

# **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**

2 In reinforcement learning, adaptive behavior requires assigning credit to the appropriate reward-predictive  
3 stimuli in the environment. The nucleus accumbens has been shown to be important for learning from  
4 feedback, and biasing and invigorating behavior in response to reward-predictive cues. Much work has  
5 emphasized nucleus accumbens coding of reward features of reward-predictive stimuli such as magnitude  
6 and identity of reward. However, not much is known as to how features of cues themselves are encoded.  
7 To test this we designed a decision making task where rats had to learn multiple sets of reward-predictive  
8 cues, and recorded single-unit activity in the nucleus accumbens during performance. We found that coding  
9 of various cue features occurred alongside coding of expected outcome. Furthermore, we found that this  
10 coding persisted during a delay period, after the rat made a decision and was waiting for an outcome, but not  
11 after the outcome was revealed. This finding of persistent coding of cue features in the nucleus accumbens  
12 could be one way in which the brain maintains an eligibility trace of outcome-predictive stimuli to be used  
13 for updating stimulus-outcome associations to inform future behavior.

## **14 Significance Statement (120 words)**

15 to do

## **16    Introduction**

17   A fundamental problem faced by all reinforcement learning agents is which features and/or actions to credit  
18   and blame for particular outcomes (Sutton and Barto, 1998). Adaptive behavior requires associating motivationally relevant outcomes with the cues that predict them through learning from feedback. Much work has  
19   focused on value signals such as reward prediction errors (RPEs), state values and action values (Lee et al.,  
20   2012; Maia, 2009). However, successful learning requires not only a RPE but also a trace of the preceding  
21   actions and/or cues. For example, if you become ill after eating food X in restaurant A, depending on if  
22   you remember the identity of the restaurant or the food at the time of illness, you may learn to avoid all  
23   restaurants; A only; X only; all foods; specific pairing of X-in-A.. Thus a complete understanding of what is  
24   learned following feedback requires understanding what trace is maintained.

26   Theories of nucleus accumbens (NAc) function generally agree that this brain structure contributes to motivated behavior, with some emphasizing a role in learning from RPEs (Averbeck and Costa, 2017; Joel et al.,  
27   2002; Khamassi and Humphries, 2012; Lee et al., 2012; Maia, 2009; Schultz, 2016)(see also the addiction  
28   literature on effects of drug rewards; Carelli, 2010; Hyman et al., 2006; Kalivas and Volkow, 2005) and  
29   others a role in the modulation of ongoing behavior through stimuli associated with motivationally relevant  
30   outcomes (invigorating, directing; Floresco, 2015; Nicola, 2010; Salamone and Correa, 2012). These proposals echo similar ideas on the functions of the neuromodulator dopamine (Berridge, 2012; Maia, 2009;  
31   Salamone and Correa, 2012; Schultz, 2016), with which the NAc is tightly linked functionally as well as  
32   anatomically (Cheer et al., 2007; du Hoffmann and Nicola, 2014; Ikemoto, 2007; Takahashi et al., 2016).

35   Much of our understanding of NAc function comes from studies of how cues that predict motivationally relevant outcomes (e.g. reward) influence behavior and neural activity in the NAc. Task designs that associate such cues with rewarding outcomes provide a convenient access point eliciting conditioned responses such as sign-tracking and goal-tracking (Hearst and Jenkins, 1974; Robinson and Flagel, 2009), pavlovian-instrumental transfer (Estes, 1943; Rescorla and Solomon, 1967) and enhanced response vigor (Nicola, 2010;

40 Niv et al., 2007), which tend to be affected by NAc manipulations (Chang et al., 2012; Corbit and Balleine,  
41 2011; Flagel et al., 2011) (although not always straightforwardly (is Hauber appropriate?); Chang and Hol-  
42 land, 2013; Hauber et al., 2000). Similarly, analysis of RPEs typically proceeds by establishing an associa-  
43 tion between a cue and subsequent reward, with NAc responses transferring from outcome to the cue with  
44 learning (Day et al., 2007; Roitman et al., 2005; Setlow et al., 2003; Schultz et al., 1997)(Day 2007 is the  
45 only paper referenced here that looks at transferring of responses from reward to CS+, other papers are in  
46 VTA, or the development of CS+ responses).

47 Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues  
48 (Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Goldstein et al., 2012;  
49 Hollerman et al., 1998; Lansink et al., 2012; McGinty et al., 2013; Nicola, 2004; Roesch et al., 2009; Roit-  
50 man et al., 2005; Saddoris et al., 2011; Setlow et al., 2003; Sugam et al., 2014; West and Carelli, 2016),  
51 much less is known about how reward-predictive cues themselves are encoded in the NAc (Sleeker et al.,  
52 2016). This is an important issue for at least two reasons. First, in reinforcement learning, motivationally  
53 relevant outcomes are typically temporally delayed relative to the cues that predict them. In order to solve  
54 the problem of assigning credit (or blame) across such temporal gaps, some trace of preceding activity needs  
55 to be maintained (Lee et al., 2012; Sutton and Barto, 1998). Since NAc is a primary target of DA signals  
56 interpretable as RPEs, and NAc lesions impair RPEs related to timing, its activity trace will help determine  
57 what can be learned when RPEs arrive (Hamid et al., 2015; Hart et al., 2014; Ikemoto, 2007; McDannald  
58 et al., 2011; Takahashi et al., 2016).

59 Second, for ongoing behavior, the relevance of cues typically depends on context. In experimental settings,  
60 context may include the identity of a preceding cue (Holland, 1992), spatial or configural arrangements  
61 (Honey et al., 2014), and unsignaled rules as occurs in set shifting and other cognitive control tasks (Grant  
62 and Berg, 1948). In such situations, the question arises how selective, context-dependent processing of  
63 reward-predictive cues is implemented. For instance, is there a gate prior to NAc, or are all cues represented  
64 in NAc but their current values dynamically updated (FitzGerald et al., 2014; Sleeker et al., 2016).

65 Thus, both from a learning and a flexible performance perspective, it is of interest to determine how cue  
66 identity is represented in the brain, with NAc of particular interest given its anatomical and functional po-  
67 sition at the center of motivational systems. We sought to determine whether cue identity is represented in  
68 the NAc, if cue identity is represented alongside other motivationally relevant variables, such as cue value,  
69 and if these representations are maintained after a behavioral decision has been made (Figure 1). To address  
70 these questions, we recorded the activity of NAc neurons as rats performed a task in which multiple, distinct  
71 sets of cues predicted the same outcome.

72 [Figure 1 about here.]

## 73 **Methods**

### 74 **Subjects:**

75 Adult male Long-Evans rats ( $n = 4$ , Charles River, Saint Constant, QC) were used as subjects. Rats were in-  
76 dividually housed with a 12/12-h light-dark cycle, and tested during the light cycle. Rats were food deprived  
77 to 85-90% of their free feeding weight (weight at time of implantation was 440 - 470 g), and water restricted  
78 4-6 hours before testing. All experimental procedures were approved by the the University of Waterloo An-  
79 imal Care Committee (protocol# 11-06) and carried out in accordance with Canadian Council for Animal  
80 Care (CCAC) guidelines.

### 81 **Overall timeline:**

82 Each rat was first handled for seven days during which they were exposed to the experiment room, the su-  
83 crose solution used as a reinforcer, and the click of the sucrose dispenser valves. Rats were then shaped to

84 run on the task for sucrose reward while being restricted to running in the clockwise direction by presenting  
85 a physical barrier to running counterclockwise. Cues signaling the availability and inavailability of reward  
86 were then introduced (described in detail below, and rats were trained for 200 trials per day until they dis-  
87 criminated between the cues. At this point they underwent hyperdrive implantation targeted at the NAc. Rats  
88 were allowed to recover for a minimum of five days before being retrained on the task, and recording began  
89 once performance returned to pre-surgery levels. Upon completion of recording, animals were euthanized  
90 and recording sites were histologically confirmed.

91 **Behavioral task and training:**

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

107 Each session consisted of both a *light block* and a *sound block* with 100 trials each. Within a block, one cue

108 signaled reward was available on that trial (L1+ or S1+), while the other signaled reward was not available  
109 (L2- or S2-). Light block cues were a flashing white light, and a constant yellow light. Sound block cues  
110 were a 2 kHz sine wave and a 8 kHz sine wave whose amplitude was modulated from 0 to maximum by a  
111 2 Hz sine wave. Reward-cue associations were counterbalanced across rats, e.g. for some rats L1+ was the  
112 flashing white light, and for others L1+ was the constant yellow light. The order of cue presentation was  
113 pseudorandomized so that the same cue could not be presented more than twice in a row. Block order within  
114 each day was also pseudorandomized, such that the rat could not begin a session with the same block for more  
115 than two days in a row. Each session consisted of a 5 minute pre-session period on a pedestal (a terracotta  
116 planter filled with towels), followed by the first block, then the second block, then a 5 minute post-session  
117 period on the pedestal. Rats were trained daily until they could distinguish between the reward-available and  
118 reward-unavailable cues for both light and sound blocks for three consecutive days, according to a chi-square  
119 test rejecting the null hypothesis of equal approaches for reward-available and reward-unavailable trials, at  
120 which point they underwent electrode implant surgery.

121 [Figure 2 about here.]

122 **Surgery:**

123 Surgical procedures were as described previously (Malhotra et al., 2015). Briefly, animals were administered  
124 analgesics and antibiotics, anesthetized with isoflurane, induced with 5% in medical grade oxygen and main-  
125 tained at 2% throughout the surgery ( 0.8 L/min). Rats were then chronically implanted with a “hyperdrive”  
126 consisting of 16 independently drivable tetrodes, either all 16 targeted for the right NAc (AP +1.4 mm and  
127 ML +1.6 mm relative to bregma; Paxinos and Watson, 2005), or 12 in the right NAc and 4 targeted at the  
128 mPFC (AP +3.0 mm and ML +0.6 mm, relative to bregma; only data from NAc tetrodes was analyzed).  
129 Following surgery, all animals were given at least five days to recover while receiving post-operative care,  
130 and tetrodes were lowered to the target (DV -6.0 mm) before being reintroduced to the behavioral task.

131 **Data acquisition and preprocessing:**

132 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with a  
133 RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex). Sig-  
134 nals were referenced against a tetrode placed in the corpus callosum above the NAc. Candidate spikes for  
135 sorting into putative single units were obtained by band-pass filtering the data between 600-9000 Hz, thresh-  
136 olding and aligning the peaks (UltraMegaSort2k, Hull et al., 2011). Spike waveforms were then clustered  
137 with KlustaKwik using energy and the first derivative of energy as features, and manually sorted into units  
138 (MClust 3.5, A.D. Redish et al.). Isolated units containing a minimum of 200 spikes within a session were  
139 included for subsequent analysis. Units were classified as fast spiking interneurons (FSIs) by an absence of  
140 interspike intervals (ISIs) > 2 s, while medium spiny neurons (MSNs) had a combination of ISIs > 2 s and  
141 phasic activity with shorter ISIs (Barnes 2005, Atallah 2014).

142 **Data analysis:**

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

150 *Neural data.* To investigate the contribution of different cue features (cue identity and cue outcome) on the  
151 firing rates of NAc single units, we first determined whether firing rates for a unit were modulated by the  
152 onset of a cue by collapsing across all cues and comparing the firing rates for the 1 s preceding cue-onset  
153 with the 1 s following cue-onset. Single units were considered to be *cue-modulated* if a Wilcoxon signed-

rank test comparing pre- and post-cue firing was significant at  $p < .01$ . Cue-modulated units were then classified as either increasing or decreasing if the post-cue activity was higher or lower than the pre-cue activity, respectively.

To determine the relative contribution of different task parameters to firing rate variance (as in Figures 5-6), a forward selection stepwise general linear model (GLM) was fit to each cue-modulated unit. Cue identity (light block, sound block), cue location (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available, reward-unavailable), behavior (approach, skip), trial length, trial number, and trial history (reward availability on the previous 2 trials) were used as predictors, and the 1 s post-cue firing rate as the response variable. Units were classified as being modulated by a given task parameter if addition of the parameter significantly improved model fit using deviance as the criterion ( $p < .01$ ). A comparison of the R-squared value between the final model and the final model minus the predictor of interest was used to determine the amount of firing rate variance explained by the addition of that predictor for a given unit.

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

To visualize NAc representations of task space within cue conditions, normalized spike trains for all units were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure

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

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

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

197 **Histology:**

198 Upon completion of the experiment, rats were anesthetized with 5% isoflurane, then asphyxiated with carbon  
199 dioxide. Transcardial perfusions were performed, and brains were fixed and removed. Brains were sliced  
200 in 50  $\mu\text{m}$  coronal sections and stained with thionin. Slices were visualized under light microscopy, tetrode

201 placement was determined, and electrodes with recording locations in the NAc were analyzed (Figure 3).

202 [Figure 3 about here.]

203 **Results**

204 **Behavior**

205 Rats were trained to discriminate between cues signaling the availability and absence of reward on a square  
206 track with four identical arms for two distinct sets of cues (Figure 2). During each session, rats were pre-  
207 sented sequentially with two behavioral blocks containing cues from different sensory modalities, a light and  
208 a sound block, with each block containing a cue that signalled the availability of reward (reward-available),  
209 and a cue that signalled the absence of reward (reward-unavailable). To maximize reward receipt, rats should  
210 approach reward sites on reward-available trials, and skip reward sites on reward-unavailable trials (see Fig-  
211 ure 4A for an example learning curve). All four rats learned to discriminate between the reward-available  
212 and reward-unavailable cues for both the light and sound blocks as determined by reaching significance ( $p <$   
213 .05) on a daily chi-square test comparing approach behavior for reward-available and reward-unavailable  
214 cues for each block, for at least three consecutive days (range for time to criterion: 22 - 57 days). Mainte-  
215 nance of behavioral performance during recording sessions was assessed using linear mixed effects models  
216 for both proportion of trials where the rat approached the receptacle, and trial length. Analyses revealed that  
217 the likelihood of a rat to make an approach was influenced by which of the cues was presented (Percentage  
218 approached: light reward-available = 97%; light reward-unavailable = 34%; sound reward-available = 91%;  
219 sound reward-unavailable 35%;  $p < .001$ ), but that cue identity did not significantly affect the length of  
220 time taken to complete a trial (Trial length: light reward-available = 1.85 s; light reward-unavailable = 1.74  
221 s; sound reward-available = 1.91 s; sound reward-unavailable 1.78 s;  $p = .13$ ; Figure 4C,D). Thus, during

222 recording, rats successfully discriminated the cues according to whether or not they signaled the availability  
223 of reward at the reward receptacle.

224 [Figure 4 about here.]

225 **NAc neurons encode behaviorally relevant and irrelevant cue features**

226 **Single unit responses discriminate cue features:**

227 We sought to address which parameters of our task were encoded by NAc activity, specifically whether the  
228 NAc encodes aspects of motivationally relevant cues not directly tied to reward, such as the identity and  
229 location of the cue, and whether this coding is independent or integrated with coding of cue outcome. To do  
230 this we recorded a total of 443 units with > 200 spikes in the NAc from 4 rats over 57 sessions while they  
231 performed a cue discrimination task (Table 1). Units that exhibited a drift in firing rate over the course of  
232 either block were excluded from further analysis, leaving 344 units for further analysis. The activity of 133  
233 (39%) of these 344 units were modulated by the cue, with more showing a decrease in firing ( $n = 103$ ) than  
234 an increase ( $n = 30$ ) around the time of cue-onset (Table 1). Within this group, 24 were classified as FSIs,  
235 while 109 were classified as SPNs. Upon visual inspection, we observed several patterns of firing activity,  
236 including units that discriminated firing upon cue-onset across various cue conditions, showed sustained  
237 differences in firing across cue conditions, had transient responses to the cue, showed a ramping of activity  
238 starting at cue-onset, and showed elevated activity immediately preceding cue-onset, for example (Figure 5).  
239 To characterize more formally whether these cue-evoked responses were modulated by various aspects of  
240 the task, we fit a GLM to each cue-modulated unit. Fitting GLMs revealed that a variety of task parameters  
241 accounted for a significant portion of firing rate variance in NAc cue-modulated units (Figure 6, Table 1).  
242 Notably, there were units that discriminated between whether the rat was performing in the light or sound  
243 block (28% of cue-modulated units, accounting for 6% of variance on average), which arm the rat was

244 currently on (38% of cue-modulated units, accounting for 6% of variance on average), and whether the rat  
245 was engaged in the common portion of a reward-available or reward-unavailable trial (26% of cue-modulated  
246 units, accounting for 4% of variance on average), suggesting that the NAc independently encodes features  
247 of reward-predictive cues separate from expected outcome (Figure 5A-F). Furthermore, overlap of coding of  
248 cue features within units was not different than expected by chance according to chi-square tests, suggesting  
249 for integrated coding across various aspects of a cue (Figure 5G,H). Fitting a GLM to all recorded units  
250 revealed a similar pattern of results (data not shown). Together, these findings show that various cue features  
251 are represented in the NAc, and that this coding has some overlap, but is mostly distinct from expected  
252 outcome (Figure 1; H2,H3).

253 [Table 1 about here.]

254 [Figure 5 about here.]

255 [Figure 6 about here.]

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

257 We observed a variety of single unit response profiles around the time of cue onset (Figure 5). To investigate  
258 whether these firing rate patterns were related to what cue features were encoded, we plotted the population  
259 level averages for units that were modulated by each feature. To do this, we normalized firing activity for  
260 each unit that was modulated by a given cue feature, such as light block, then generated the cue-onset aligned  
261 population average firing rate for each of the cue features (Figure 7). Overall, this analysis revealed that cells  
262 that showed an increase upon cue presentation had stronger responses for the preferred cue condition (Figure  
263 7A,C,E). Interestingly, units that were classified as decreasing in response to the cue showed a biphasic  
264 response at the population level, with a small peak at a time in alignment with entry into the arm, followed by  
265 a sustained dip after cue-onset (Figure 7B,D,F). Units that were modulated by cue identity showed a stronger

266 increase in response to the preferred task block, as well as a higher tonic firing rate to the preferred task block,  
267 most notably in units that decreased in firing rate to the cue (Figure 7A,B). Units that were modulated by cue  
268 location showed a graded response to locations of decreasing preference, with peak firing occurring around  
269 cue-onset (Figure 7C,D). Units that were modulated by cue outcome showed a ramping of activity after  
270 cue-onset for their preferred cue type. Additionally, units that exhibited a decrease in firing in response to  
271 the cue and whose activity was modulated by cue outcome, showed a sustained discriminatory response to  
272 reward-available and reward-unavailable cues that extended beyond cue-onset (Figure 7F). Together, these  
273 visualizations of the averaged population responses revealed nuanced differences in the way NAc units are  
274 modulated by cue conditions across cue features.

275 [Figure 7 about here.]

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

277 Given the varied time courses and response profiles of NAc units to various aspects of the cue, the NAc  
278 may be computing a temporally evolving state value signal (Pennartz., 2011). If this is the case, then the  
279 recruitment of NAc units should vary alongside changes in the environment. To look at the distribution of  
280 responses throughout our task space and see if this distribution is modulated by cue features, we z-scored  
281 the firing rate of each unit and plotted the normalized firing rates of all units aligned to cue-onset and sorted  
282 them according to the time of peak firing rate (Figure 8). We did this separately for both the light and  
283 sound blocks, and found a nearly uniform distribution of firing fields in task space that was not limited to  
284 alignment to the cue (Figure 8A). Furthermore, to determine if this population level activity was similar  
285 across blocks, we also organized firing during the sound blocks according to the ordering derived from the  
286 light blocks. (**JIMMIE EDIT THIS SECTION!**) This revealed that while there was some preservation of  
287 order, the overall firing was qualitatively different across the two blocks. To control for the possibility that  
288 any comparison of trials would produce this effect, we did a within block comparison, comparing half of  
289 the trials in the light block against the other half. This comparison looked similar to our test comparison of

290 sound block trials ordered by light block trials. Additionally, given that the majority of our units showed  
291 an inhibitory response to the cue, we also plotted the firing rates according to the lowest time in firing, and  
292 again found some maintenance of order, but largely different ordering across the two blocks, and the within  
293 block comparison (Figure 8B). To further test this, we divided each block into two halves and looked at the  
294 correlation of the average smoothed firing rates across various combinations of these halves across our cue-  
295 aligned centered epoch. A linear mixed effects model revealed that within block correlations (e.g. one half  
296 of light trials vs other half of light trials) were higher and more similar than across block correlations (e.g.  
297 half of light trials vs half of sound trials), suggesting that activity in the NAc discriminates across various cue  
298 conditions (within block correlations = .381, across block correlations = .342). This process was repeated  
299 for cue location and cue outcome, showing that NAc segmentation of the task is qualitatively different even  
300 during those parts of the task not immediately associated with a specific cue, action, or outcome, although  
301 the within condition comparison of reward-unavailable trials was less correlated than reward-available trials,  
302 and more similar to the across condition comparisons (Figure 8C-F).

303 [Figure 8 about here.]

304 **Encoding of cue features is maintained until outcome:**

305 In order to be useful for credit assignment in reinforcement learning, a trace of the cue must be maintained  
306 until the outcome, so that information about the outcome can be associated with the outcome-predictive cue.  
307 To test whether representations of cue features were maintained post-approach until the outcome was re-  
308 vealed, we fit a GLM to the post-approach firing rates of cue-modulated units aligned to the time of nosepoke  
309 into the reward receptacle. This analysis showed that a variety of units still discriminated firing according  
310 to various cue features, but not other task parameters, showing that NAc activity discriminates various cue  
311 conditions well into a trial (Table 1, Figures 9,10). Additionally, these units were a mix between most of the  
312 units that encoded cue features at cue-onset (observed overlap greater than expected by chance according to  
313 chi-square tests), and those that did not previously have a cue feature as a predictor (29, 48, and 30 units en-

314 coded both time points for cue identity, location, and outcome, respectively). Fitting a GLM to all recorded  
315 units revealed a similar pattern (data not shown). Population level averages for units that increased to cue-  
316 onset showed a ramping up of activity that peaked upon nosepoke, whereas units that decreased to cue-onset  
317 showed a gradual reduction of firing activity that reached a minimum upon nosepoke (Figure 11). Addition-  
318 ally, a peak is seen for preferred cue outcome in decreasing units at 1 second post cue-onset when reward was  
319 received, demonstrating an integration of expected and received reward (Figure 11F). Furthermore, aligning  
320 normalized peak firing rates to nosepoke onset, revealed a clustering of responses around reward receipt for  
321 all cue conditions where the rat would have received reward (Figure 12). To determine whether coding of cue  
322 features was maintained after the outcome was revealed, a GLM was fit to the firing rates of cue-modulated  
323 units at the time of reward receipt, during which the cue was still present. Fitting a GLM revealed 10 units  
324 (8%) where cue outcome accounted for an average of 32% of firing rate variance (Table 1, data not shown).  
325 However, an absence of cue identity or cue location coding was observed, suggesting that the NAc does not  
326 maintain a representation of these cue features once the rat receives behavioral feedback for its decision.

327 [Figure 9 about here.]

328 [Figure 10 about here.]

329 [Figure 11 about here.]

330 [Figure 12 about here.]

331 **Discussion**

332 The present study found evidence for coding, by partially overlapping populations of NAc units, of multiple  
333 identifying features of motivationally relevant stimuli in addition to expected outcome; the sensory modality  
334 of the presented cue, as well as its physical location on the track (H2 and H3 in Figure 1A). Importantly, this  
335 identity coding was maintained on approach trials during a delay period where the rat held a nosepoke until  
336 receipt of outcome (H2 in Figure 1B). Cells that coded different cue features exhibited different temporal  
337 profiles as a whole, although across all recorded cells a tiling of task structure was observed such that  
338 all points within our analyzed task space were accounted for by the ordered peak firing rates of all cells.  
339 Furthermore, this tiling differed between various conditions with a cue feature, such as light versus sound  
340 blocks.

341 **Cue identity:**

342 Our finding that NAc units can discriminate between outcome-predictive stimuli with similar biological  
343 significance, in the current study cues from different sensory modalities, expands upon an extensive rodent  
344 literature examining NAc correlates of conditioned stimuli (Atallah et al., 2014; Bissonette et al., 2013;  
345 Cooch et al., 2015; Day et al., 2006; Dejean et al., 2017; Goldstein et al., 2012; Lansink et al., 2012; McGinty  
346 et al., 2013; Nicola, 2004; Roesch et al., 2009; Roitman et al., 2005; Saddoris et al., 2011; Setlow et al., 2003;  
347 Sugam et al., 2014; West and Carelli, 2016). Units that discriminated across blocks were not simply due to  
348 drifting of the signal across trials, as units that showed a drift in firing between the first and second half within  
349 a block were excluded from the analysis. Furthermore, even though actions were stereotyped during correct  
350 trials, such that the rat always turned left at the decision point to approach for reward, and right to skip the  
351 receptacle and initiate the next trial, units that were modulated by the expected value of the cue maintained  
352 their specific firing patterns even during error trials where the rat turned left after presentation of the reward-  
353 unavailable cue, suggesting that these signals did not represent action values. Additionally, NAc signals have  
354 been shown to be modulated by response vigor, to detangle this from our results we included the trial length

355 as a predictor in our GLMs, and found units with cue feature correlates independent of trial length (McGinty  
356 et al., 2013). Perhaps the most comparable work in rodents comes from a study that found distinct coding  
357 for an odor when it predicted separate but equally valued rewards (Cooch et al., 2015). The present work is  
358 complementary to this as it shows that NAc units have representations of identifiable aspects of the cue itself,  
359 both separately and integrated with the reward it predicts (H2 and H3 in Figure 1A). Another study paired  
360 separate cues with appetitive or aversive outcomes, and found separate populations of units that encoded  
361 each cue, with many switching selectivity after reversal of the associations between the cues and outcomes,  
362 providing evidence that the NAc encodes the biological significance of stimuli (Setlow et al., 2003). Once  
363 again, our study was different as we recorded neural responses to distinct cues encoding the same anticipated  
364 outcome, suggesting that even when the biological relevance of these stimuli is similar the NAc dissociates  
365 their representations at the level of the single-unit, although this doesn't explain the difference in population  
366 averaged firing rate for the epochs when the cue is not present. A possible interpretation of this coding of cue  
367 features alongside expected outcome is that these representations are used to associate reward with relevant  
368 features of the environment, so-called credit assignment in the reinforcement learning literature (Sutton and  
369 Barto, 1998). A burgeoning body of human and non-human primate work has started to elucidate neural  
370 correlates of credit assignment in the PFC, particularly in the lateral orbitofrontal cortex (Akaishi et al., 2016;  
371 Asaad et al., 2017; Chau et al., 2015; Noonan et al., 2017). Given the importance of cortical inputs in NAc  
372 associative representations, it is possible that information related to credit assignment is relayed from the  
373 cortex to NAc (Cooch et al., 2015; Ishikawa et al., 2008).

374 Another possibility is that these modality specific units were encoding the context, rule, or sequence within  
375 a session as some units responded similarly for both reward-available and reward-unavailable cues within  
376 a block. This interpretation is in alignment with a recent paper from the non-human primate literature that  
377 recorded NAc responses during the Wisconsin Card Sorting Task (WCST), a common set-shifting task used  
378 in both the laboratory and clinic, and found units that preferred firing to stimuli when a certain rule, or rule  
379 category was currently active (Sleeker et al., 2016). Indeed, an encoding of the current strategy could be an  
380 explanation as to why a sustained difference in population averaged firing was seen across stimulus blocks, as

381 well as a potential explanation for the differentially tiling of task structure across blocks in the current study.  
382 Further support for a modulation of NAc responses by strategy comes from an fMRI study that examined  
383 BOLD levels during a set-shifting task (FitzGerald et al., 2014). In this task, participants learned two sets  
384 of stimulus-outcome contingencies, a visual set and auditory set. During testing they were presented with  
385 both simultaneously, and the stimulus dimension that was relevant was periodically shifted between the two.  
386 Here, they found that bilateral NAc activity reflected value representations for the currently relevant stimulus  
387 dimension, and not the irrelevant stimulus. The current finding of separate, but overlapping, populations of  
388 units encoding cue modality and expected value, suggests that the fMRI finding is generated by the combined  
389 activity of several different functional cell types.

390 A caveat of the current study is that rats were never presented with both sets of cues simultaneously, and  
391 thus never had to switch strategies, although extrapolating the data from the primate study suggests that the  
392 activity of the cue identity units would be modulated by relevance. Keeping along this theme, the current  
393 data set is unable to identify precisely what the modality-sensitive neurons were encoding, that is were they  
394 tracking representations of stimulus identity, a preferred context, or even a macroscale representation of  
395 progress through the session. Furthermore, their relevance for ongoing behavior is also uncertain. NAc core  
396 lesions have been shown to impair shifting between different behavioral strategies, and it is possible that  
397 selectively silencing the units that prefer responding for a given modality or rule would impair performance  
398 when the animal is required to use that information, or artificial enhancement of those units would cause  
399 them to use the rule when it is the inappropriate strategy.

400 **Encoding of position:**

401 Our finding that cue-evoked activity was modulated by cue location is in alignment with several previous  
402 reports (Lavoie and Mizumori, 1994; Mulder et al., 2005; Strait et al., 2016; Wiener et al., 2003). The NAc  
403 receives strong inputs from the hippocampus, and the communication of place-reward information across the  
404 two structures suggests that the NAc might be keeping track of locations associated with reward (Lansink

405 et al., 2008, 2009, 2016; Pennartz, 2004; Sjulson et al., 2017; Tabuchi et al., 2000; van der Meer and Redish,  
406 2011). Alternatively, it is possible that this positional modulation is part of a representation of task space,  
407 keeping track of where in the task the rat is. A previous non-human primate experiment showed that when  
408 reward is contingent upon completion of a series of trials, separate populations of NAc neurons signal the  
409 start of a schedule, subsequent trials in the schedule, and the first trial in extended schedules (Shidara et al.,  
410 1998). This signalling of position within a sequence has been observed in subsequent studies, and it is  
411 possible that the rats were keeping track of which specific arm they were in as part of a sequence of arms,  
412 especially since they started from the same arm during the start of each session (Atallah et al., 2014; Berke  
413 et al., 2009; Khamassi et al., 2008; Lansink et al., 2012; Mulder et al., 2004). Given that the current task was  
414 pseudo-random, it is possible that the rats learned the structure of sequential cue presentation, and the neural  
415 activity could reflect this. However, this is unlikely as including a previous trial variable in the analysis did  
416 not explain a significant amount of firing rate variance in response to the cue for the vast majority of cells.  
417 A limitation of the current study is that the behavioral relevance of cue location was not manipulated in the  
418 sense that varying cue location did not change the outcome-predictive nature of the cue.

419 **Potential functional consequences of maintained coding:**

420 In the current study we found that the coding of cue features persisted after the choice point, during a delay  
421 period while the rat waited at the receptacle for the outcome (H2 in Figure 1B). Having an enriched repre-  
422 sentation that includes details about the environmental context the animal finds itself in alongside expected  
423 outcome, and is maintained online until the outcome is revealed could be useful for assigning credit of a re-  
424 ward to the appropriate elements of an environment, to develop an accurate value function and optimize long  
425 term acquisition of reward (Lee et al., 2012). A recent non-human primate experiment found evidence for  
426 simultaneous coding of outcome and outcome-predictive stimuli at the time of feedback in the dorsolateral  
427 PFC, and that this coding of stimulus information was strongest before behavioral performance stabilized  
428 (Asaad et al., 2017). Interestingly, we found coding of cue identity and location after presentation of the  
429 cue, and during a delay period between nosepoke and outcome, but not after the outcome was received. The

absence of coding during feedback, and that these recordings were done after behavioral performance had stabilized suggests the possibility that these NAc representations represent an integrated value estimate that is the outcome of earlier upstream credit assignment in cortical inputs. From a motivated behavior standpoint, the presence of these enriched representations can help inform action selection involving Pavlovian behaviors such as conditioned approach (Ciano et al., 2001; Parkinson et al., 2000; Saunders and Robinson, 2012), and after a decision has been made the maintained representation can help to hold the response in the face of competing alternatives while waiting for an outcome (Di Ciano et al., 2008; Floresco et al., 2008; Floresco, 2015; Peters et al., 2008).

#### 438 **Tiling of task structure:**

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

455 **Implications:**

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

476 Orphan paragraph? (also removed mixed selectivity paragraphs)

477 If these representations function as an eligibility trace, then there should be a relationship between the con-  
478 tent and robustness of these representations and the degree to which organisms form adaptive associations

479 between reinforcers and their environment, which could contribute to the presence of prediction errors in the  
480 NAc in those that learned an experimental task in humans versus those that did not (Schonberg 2007). In  
481 the present experiment, understanding what information about the environment the NAc is keeping track of  
482 when reward arrives could be useful for potentially determining what representations are being reinforced.  
483 Would be interesting to see if selectively reward an animal when certain representations are active (e.g. arm  
484 1), would help shape their assignment of reward to that arm. Could be useful for altering a maladaptive pref-  
485 erence back to a more adaptive one, as the loss of certain aspects of coding via BLA and PFC inputs leads to  
486 the loss of outcome-predictive activity to discrete cues in the NAc. Furthermore, non-discriminatory coding  
487 in the NAc could hypothetically be correlated to over-generalization of situations, in which the animal may  
488 not recognize it is in a different situation and perform actions that are inefficient or maladaptive for reward  
489 procurement.

## 490 References

- 491 Akaiishi, R., Kolling, N., Brown, J. W., and Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue  
492 Environment. *Journal of Neuroscience*, 36(4):1096–1112.
- 493 Asaad, W. F., Lauro, P. M., Perge, J. A., and Eskandar, E. N. (2017). Prefrontal Neurons Encode a Solution to the Credit-Assignment  
494 Problem. *The Journal of Neuroscience*, 37(29):6995–7007.
- 495 Atallah, H. E., McCool, A. D., Howe, M. W., and Graybiel, A. M. (2014). Neurons in the ventral striatum exhibit cell-type-specific  
496 representations of outcome during learning. *Neuron*, 82(5):1145–1156.
- 497 Averbeck, B. B. and Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*,  
498 20(4):505–512.
- 499 Berke, J. D., Breck, J. T., and Eichenbaum, H. (2009). Striatal Versus Hippocampal Representations During Win-Stay Maze  
500 Performance. *Journal of Neurophysiology*, 101(3):1575–1587.
- 501 Berridge, K. C. (2012). From prediction error to incentive salience: Mesolimbic computation of reward motivation. *European  
502 Journal of Neuroscience*, 35(7):1124–1143.
- 503 Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., Kashtelyan, V., and Roesch, M. R.  
504 (2013). Separate Populations of Neurons in Ventral Striatum Encode Value and Motivation. *PLoS ONE*, 8(5):e64673.

- 505 Carelli, R. M. (2010). Drug Addiction: Behavioral Neurophysiology. In *Encyclopedia of Neuroscience*, pages 677–682. Elsevier.
- 506 Chang, S. E. and Holland, P. C. (2013). Effects of nucleus accumbens core and shell lesions on autoshaped lever-pressing. *Be-*  
507 *havioural Brain Research*, 256:36–42.
- 508 Chang, S. E., Wheeler, D. S., and Holland, P. C. (2012). Roles of nucleus accumbens and basolateral amygdala in autoshaped lever  
509 pressing. *Neurobiology of Learning and Memory*, 97(4):441–451.
- 510 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).  
511 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5):1106–  
512 1118.
- 513 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., and Wightman, R. M. (2007). Coordinated Accumbal  
514 Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2):237–244.
- 515 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.,  
516 and Everitt, B. J. (2001). Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens  
517 core in the acquisition and performance of pavlovian approach behavior. *The Journal of neuroscience*, 21(23):9471–9477.
- 518 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., and Schoenbaum, G. (2015).  
519 Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature Communi-*  
520 *cations*, 6:7195.
- 521 Corbit, L. H. and Balleine, B. W. (2011). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are  
522 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33):11786–11794.
- 523 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., and Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in*  
524 *Neurobiology*, 92:345–369.
- 525 Day, J. J., Roitman, M. F., Wightman, R. M., and Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine  
526 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8):1020–1028.
- 527 Day, J. J., Wheeler, R. A., Roitman, M. F., and Carelli, R. M. (2006). Nucleus accumbens neurons encode Pavlovian approach  
528 behaviors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5):1341–1351.
- 529 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., Boraud, T., and Le Moine, C. (2017). Memories of Opiate  
530 Withdrawal Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychopharmacol-*  
531 *ogy*, 42(5):1157–1168.
- 532 Di Ciano, P., Robbins, T. W., and Everitt, B. J. (2008). Differential effects of nucleus accumbens core, shell, or dorsal striatal  
533 inactivations on the persistence, reacquisition, or reinstatement of responding for a drug-paired conditioned reinforcer. *Neu-*  
534 *ropsychopharmacology*, 33(6):1413–1425.
- 535 du Hoffmann, J. and Nicola, S. M. (2014). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the  
536 Nucleus Accumbens. *Journal of Neuroscience*, 34(43):14349–14364.
- 537 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental*

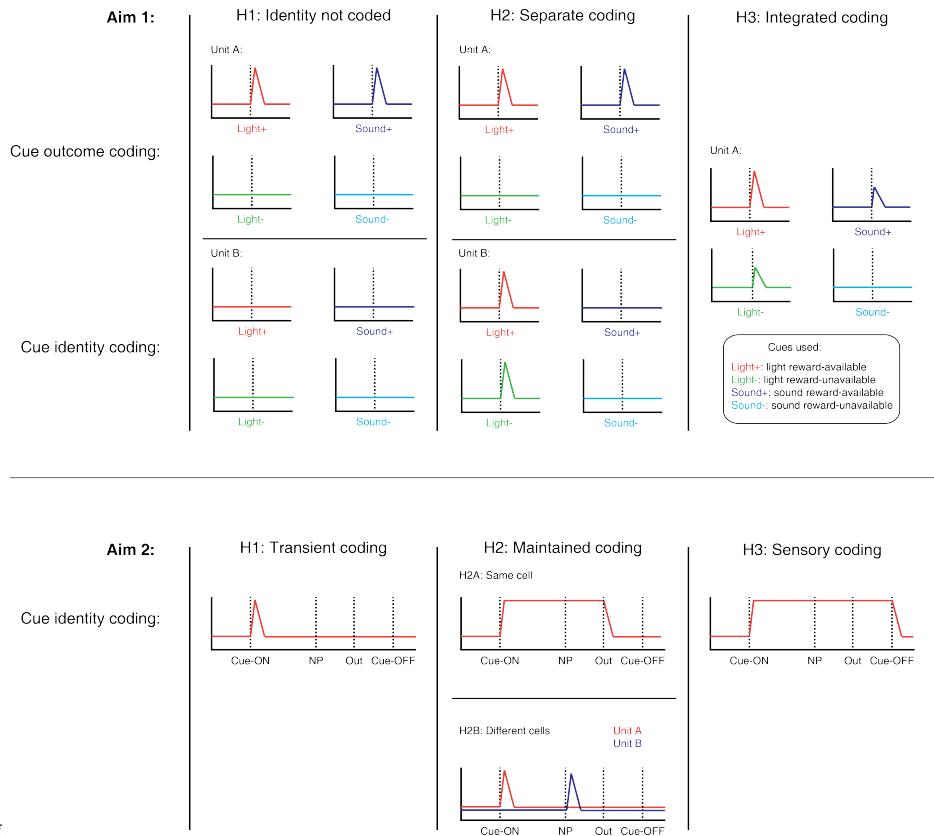
- 538      *Psychology*, 32(2):150–155.
- 539      Fitzgerald, T. H. B., Schwartenbeck, P., and Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent  
540      and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4):1271–1279.
- 541      Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., Akers, C. A., Clinton, S. M., Phillips, P. E. M., and Akil,  
542      H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, 469(7328):53–59.
- 543      Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of  
544      Psychology*, 66(1):25–52.
- 545      Floresco, S. B., McLaughlin, R. J., and Haluk, D. M. (2008). Opposing roles for the nucleus accumbens core and shell in cue-induced  
546      reinstatement of food-seeking behavior. *Neuroscience*, 154(3):877–884.
- 547      Frank, M. J., Seeberger, L. C., and O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism.  
548      *Science*, 306(5703):1940–1943.
- 549      Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., Bryden, D. W., and Roesch, M. R.  
550      (2012). Ventral Striatum Encodes Past and Predicted Value Independent of Motor Contingencies. *Journal of Neuroscience*,  
551      32(6):2027–2036.
- 552      Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., and Steele, J. D. (2011). Expected value  
553      and prediction error abnormalities in depression and schizophrenia. *Brain*, 134(6):1751–1764.
- 554      Grant, D. A. and Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a  
555      Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4):404–411.
- 556      Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J.,  
557      and Berke, J. D. (2015). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1):117–126.
- 558      Hart, A. S., Rutledge, R. B., Glimcher, P. W., and Phillips, P. E. M. (2014). Phasic Dopamine Release in the Rat Nucleus Accumbens  
559      Symmetrically Encodes a Reward Prediction Error Term. *The Journal of Neuroscience*, 34(3):698–704.
- 560      Hauber, W., Bohn, I., and Giertler, C. (2000). NMDA, but not dopamine D(2), receptors in the rat nucleus accumbens are involved  
561      in guidance of instrumental behavior by stimuli predicting reward magnitude. *J Neurosci*, 20(16):6282–6288.
- 562      Hearst, E. and Jenkins, H. M. (1974). *Sign-tracking: the stimulus-reinforcer relation and directed action*. Psychonomic Society.
- 563      Holland, P. C. (1992). Occasion setting in pavlovian conditioning. *Psychology of Learning and Motivation - Advances in Research  
564      and Theory*, 28(C):69–125.
- 565      Hollerman, J. R., Tremblay, L., and Schultz, W. (1998). Influence of Reward Expectation on Behavior-Related Neuronal Activity in  
566      Primate Striatum. *Journal of Neurophysiology*, 80(2):947–963.
- 567      Honey, R. C., Iordanova, M. D., and Good, M. (2014). Associative structures in animal learning: Dissociating elemental and  
568      configural processes. *Neurobiology of Learning and Memory*, 108:96–103.
- 569      Hyman, S. E., Malenka, R. C., and Nestler, E. J. (2006). NEURAL MECHANISMS OF ADDICTION: The Role of Reward-Related

- 570 Learning and Memory. *Annual Review of Neuroscience*, 29(1):565–598.
- 571 Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-  
572 olfactory tubercle complex. *Brain Research Reviews*, 56(1):27–78.
- 573 Ishikawa, A., Ambroggi, F., Nicola, S. M., and Fields, H. L. (2008). Dorsomedial Prefrontal Cortex Contribution to Behavioral and  
574 Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, 28(19):5088–5098.
- 575 Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives.  
576 *Neural Networks*, 15(4-6):535–547.
- 577 Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., Sponheim, S. R., and Lissek,  
578 S. (2017). Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174(2):125–134.
- 579 Kalivas, P. W. and Volkow, N. D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal  
580 of Psychiatry*, 162(8):1403–1413.
- 581 Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in  
582 schizophrenia. *American Journal of Psychiatry*, 160(1):13–23.
- 583 Khamassi, M. and Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and  
584 model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, 6:79.
- 585 Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., and Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal  
586 neurons of behaving rats. *European Journal of Neuroscience*, 28(9):1849–1866.
- 587 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., and Pennartz, C. M. A. (2008).  
588 Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, 28(25):6372–  
589 6382.
- 590 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., and Pennartz, C. M. (2009). Hippocampus leads ventral  
591 striatum in replay of place-reward information. *PLoS Biology*, 7(8):e1000173.
- 592 Lansink, C. S., Jackson, J. C., Lankelma, J. V., Ito, R., Robbins, T. W., Everitt, B. J., and Pennartz, C. M. A. (2012). Reward  
593 Cues in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles. *Journal of  
594 Neuroscience*, 32(36):12444–12459.
- 595 Lansink, C. S., Meijer, G. T., Lankelma, J. V., Vinck, M. A., Jackson, J. C., and Pennartz, C. M. A. (2016). Reward Expectancy  
596 Strengthens CA1 Theta and Beta Band Synchronization and Hippocampal-Ventral Striatal Coupling. *Journal of Neuroscience*,  
597 36(41):10598–10610.
- 598 Lavoie, A. M. and Mizumori, S. J. (1994). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons  
599 of rats. *Brain Research*, 638(1-2):157–168.
- 600 Lee, D., Seo, H., and Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of  
601 Neuroscience*, 35(1):287–308.

- 602 Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., and Grillon, C. (2014). Generalized anxiety disorder is associated  
603 with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75(11):909–915.
- 604 Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and*  
605 *Behavioral Neuroscience*, 9(4):343–364.
- 606 Maia, T. V. and Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature*  
607 *Neuroscience*, 14(2):154–162.
- 608 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., and Schoenbaum, G. (2011). Ventral Striatum and Orbitofrontal Cortex  
609 Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, 31(7):2700–2705.
- 610 McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., and Nicola, S. M. (2013). Invigoration of reward seeking by cue and proximity  
611 encoding in the nucleus accumbens. *Neuron*, 78(5):910–922.
- 612 Mulder, A. B., Shibata, R., Trullier, O., and Wiener, S. I. (2005). Spatially selective reward site responses in tonically active neurons  
613 of the nucleus accumbens in behaving rats. *Experimental Brain Research*, 163(1):32–43.
- 614 Mulder, A. B., Tabuchi, E., and Wiener, S. I. (2004). Neurons in hippocampal afferent zones of rat striatum parse routes into  
615 multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7):1923–1932.
- 616 Nicola, S. M. (2004). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a Discrimi-  
617 native Stimulus Task. *Journal of Neurophysiology*, 91(4):1840–1865.
- 618 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of  
619 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49):16585–16600.
- 620 Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor.  
621 *Psychopharmacology*, 191(3):507–520.
- 622 Noonan, M. P., Chau, B. K., Rushworth, M. F., and Fellows, L. K. (2017). Contrasting Effects of Medial and Lateral Orbitofrontal  
623 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29):7023–7035.
- 624 Parkinson, J. A., Willoughby, P. J., Robbins, T. W., and Everitt, B. J. (2000). Disconnection of the anterior cingulate cortex  
625 and nucleus accumbens core impairs pavlovian approach behavior: Further evidence for limbic cortical-ventral striatopallidal  
626 systems. *Behavioral Neuroscience*, 114(1):42–63.
- 627 Pennartz, C. M. A. (2004). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation by  
628 Hippocampal Ripples. *Journal of Neuroscience*, 24(29):6446–6456.
- 629 Peters, J., LaLumiere, R. T., and Kalivas, P. W. (2008). Infralimbic Prefrontal Cortex Is Responsible for Inhibiting Cocaine Seeking  
630 in Extinguished Rats. *Journal of Neuroscience*, 28(23):6046–6053.
- 631 Rescorla, R. A. and Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and  
632 Instrumental Learning. *Psychological Review*, 74(3):151–182.
- 633 Robinson, T. E. and Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues

- 634 Through the Study of Individual Differences. *Biological Psychiatry*, 65(10):869–873.
- 635 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., and Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the  
636 Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42):13365–13376.
- 637 Roitman, M. F., Wheeler, R. A., and Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for rewarding and  
638 aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4):587–597.
- 639 Saddoris, M. P., Stamatakis, A., and Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nu-  
640 cleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*,  
641 33(12):2274–2287.
- 642 Salamone, J. D. and Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3):470–485.
- 643 Saunders, B. T. and Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of pavlovian-conditioned  
644 responses. *European Journal of Neuroscience*, 36(4):2521–2532.
- 645 Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1):23–32.
- 646 Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.
- 647 Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning.  
648 *Neuron*, 38(4):625–636.
- 649 Shidara, M., Aigner, T. G., and Richmond, B. J. (1998). Neuronal signals in the monkey ventral striatum related to progress through  
650 a predictable series of trials. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18(7):2613–25.
- 651 Shmuel Lissek, A. N. K. (2013). Generalization of Conditioned Fear and Obsessive-Compulsive Traits. *Journal of Psychology &*  
652 *Psychotherapy*, 7:3.
- 653 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., and Buzsáki, G. (2017). Cocaine place conditioning strengthens location-  
654 specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, pages 1–10.
- 655 Sleezer, B. J., Castagno, M. D., and Hayden, B. Y. (2016). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.  
656 *Journal of Neuroscience*, 36(44):11223–11237.
- 657 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., and Hayden, B. Y. (2016). Neuronal selectivity for spatial  
658 positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3):1098–1111.
- 659 Sugam, J. A., Saddoris, M. P., and Carelli, R. M. (2014). Nucleus accumbens neurons track behavioral preferences and reward  
660 outcomes during risky decision making. *Biological Psychiatry*, 75(10):807–816.
- 661 Sutton, R. and Barto, A. (1998). Reinforcement Learning: An Introduction.
- 662 Tabuchi, E. T., Mulder, A. B., and Wiener, S. I. (2000). Position and behavioral modulation of synchronization of hippocampal and  
663 accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6):717–728.
- 664 Takahashi, Y. K., Langdon, A. J., Niv, Y., and Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled  
665 by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1):182–193.

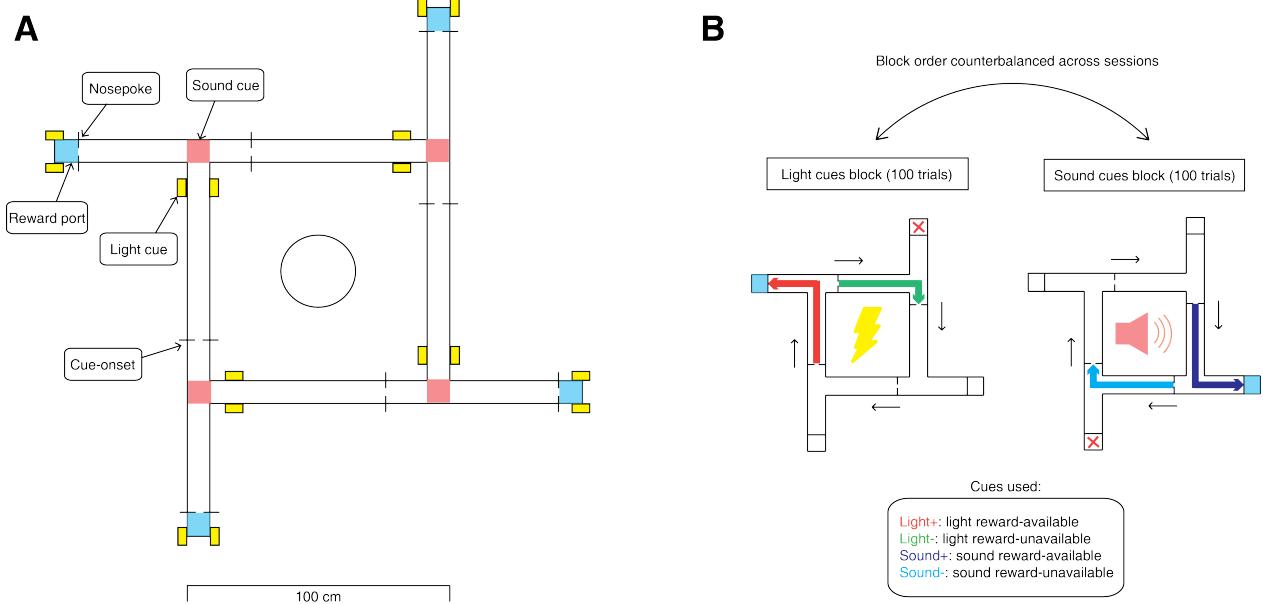
- 666 van der Meer, M. A. A. and Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward  
667 Information. *Journal of Neuroscience*, 31(8):2843–2854.
- 668 West, E. A. and Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after  
669 Reinforcer Devaluation. *Journal of Neuroscience*, 36(4):1128–1139.
- 670 Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., and Mulder, A. B. (2003). Spatial and behavioral correlates  
671 in nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,  
672 1250(C):275–292.



1 - Schematic neural.png

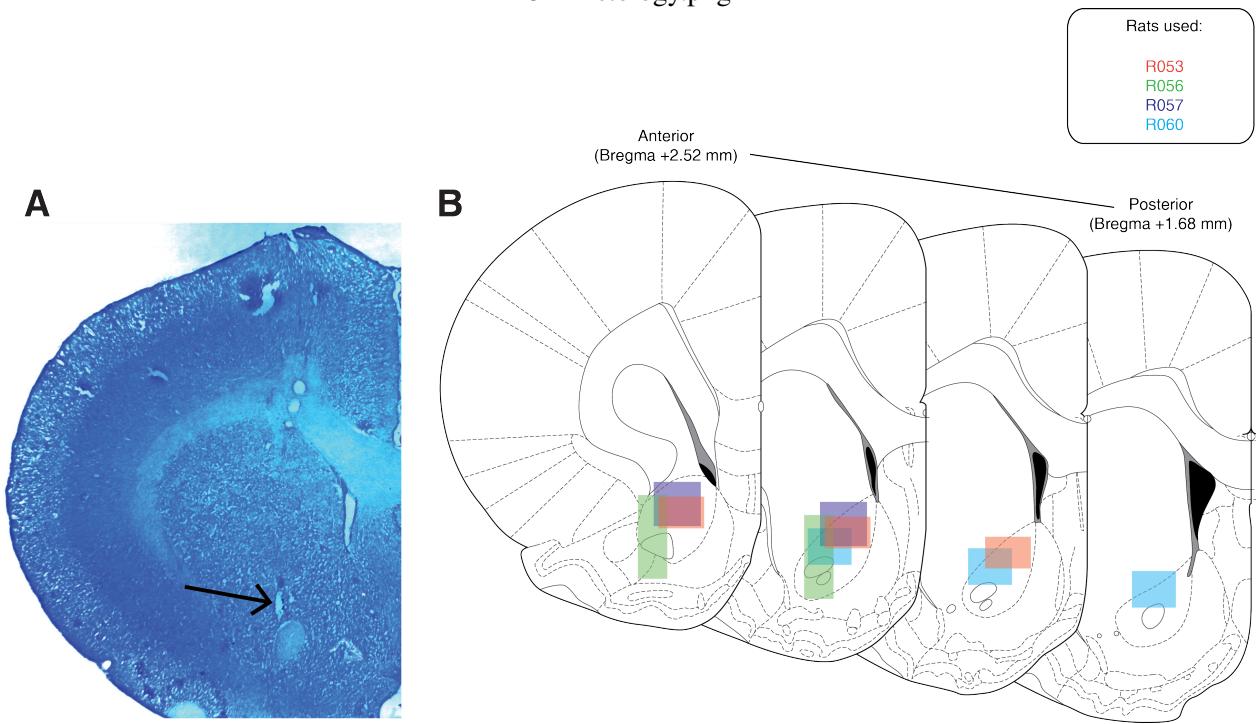
**Figure 1:** Schematic of potential coding strategies for cue identity (light, sound) and cue outcome (reward-available, reward-unavailable) employed by single units in the NAc. **A:** Displayed are schematic PETHs illustrating putative responses to different cues under different hypotheses of how cue identity and value 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 reward-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, 'Cue-OFF' signifies when the cue turns off. H2 (middle panel): Coding of cue identity is maintained 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 is maintained during entire duration of cue, including after outcome is revealed. 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



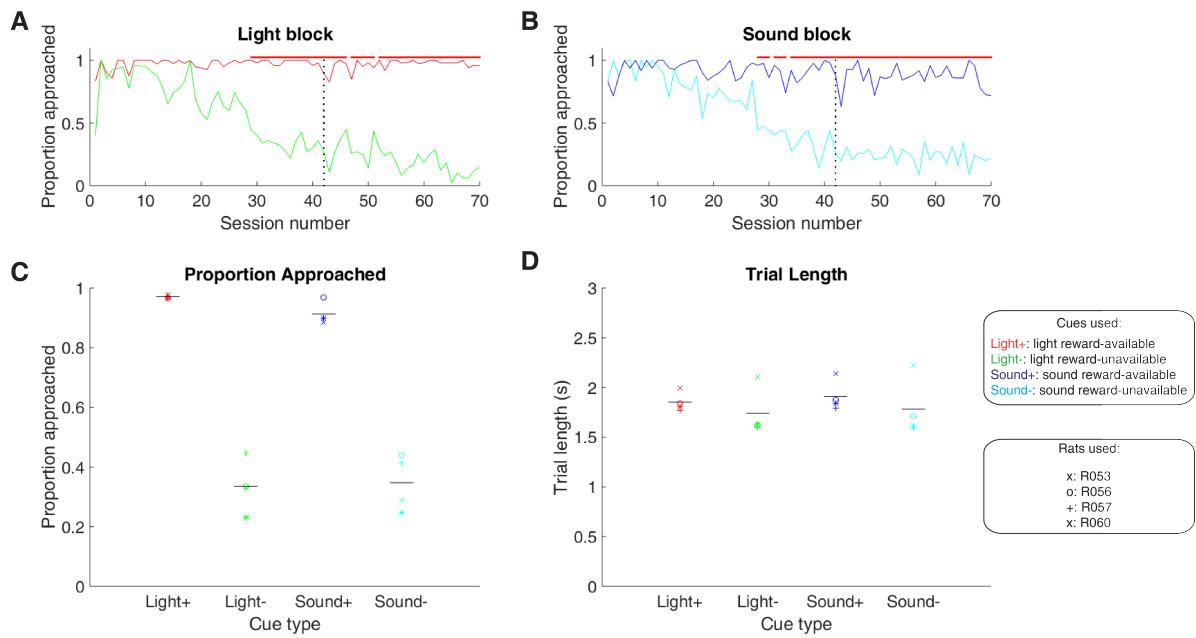
**Figure 2:** Schematic of behavioral task. **A:** To scale depiction of square track consisting of multiple identical T-choice points. At each choice point, the availability of sucrose reward at the nearest reward receptacle (light blue fill) was signaled by one of four possible cues, presented when 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. 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. Light blue fill on ends of arms indicate location of receptacles where sucrose was dispensed on rewarded trials. 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 two blocks of a cue discrimination task on a square track where they had to dissociate between a reward-available and reward-unavailable cue for a set of light cues and a set of sound cues for a target of 100 trials each. Rats ran clockwise on the square track, and upon presentation of the cue could either turn left at the choice point to check if reward was available at the reward receptacle (approach trial; red and navy blue in figure), or turn right at the choice point to initiate the following trial (skip trial; green and light blue in figure). 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. Left in figure depicts a light block, showing an example trajectory for a correct reward-available (red) and reward-unavailable (green) trial. Right in figure depicts a sound block, with a reward-available (navy blue) and reward-unavailable (light blue) trial. Ordering of the light and sound blocks was counterbalanced across sessions. Dashed line within track indicates photobeam that triggered cue onset. Solid line within track indicates photobeam that registered a nosepoke into the reward receptacles. Receptacles containing sucrose solution are present within the squares at the end of each arm. Light blue fill signifies reward receipt on a correct reward-available trial. Arrows outside of track indicate correct running direction. A session ended with another 5 minute recording period on the pedestal.

3 - Histology.png



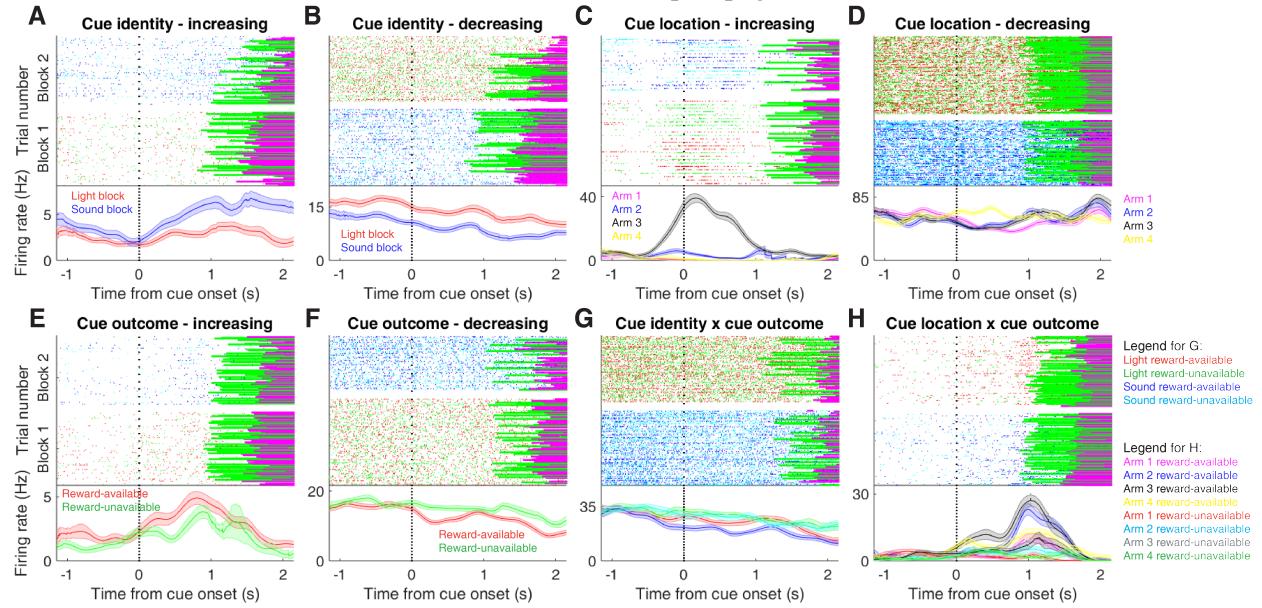
**Figure 3:** Histological verification of recording sites. Upon completion of experiments, brains were sectioned and tetrode placement was confirmed. **A:** Example section from R060 showing a recording site in the NAc core just dorsal to the anterior commissure (white circle). **B:** Schematic showing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

4 - Behavioral results.png



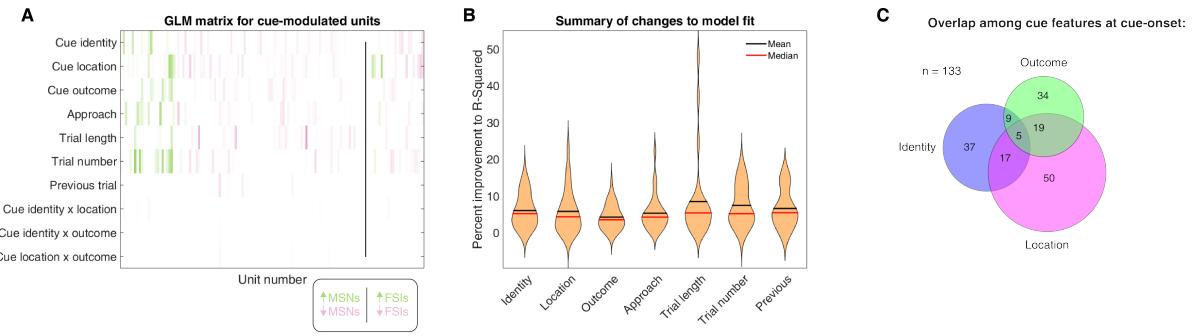
**Figure 4:** Performance on the behavioral task. **A-B.** Example learning curves across sessions from a single subject (R060) showing the proportion approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light block, light blue line for sound block). Fully correct performance corresponds to 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. **C-D:** Summary of performance during recording sessions for each rat. **C:** 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). **D:** Average trial length for each cue. Note that the time to complete a trial was comparable for the different cues.

5 - Neural examples.png



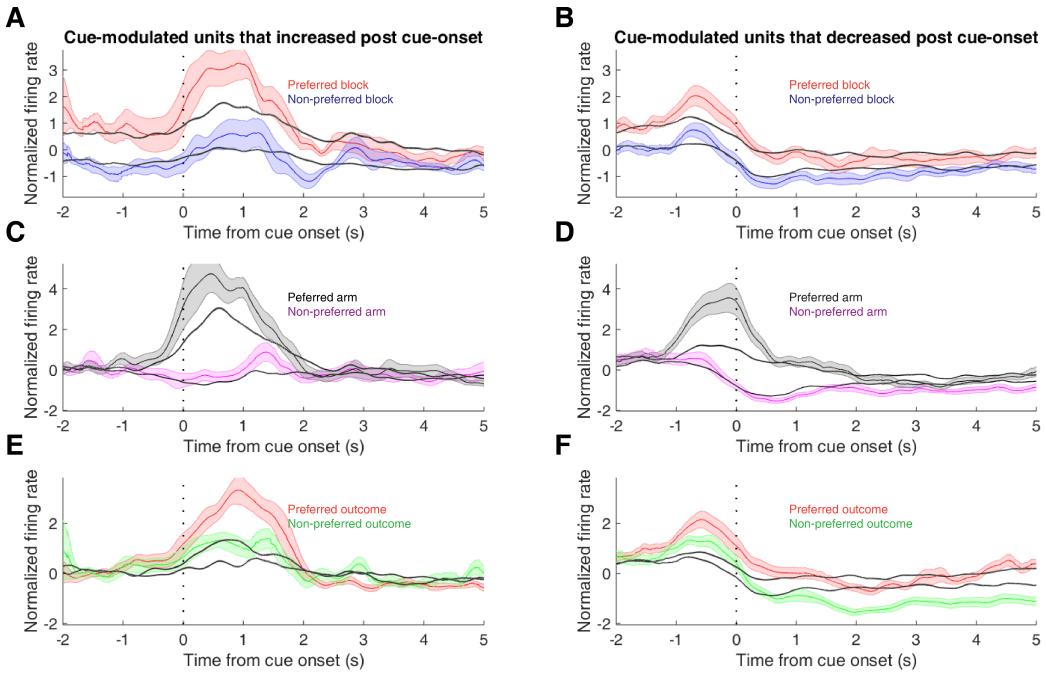
**Figure 5:** Examples of different cue-modulated NAc units influenced by various task parameters. **A:** Example of a cue-modulated NAc unit that showed an increase in firing following the cue, and encoded cue identity. Top: rasterplot showing the spiking activity across all trials aligned to cue-onset. 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

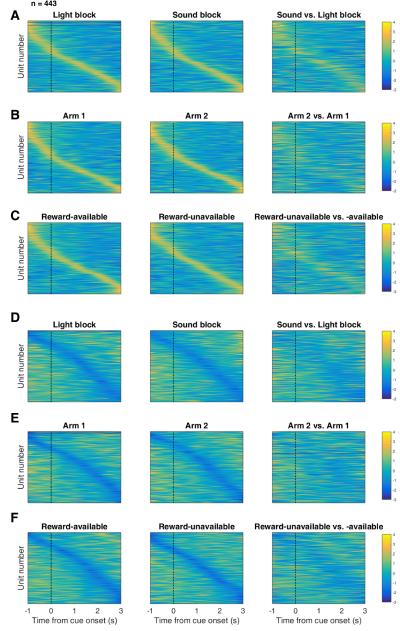


**Figure 6:** Summary of influence of various task parameters on cue-modulated NAc units after cue-onset. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank 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. **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.

## 7 - Population averages.png



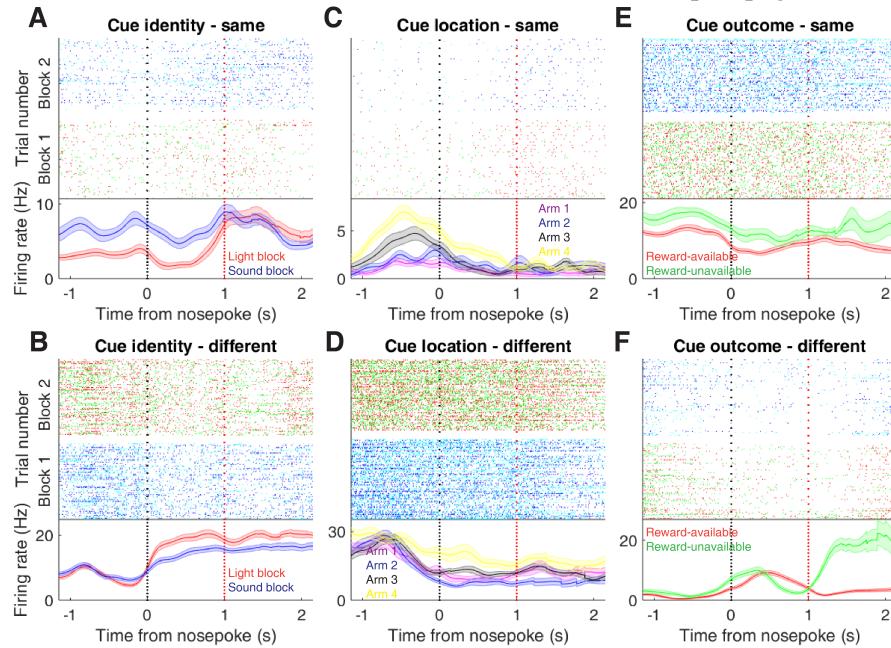
**Figure 7:** Population-level averages of cue feature sensitive NAc units. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to cue-onset. Activity is plotted for preferred stimulus block (red) and non-preferred stimulus block (blue). Dashed line indicates onset of cue. Lightly shaded area indicates standard error of the mean. Note larger increase to preferred stimulus block to 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 cells 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 from most preferred arm (yellow), in decreasing order to least preferred arm (black, navy blue, magenta, respectively). Note the graded response to arms of decreasing preference. **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 cells preferred outcome (E), and the sustained decrease to the nonpreferred outcome (F).



8 - Task tiling.png

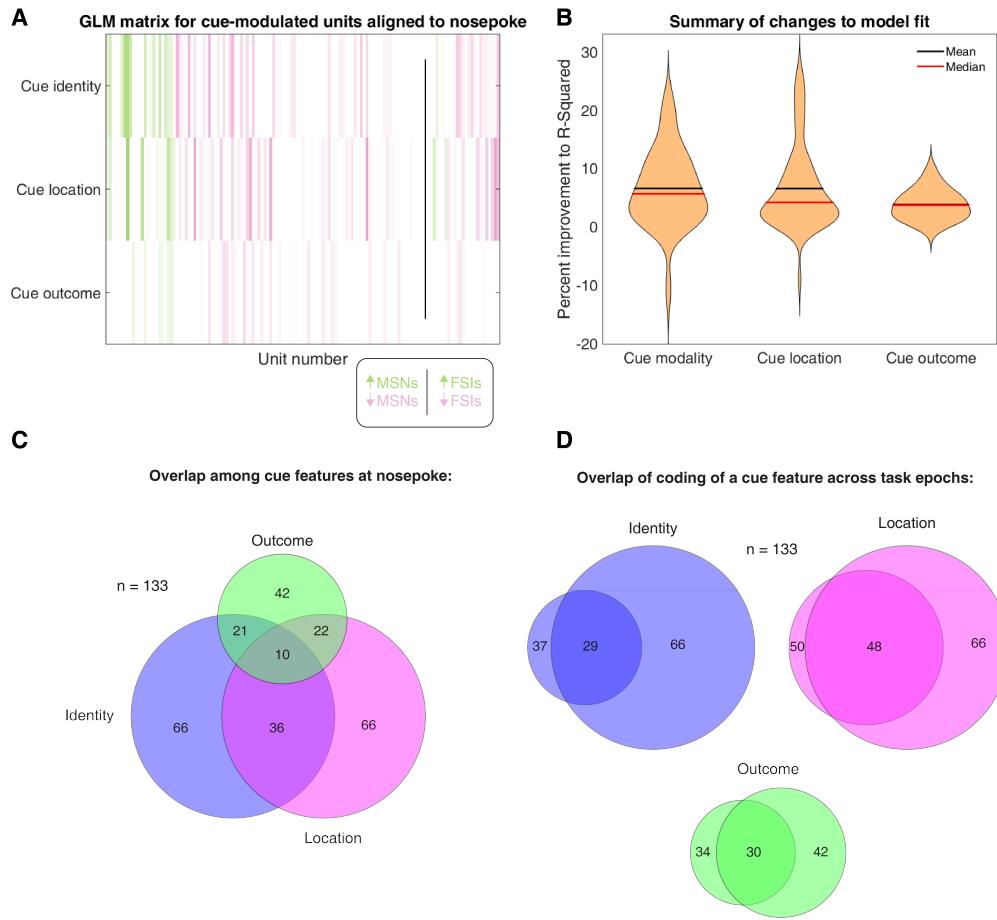
**Figure 8:** Distribution of NAc firing rates across time surrounding cue onset. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue onset), averaged across all trials for a specific cue type, indicated by text labels. **A-C:** Heat plots aligned to normalized peak firing rates. A, 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. 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 left: 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, middle 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, far right: Unit firing rates taken from half of the light block, ordered according to peak firing rate taken from the other half of light trials. Note similar pattern to middle right suggests 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



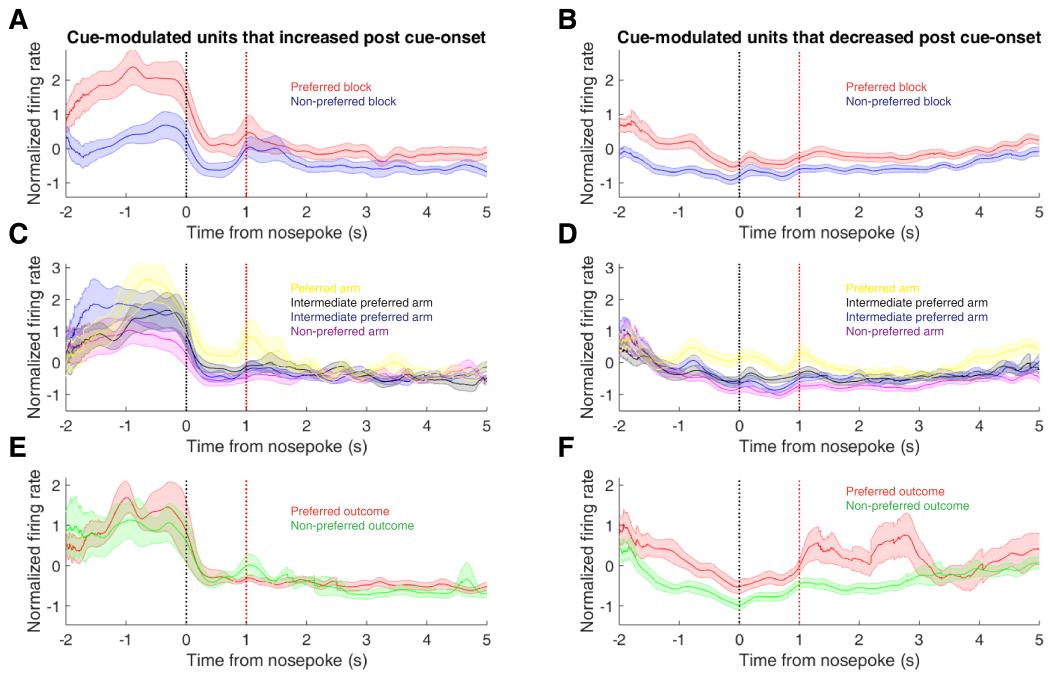
**Figure 9:** Examples of cue-modulated NAc units influenced by various task parameters at time of nosepoke. **A:** Example of a cue-modulated NAc unit that encoded cue identity at both cue-onset and during nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to nosepoke. 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 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

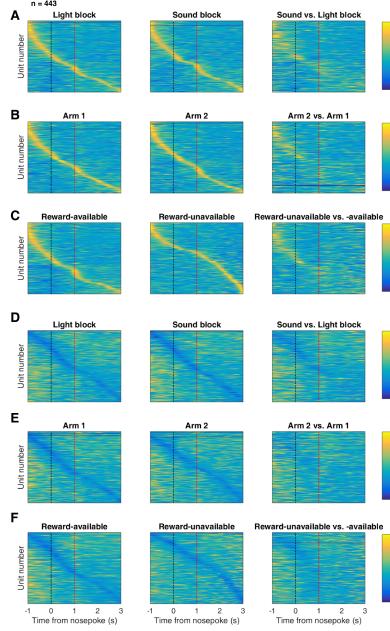


**Figure 10:** Summary of influence of various task parameters of cue-modulated NAc units during nosepoke. **A:** GLM matrix demonstrating impact of various task parameters on NAc firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank test. Each row represents a given task parameter, and each column is the influence of that task parameter on a given unit. Response variable is 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.

## 11 - NP population averages.png



**Figure 11:** Population-level averages of cue feature sensitive NAc units during a nosepoke. **A:** Average smoothed normalized 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 dashed line indicates nosepoke. Red dashed line indicates reward delivery occurring 1 s after nosepoke for reward-available trials. Lightly shaded area indicates standard error of the mean. Note larger increase leading up to nosepoke to preferred stimulus block to 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 from most preferred arm (yellow), in decreasing order to least preferred arm (black, navy blue, magenta, respectively). Note the subtle but graded response to arms of decreasing preference. **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 reward receipt for preferred outcome in decreasing cells (F).



12 - NP task tiling.png

**Figure 12:** Distribution of NAc firing rates across task space during approach trials. Each panel shows normalized 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**, 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 left: 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 reward receipt. A, middle 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, far right: Unit firing rates taken from half of the light block, ordered according to peak firing rate taken from the other half of light trials. Note similar pattern to middle right suggests 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, but with units ordered according to the time of their minimum firing rate. **C**: 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. **D**: As for C, but with units ordered by their minimum firing rate. **E**: Same layout as in A, except that panels now compare reward-available and reward-unavailable trials. Notice the absence of a cluster of tiling around reward receipt for reward-unavailable trials. **F**: As for A, 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.

Task parameter	Total	↑ MSN	↓ MSN	↑ FSI	↓ FSI
All units	443	155	216	27	45
<b>Rat ID</b>					
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	?	?	?	?
Cue modulated units	133	24	85	6	18
<b>GLM aligned to cue-onset</b>					
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
Recent trial history	5	0	5	0	0
<b>GLM aligned to nosepoke</b>					
Cue identity	66	14	36	2	14
Cue location	66	14	40	3	9
Cue outcome	42	8	29	0	5
<b>GLM aligned to outcome</b>					
Cue outcome	10	0	6	0	4

**Table 1:** Cells overview