tiled" experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

9 - NP Neural examples.png



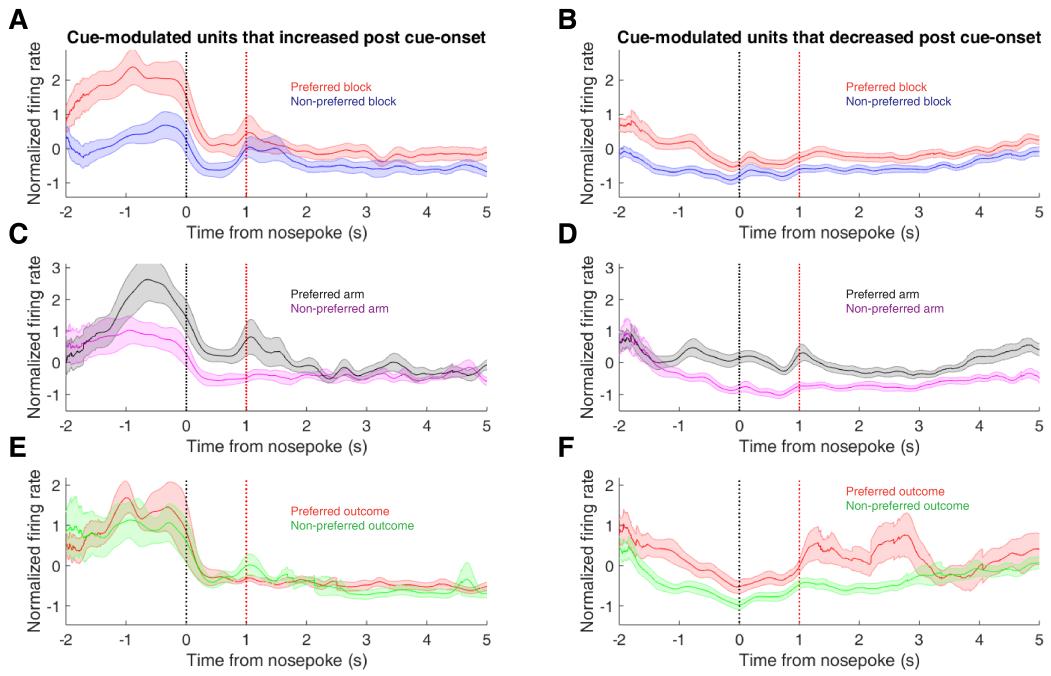
**Figure 9:** Examples of cue-modulated NAc units influenced by various task parameters at time of nosepoke. **A:** Example of a cue-modulated NAc unit that encoded cue identity at both cue-onset and during nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to nosepoke. Spikes across trials are color coded according to cue type (red: reward-available light; green: reward-unavailable light; navy blue: reward-available sound; light blue: reward-unavailable sound). White space halfway up the rasterplot indicates switching from one block to the next. Black dashed line indicates nosepoke. Red dashed line indicates receipt of outcome. Bottom: PETHs showing the average smoothed firing rate for the unit for trials during light (red) and sound (blue) blocks, aligned to nosepoke. Lightly shaded area indicates standard error of the mean. Note this unit showed a sustained increase in firing to sound cues during the trial. **B:** An example of a unit that was responsive to cue identity at time of nosepoke but not cue-onset. **C-D:** Cue-modulated units that encoded cue location, at both cue-onset and nosepoke (C), and only nosepoke (D). Each color in the PETHs represents average firing response for a different cue location. **E-F:** Cue-modulated units that encoded cue outcome, at both cue-onset and nosepoke (E), and only nosepoke (F), with the PETHs comparing reward-available (red) and reward-unavailable (green) trials.



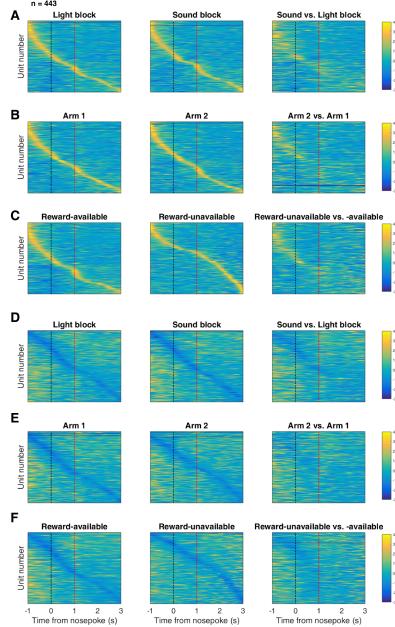
10 - NP GLM.png

**Figure 10:** Summary of influence of various task parameters of cue-modulated NAc units during nosepoke. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank test. Each row represents a given task parameter, and each column corresponds to a single unit. Colors indicate how much of the firing rate variance an individual predictor contributed to the model, as measured by differences in R-squared between the final model and the model minus the predictor of Interest. Ordering from left to right: MSNs that increased firing in response to the cue (green, left of line), MSNs with a decreasing response (red, left of line), FSIs with an increasing response (green, right of line), FSIs with a decreasing response (red, right of line). Darker shades indicate more firing rate variance explained by a given predictor. Black line indicates separation of MSNs and FSIs. **B:** Violin plots demonstrating changes in R-squared values with the addition of each of the individual predictors. The mean, median, and distribution of changes in R-squared values is plotted for each of the three task parameters that were significant predictors in the GLM. **C:** Venn diagram illustrating the number of cue-modulated units encoding cue identity (blue circle), cue location (green circle), cue outcome (pink circle), as well as the overlap among units that encoded multiple cue features. **D:** Venn diagram illustrating the number of cue-modulated units encoding cue identity, cue location, and cue outcome during cue-onset (left circle), during nosepoke hold (right circle), and during both epochs (overlap). **E-G:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of cue-onset (E), nosepoke (F), and outcome (G). **E:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), cue location (red solid line), and cue outcome (yellow solid line) significantly contributed to the model at various time epochs relative to cue-onset. Dashed colored lines indicate the average of shuffling the firing rate order that went into the GLM 100 times. Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the GLM. **F:** Same as E, but for time epochs relative to nosepoke where the rat waited for the outcome. **G:** Same as E, but for time epochs relative to receipt of outcome after the rat got feedback about his approach. **H-J:** Average improvement to model fit. **H:** Average percent improvement to R-squared for units where cue identity, cue location, or cue outcome were significant contributors to the final model for time epochs relative to cue-onset. Shaded area around mean represents the standard error of the mean. **I:** Same as H, but for time epochs relative to nosepoke. **J:** Same H, but for time epochs relative to receipt of outcome.

## 11 - NP population averages.png



**Figure 11:** Population-level averages of cue feature sensitive NAc units during a nosepoke. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to nosepoke with reward delivery occurring 1 s after nosepoke. Activity is plotted for preferred stimulus block (red) and nonpreferred stimulus block (blue). Black vertical dashed line indicates nosepoke. Red vertical dashed line indicates reward delivery occurring 1 s after nosepoke for reward-available trials. Dashed color lines indicate the result of shuffling the identity of the units used for this average 1000 times. Lightly shaded area indicates standard error of the mean. Note larger increase leading up to nosepoke to preferred stimulus block over nonpreferred stimulus block. **B:** Same as A but for units that decreased in firing. Note the sustained difference in firing between the two blocks. **C-D:** Same as A-B for cue location. Activity is plotted for most preferred arm (black) and least preferred arm (magenta). **E-F:** Same as A-B for cue outcome. Activity is plotted for preferred expected outcome (red), and nonpreferred outcome (green). Note the peak after outcome receipt for preferred outcome in decreasing units (F).



12 - NP task tiling.png

**Figure 12:** Distribution of NAc firing rates across time surrounding nosepoke for approach trials. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates nosepoke), averaged across all approach trials for a specific cue type, indicated by text labels. **A-C:** Heat plots aligned to normalized peak firing rates. **A, far left:** Heat plot showing smoothed normalized firing activity of all recorded NAc units ordered according to the time of their peak firing rate during the light block. Each row is a units average activity across time to the light block. Black dashed line indicates nosepoke. Red dashed line indicates reward delivery occurring 1 s after nosepoke for reward-available trials. Notice the yellow band across time, indicating all aspects of visualized task space were captured by the peak firing rates of various units. **A, middle:** Same units ordered according to the time of the peak firing rate during the sound block. Note that for both blocks, units tile time approximately uniformly with a clear diagonal of elevated firing rates, and a clustering around outcome receipt. **A, right:** Unit firing rates taken from the sound block, ordered according to peak firing rate taken from the light block. Note that a weaker but still discernible diagonal persists, indicating partial similarity between firing rates in the two blocks. A similar pattern exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text). **B:** Same layout as in A, except that the panels now compare two different locations on the track instead of two cue modalities. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C:** Same layout as in A, except that panels now compare correct reward-available and incorrect reward-unavailable trials. The disproportionate tiling around outcome receipt for reward-available, but not reward-unavailable trials suggests encoding of reward receipt by NAc units. **D-F:** Heat plots aligned to normalized minimum firing rates. **D:** Responses during different stimulus blocks as in A, but with units ordered according to the time of their minimum firing rate. **E:** Responses during trials on different arms as in B, but with units ordered by their minimum firing rate. **F:** Responses during cues signalling different outcomes as in C, but with units ordered by their minimum firing rate. Overall, NAc units "tiled" experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

<b>Task parameter</b>	<b>Total</b>	$\uparrow$ <b>MSN</b>	$\downarrow$ <b>MSN</b>	$\uparrow$ <b>FSI</b>	$\downarrow$ <b>FSI</b>
All units	443	155	216	27	45
<i>Rat ID</i>					
R053	145	51	79	4	11
R056	70	12	13	17	28
R057	136	55	75	3	3
R060	92	37	49	3	3
Analyzed units	344	117	175	18	34
Cue modulated units	133	24	85	6	18
<i>GLM aligned to cue-onset</i>					
Cue identity	37	7	21	1	8
Cue location	50	13	27	3	7
Cue outcome	34	10	18	0	6
Approach behavior	31	8	18	1	4
Trial length	25	5	18	0	2
Trial number	32	11	12	1	8
Previous trial	5	0	5	0	0
<i>GLM aligned to nosepoke</i>					
Cue identity	66	14	36	2	14
Cue location	66	14	40	3	9
Cue outcome	42	8	29	0	5
<i>GLM aligned to outcome</i>					
Cue outcome	10	0	6	0	4

**Table 1:** Units overview