

Maintained coding of environmental features alongside outcome-predictive signals in the NAc (old: Coding of behaviorally relevant and irrelevant cue features in the nucleus accumbens)

Authors: Jimmie M. Gmaz¹, James E. Carmichael¹, Matthijs A. A. van der Meer^{1*}

¹Department 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 to do

3 Significance Statement (120 words)

4 to do

5 Introduction

6 Theories of nucleus accumbens (NAc) function generally agree that this brain structure contributes to moti-
7 vated behavior, with some emphasizing a role in learning from reward prediction errors (Averbeck and Costa,
8 2017; Joel et al., 2002; Khamassi and Humphries, 2012; Lee et al., 2012; Maia, 2009; Schultz, 2016)(see also
9 the addiction literature on effects of drug rewards; Carelli, 2009; Hyman et al., 2006; Kalivas and Volkow,
10 2005) and others a role in the modulation of ongoing behavior through stimuli associated with motivationally
11 relevant outcomes (invigorating, directing; Floresco, 2015; Nicola, 2010; Salamone and Correa, 2012).

12 These proposals echo similar ideas on the functions of the neuromodulator dopamine (Berridge, 2012; Maia,
13 2009; Salamone and Correa, 2012; Schultz, 2016), with which the NAc is tightly linked functionally as well
14 as anatomically (Cheer et al., 2007; du Hoffmann and Nicola, 2014; Ikemoto, 2007; Takahashi et al., 2016).

15 Much of our understanding of NAc function comes from studies of how cues that predict motivationally
16 relevant outcomes (e.g. reward) influence behavior and neural activity in the NAc. Task designs that asso-
17 ciate such cues with rewarding outcomes provide a convenient access point eliciting conditioned responses
18 such as sign-tracking and goal-tracking (Hearst and Jenkins, 1974; Robinson and Flagel, 2009), pavlovian-
19 instrumental transfer (Estes, 1943; Rescorla and Solomon, 1967) and enhanced response vigor (Nicola, 2010;
20 Niv et al., 2007), which tend to be affected by NAc manipulations (Chang et al., 2012; Corbit and Balleine,
21 2011; Flagel et al., 2011) (although not always straightforwardly (is Hauber appropriate?); Chang and Hol-
22 land, 2013; Hauber et al., 2000). Similarly, analysis of reward prediction errors typically proceeds by estab-
23 lishing an association between a cue and subsequent reward, with NAc responses transferring from outcome
24 to the cue with learning (Day et al., 2007; Roitman et al., 2005; Setlow et al., 2003; Schultz et al., 1997)(Day
25 2007 is the only paper referenced here that looks at transferring of responses from reward to CS+, other
26 papers are in VTA, or the development of CS+ responses).

29 how reward-predictive cues themselves are encoded in the NAc (Hayden from primate realm). This is an im-
30 portant issue for at least two reasons. First, in reinforcement learning, motivationally relevant outcomes are
31 typically temporally delayed relative to the cues that predict them. In order to solve the problem of assigning
32 credit (or blame) across such temporal gaps, some trace of preceding activity needs to be maintained (Lee
33 et al., 2012) (Maia?). Since NAc is a primary target of DA signals interpretable as RPEs, and NAc lesions
34 impair RPEs related to timing, its activity trace will help determine what can be learned when RPEs arrive
35 (Takahashi).

36 Second, for ongoing behavior, the relevance of cues typically depends on context. In experimental settings,
37 context may include the identity of a preceding cue (occasion setter, Holland, Kesner), spatial or configural
38 arrangements (Good/Honey, Eichenbaum), and unsignaled rules as occurs in set shifting and other cognitive
39 control tasks (CITE). In such situations, the question arises how selective, context-dependent processing of
40 reward-predictive cues is implemented. For instance, is there a gate prior to NAc, or are all cues represented
41 in NAc but their current values dynamically updated (FitzGerald; **WHAT ARE SOME THEORETICAL**
42 **POSITIONS ON THIS?**)

43 Thus, both from a learning and a flexible performance perspective, it is of interest to determine how cue iden-
44 tity is represented in the brain, with NAc of particular interest given its anatomical and functional position
45 at the center of motivational systems. We sought to determine whether cue features signalling identity are
46 represented in the NAc, and if cue identity is represented alongside other motivationally relevant variables,
47 such as cue value (Figure 1). To address this question, we designed an experiment in which multiple, distinct
48 sets of cues predicted the same outcome. We recorded the activity of NAc neurons as rats performed this
49 task.

50 [Figure 1 about here.]

51 **Methods**

52 **Subjects:**

53 Adult male Long-Evans rats ($n = 4$, Charles River, Saint Constant, QC) were used as subjects. Rats were in-
54 dividually housed with a 12/12-h light-dark cycle, and tested during the light cycle. Rats were food deprived
55 to 85-90% of their free feeding weight (weight at time of implantation was 440 - 470 g), and water restricted
56 4-6 hours before testing. All experimental procedures were approved by the the University of Waterloo An-
57 imal Care Committee (protocol# 11-06) and carried out in accordance with Canadian Council for Animal
58 Care (CCAC) guidelines.

59 **Overall timeline:**

60 Each subject was first handled for seven days during which they were exposed to the running room, the
61 sucrose solution used as a reinforcer, and the click of the valves upon approach to the receptacles. They were
62 then shaped to run on the task for seven days where they were restricted to running in the clockwise direction
63 by presenting a physical barrier to running counterclockwise. Rats then continued to be trained each day for
64 100 trials in both a light and sound block until approach behavior discriminated between reward-available
65 and reward-unavailable cues for three consecutive days according to a chi square test, at which point they
66 underwent hyperdrive implantation targeted at the NAc. Rats were allowed to recover for a minimum of five
67 days before being retrained on the task, and recording began once performance returned to pre-surgery levels.
68 Upon completion of recording, animals were sacrificed and recording sites were histologically confirmed.

69 **Behavioral task and training:**

70 Rats were trained to run clockwise on an elevated, square-shaped track (100x100 cm) containing four possi-

ble reward locations (Figure 2). Rats initiated a trial by running down the length of an arm, and triggering a photobeam located 24 cm from the start of each arm. Upon trial initiation, a light or sound cue was presented that signaled the presence (reward-available trial) or absence (reward-unavailable trial) of a 12% sucrose water reward (0.1 mL) at the upcoming site. A trial was classified as an approach trial if the rat turned left at the decision point and made a nosepoke at the reward receptacle (40 cm from the decision point), while trials were classified as a skip trial if the rat instead turned right at the decision point and triggered the photobeam to initiate the following trial. In reward-available trials there was a 1 second delay between a nosepoke and subsequent reward delivery. Trial length was determined by measuring the length of time from cue onset until nosepoke (for approach trials), or from cue onset until the start of the following trial (for skip trials). Trials could only be initiated through clockwise progression through the series of arms, and each entry into the subsequent arm on the track counted as a trial. To get enough trials for neural analysis rats were trained each day in both a light and sound block for 100 trials each. Within a block, one cue signaled reward was available on that trial, while the other signaled reward was not available. Light block cues were a flashing white light, and a constant yellow light. Sound block cues were a 2 kHz sine wave and a 8 kHz sine wave whose amplitude was modulated from 0 to maximum by a 2 Hz sine wave. Reward-cue associations were counterbalanced across rats. Cue presentation was pseudorandomized so that the same cue could not be presented more than twice in a row. Block order within each day was also pseudorandomized, such that the rat could begin a session with the same block for more than two days in a row. Each training or testing day consisted of a 5 minute pre-session period on a pedestal, followed by the first block, then the second block, then a 5 minute post-session period on the pedestal. Performance was determined by the proportion of trials a rat approached each cue. Perfect performance would be 100% approach on reward-available trials, and 0% approach on reward-unavailable trials. Rats were trained daily until they could distinguish between the reward-available and reward-unavailable cues for both light and sound blocks for three consecutive days, according to a chi-square test rejecting the null hypothesis of equal approaches for reward-available and reward-unavailable trials, at which point they underwent surgery.

96

[Figure 2 about here.]

97 **Surgery:**

98 Surgical procedures were as described previously (Malhotra et al., 2015). Briefly, animals were anesthetized
99 with isoflurane, induced with 5% in medical grade oxygen and maintained at 2% throughout the surgery
100 (0.8 L/min). Rats were then chronically implanted with a hyperdrive consisting of 16 independently drivable
101 tetrodes, either all 16 targeted for the right NAc (AP +1.4 mm and ML +1.6 mm, relative to bregma; Paxinos
102 and Watson, 2005), or 12 in the right NAc and 4 targeted at the mPFC (AP +3.0 mm and ML +0.6 mm,
103 relative to bregma; only data from NAc tetrodes were analyzed). Following surgery, all animals were given
104 at least five days to recover and lower tetrodes to the target (DV -6.0 mm) before being reintroduced to the
105 behavioral task.

106 **Data acquisition and preprocessing:**

107 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with
108 a RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex).
109 Signals were referenced against a tetrode placed in the corpus callosum above the NAc.

110 Candidate spikes for sorting into putative single units were obtained by band-pass filtering the data between
111 600-9000 Hz, thresholding and aligning the peaks (UltraMegaSort2k, Hull et al., 2011). Spike waveforms
112 were then clustered with KlustaKwik using energy and the first derivative of energy as features, and manually
113 sorted into units (MClust 3.5, A.D. Redish et al.). Isolated units containing a minimum of 200 spikes within
114 a session were included for subsequent analysis. Units were classified as fast spiking interneurons (FSIs) by
115 an absence of interspike intervals (ISIs) > 2 s, while medium spiny neurons (MSNs) had a combination of
116 ISIs > 2 s and phasic activity with shorter ISIs (Barnes 2005, Atallah 2014).

117 **Data analysis:**

118 For behavioral performance we generated linear mixed effects models to investigate the relationships be-
119 tween cue type and our behavioral variables, with cue type as a fixed effect, and the addition of an intercept
120 for rat identity as a random effect. Average proportion of trials approached and trial length for a session
121 were used as response variables. Contribution of cue type to behavior was determined by comparing the full
122 model to a model with cue type removed for each behavioral variable.

123 To investigate the contribution of various cue features on NAc firing rates we first determined whether firing
124 rates for a unit were modulated by the onset of a cue by collapsing across all cues and comparing the firing
125 rates for the 1 s preceding cue-onset with the 1 s following cue-onset. Single units were considered to be
126 cue-responsive if a Wilcoxon signed-rank test comparing pre- and post-cue firing had a $p < .01$. (Excluded:
127 the mean firing rate difference between pre- and post-cue onset was within the lower or upper 2.5% of a
128 shuffled distribution, as there is redundancy using both, and I did both because I was paranoid about just
129 using one.). Cue-modulated units were then classified as either increasing or decreasing in response to the
130 cue if the post-cue activity was higher or lower than the pre-cue activity, respectively.

131 To determine the relative contribution of various task parameters to firing rate variance for units whose firing
132 was modulated by cue-onset (as in Figures 5,6), a forward selection stepwise general linear model (GLM)
133 was fit to each cue-responsive unit. Cue identity (light block, sound block), cue location (arm 1, arm 2, arm
134 3, arm 4), cue outcome (reward-available, reward-unavailable), approach behavior, trial length, trial number,
135 and trial history were used as predictors, and the 1 s post-cue firing rate as the response variable. Units were
136 classified as being modulated by a given task parameter if addition of the parameter significantly improved
137 model fit using deviance as the criterion ($p < .01$). A comparison of the R-squared value between the final
138 model and the final model minus the predictor of interest was used to determine the amount of firing rate
139 variance explained by the addition of that predictor for a given unit.

140 To better visualize responses to cues and enable subsequent population level analyses (as in Figures 5,7,8),
141 spike trains were convolved with a Gaussian kernel ($\sigma = 100$ ms), and peri-event time histograms (PETHs)

142 were generated by taking the average of the convolved spike trains across trials for a given task condition.
143 For analysis of population-level responses for cue features (Figure 7), convolved spike trains for all units
144 where cue identity, cue location, or cue outcome explained a significant portion of firing rate variance were
145 z-scored. Within a given cue feature, normalized spike trains were then separated according to the preferred
146 and non-preferred cue condition (e.g. light vs sound block), and averaged across units to generate population-
147 level averages. To account for separation that would result from any random selection of units, cell identity
148 was shuffled and the shuffled average for preferrend and non-preferred cue conditions was generated for
149 1000 shuffles.

150 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units
151 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure
152 8). To compare representations of task space across cue conditions for a cue feature, the ordering of units
153 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other
154 condition (e.g. sound block). For control comparisons within cue conditions, half of the trials for a condition
155 were compared against the other half. To look at the correlation of firing rates of all cells within and across
156 various cue conditions, trials for each cue condition for a cell were shuffled and divided into two averages,
157 and averages within and across cue conditions were correlated. A linear mixed effects model was run for
158 each cue condition to determine if correlations of firing rates within cue conditions were more similar than
159 correlations across cue conditions.

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

163 Given that some of our analyses compare firing rates across time, particularly comparisons across blocks,
164 we sought to exclude units with unstable firing rates that would generate spurious results reflecting a drift
165 in firing rate over time unrelated to our task. To do this we ran a Mann-Whitney U test comparing the cue-

¹⁶⁶ evoked firing rates for the first and second half of trials within a block, and excluded 99 units from analysis
¹⁶⁷ that showed a significant change for either block. All analyses were completed in MATLAB R2015a, the
¹⁶⁸ code is available on our public GitHub repository (<http://github.com/vandermeerlab/papers>), and the data
¹⁶⁹ can be accessed through DataLad.

¹⁷⁰ **Histology:**

¹⁷¹ Upon completion of the experiment, rats were anesthetized with 5% isoflurane, then asphyxiated with carbon
¹⁷² dioxide. Transcardial perfusions were performed, and brains were fixed and removed. Brains were sliced
¹⁷³ in 50 μ m coronal sections and stained with thionin. Slices were visualized under light microscopy, tetrode
¹⁷⁴ placement was determined, and electrodes with recording locations in the NAc were analyzed (Figure 3).

¹⁷⁵ [Figure 3 about here.]

¹⁷⁶ **Results**

¹⁷⁷ **Behavior**

¹⁷⁸ Rats were trained to discriminate between cues signaling the availability and absence of reward on a square
¹⁷⁹ track with four identical arms for two distinct sets of cues. During each session, rats were presented se-
¹⁸⁰ quentially with two behavioral blocks containing cues from different sensory modalities, a light and a sound
¹⁸¹ block, with each block containing a cue that signalled the availability of reward (reward-available), and a
¹⁸² cue that signalled the absence of reward (reward-unavailable). An example learning curve is seen in Figure
¹⁸³ 4A,B. All four rats learned to discriminate between the reward-available and reward-unavailable cues for
¹⁸⁴ both the light and sound blocks as determined by reaching significance ($p < .05$) on a daily chi-square test

185 comparing approach behavior for reward-available and reward-unavailable cues for each block, for at least
186 three consecutive days (range for time to criterion: 22 - 57 days). Maintenance of behavioral performance
187 during recording sessions was assessed using linear mixed effects models for both proportion of trials where
188 the rat approached the receptacle, and trial length. Analyses revealed that the likelihood of a rat to make an
189 approach was influenced by which of the cues was presented (Percentage approached: light reward-available
190 = 97%; light reward-unavailable = 34%; sound reward-available = 91%; sound reward-unavailable 35%; p
191 < .001), but that cue identity did not significantly affect the length of time taken to complete a trial (Trial
192 length: light reward-available = 1.85 s; light reward-unavailable = 1.74 s; sound reward-available = 1.91 s;
193 sound reward-unavailable 1.78 s; p = .13)(Figure 4C,D). Thus, during recording, rats successfully discrim-
194 inated the various cues according to whether or not they signalled the availability of reward at the reward
195 receptacle.

196 [Figure 4 about here.]

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

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

199 We sought to address which parameters of our task were encoded by NAc activity, specifically whether the
200 NAc encodes aspects of motivationally relevant cues not directly tied to reward, such as the identity and
201 location of the cue, and whether this coding is independent or integrated with coding of cue outcome. To do
202 this we recorded a total of 443 units with > 200 spikes in the NAc from 4 rats over 57 sessions while they
203 performed a cue discrimination task (Table 1). Units that exhibited a drift in firing rate over the course of
204 either block were excluded from further analysis, leaving 344 units for further analysis. The activity of 133
205 (39%) of these 344 units were modulated by the cue, with more showing a decrease in firing (n = 103) than
206 an increase (n = 30) around the time of cue-onset (Table 1). Within this group, 24 were classified as FSIs,

207 while 109 were classified as SPNs. Upon visual inspection, we observed several patterns of firing activity,
208 including units that; discriminated firing upon cue-onset across various cue conditions, showed sustained
209 differences in firing across cue conditions, had transient responses to the cue, showed a ramping of activity
210 starting at cue-onset, and showed elevated activity immediately preceding cue-onset, for example (Figure 5).
211 To characterize more formally whether these cue-evoked responses were modulated by various aspects of
212 the task, we fit a GLM to each cue-modulated unit. Fitting GLMs revealed that a variety of task parameters
213 accounted for a significant portion of firing rate variance in NAc cue-modulated units (Figure 6, Table 1).
214 Notably, there were units that discriminated between whether the rat was performing in the light or sound
215 block (28% of cue-modulated units, accounting for 6% of variance on average), which arm the rat was
216 currently on (38% of cue-modulated units, accounting for 6% of variance on average), and whether the rat
217 was engaged in the common portion of a reward-available or reward-unavailable trial (26% of cue-modulated
218 units, accounting for 4% of variance on average), suggesting that the NAc independently encodes features
219 of reward-predictive cues separate from expected outcome (Figure 5A-F). Furthermore, overlap of coding of
220 cue features within units was not different than expected by chance according to chi-square tests, suggesting
221 for integrated coding across various aspects of a cue (Figure 5G,H). Fitting a GLM to all recorded units
222 revealed a similar pattern of results (data not shown). Together, these findings show that various cue features
223 are represented in the NAc, and that this coding has some overlap, but is mostly distinct from expected
224 outcome (Figure 1; H2,H3).

225 [Table 1 about here.]

226 [Figure 5 about here.]

227 [Figure 6 about here.]

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

229 A variety of single unit response profiles were observed around the time of cue-onset. To investigate whether
230 these firing rate patterns were related to the cue information they represented, we looked at population level
231 averages for units that were modulated by each feature. To do this we normalized firing activity for each
232 unit that was modulated by a given cue feature, such as light block, then generated the cue-onset aligned
233 population average firing rate for each of the cue features (Figure 7). This analyses revealed overall that cells
234 that showed an increase upon cue presentation had stronger responses for the preferred cue condition (Figure
235 7A,C,E). Interestingly, units that were classified as decreasing in response to the cue showed a biphasic
236 response at the population level, with a small peak at a time in alignment with entry into the arm, followed by
237 a sustained dip after cue-onset (Figure 7B,D,F). Units that were modulated by cue identity showed a stronger
238 increase in response to the preferred task block, as well as a higher tonic firing rate to the preferred task block,
239 most notably in units that decreased in firing rate to the cue (Figure 7A,B). Units that were modulated by cue
240 location showed a graded response to locations of decreasing preference, with peak firing occurring around
241 cue-onset (Figure 7C,D). Units that were modulated by cue outcome showed a ramping of activity after
242 cue-onset for their preferred cue type. Additionally, units that exhibited a decrease in firing in response to
243 the cue and whose activity was modulated by cue outcome, showed a sustained discriminatory response to
244 reward-available and reward-unavailable cues that extended beyond cue-onset (Figure 7F). Together, these
245 visualizations of the averaged population responses revealed nuanced differences in the way NAc units are
246 modulated by cue conditions across cue features.

247 [Figure 7 about here.]

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

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

253 the firing rate of each unit and plotted the normalized firing rates of all units aligned to cue-onset and sorted
254 them according to the time of peak firing rate (Figure 8). We did this separately for both the light and sound
255 blocks, and found a nearly uniform distribution of firing fields in task space that was not limited to alignment
256 to the cue (Figure 8A). Furthermore, to determine if this population level activity was similar across blocks,
257 we also organized firing during the sound blocks according to the ordering derived from the light blocks.
258 This revealed that while there was some preservation of order, the overall firing was qualitatively different
259 across the two blocks. To control for the possibility that any comparison of trials would produce this effect,
260 we did a within block comparison, comparing half of the trials in the light block against the other half.
261 This comparison looked similar to our test comparison of sound block trials ordered by light block trials.
262 Additionally, given that the majority of our units showed an inhibitory response to the cue, we also plotted
263 the firing rates according to the lowest time in firing, and again found some maintenance of order, but largely
264 different ordering across the two blocks, and the within block comparison (Figure 8B). To further test this,
265 we divided each block into two halves and looked at the correlation of the average smoothed firing rates
266 across various combinations of these halves across our cue-aligned centered epoch. A linear mixed effects
267 model revealed that within block correlations (e.g. one half of light trials vs other half of light trials) were
268 higher and more similar than across block correlations (e.g. half of light trials vs half of sound trials),
269 suggesting that activity in the NAc discriminates across various cue conditions (within block correlations
270 = .381, across block correlations = .342). This process was repeated for cue location and cue outcome,
271 showing that NAc segmentation of the task is qualitatively different even during those parts of the task not
272 immediately associated with a specific cue, action, or outcome, although the within condition comparison
273 of reward-unavailable trials was less correlated than reward-available trials, and more similar to the across
274 condition comparisons (Figure 8C-F).

275 [Figure 8 about here.]

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

277 In order to be useful for reinforcement learning, a trace of the cue must be maintained until the outcome
278 to link the outcome to the outcome-predictive cue. To test whether representations of cue features were
279 maintained post-approach until the outcome was revealed, we fit a GLM to the post-approach firing rates of
280 cue-modulated units aligned to the time of nosepoke into the reward receptacle. This analysis showed that
281 a variety of units still discriminated firing according to various cue features, but not other task parameters,
282 showing that NAc activity discriminates various cue conditions well into a trial (Table 1, Figures 9,10). Ad-
283 ditionally, these units were a mix between most of the units that encoded cue features at cue-onset (observed
284 overlap greater than expected by chance according to chi-square tests), and those that did not previously have
285 a cue feature as a predictor (29, 48, and 30 units encoded both time points for cue identity, location, and
286 outcome, respectively). Fitting a GLM to all recorded units revealed a similar pattern (data not shown).
287 Population level averages for units that increased to cue-onset showed a ramping up of activity that peaked
288 upon nosepoke, whereas units that decreased to cue-onset showed a gradual reduction of firing activity that
289 reached a minimum upon nosepoke (Figure 11). Additionally, a peak is seen for preferred cue outcome
290 in decreasing units at 1 second post cue-onset when reward was received, demonstrating an integration of
291 expected and received reward (Figure 11F). Furthermore, aligning normalized peak firing rates to nosepoke
292 onset, revealed a clustering of responses around reward receipt for all cue conditions where the rat would
293 have received reward (Figure 12). To determine whether coding of cue features was maintained after the
294 outcome was revealed, a GLM was fit to the firing rates of cue-modulated units at the time of reward receipt,
295 during which the cue was still present. Fitting a GLM revealed 10 units (8%) where cue outcome accounted
296 for an average of 32% of firing rate variance (Table 1, data not shown). However, an absence of cue identity
297 or cue location coding was observed, suggesting that the NAc does not maintain a representation of these
298 cue features once the rat receives behavioral feedback for its decision.

299 [Figure 9 about here.]

300 [Figure 10 about here.]

301

[Figure 11 about here.]

302

[Figure 12 about here.]

303 **Discussion**

304 The present study found evidence for coding of multiple identifying features of motivationally relevant stim-
305 uli in addition to expected outcome; the sensory modality of the presented cue, as well as its physical location
306 within the track, and that this coding was maintained on approach trials during a delay period where the rat
307 held a nosepoke until receipt of outcome. Furthermore, this coding was both independent, and intermixed
308 across cue features and behavioral measures. Coding at the population level revealed qualitatively distinct
309 comparison profiles across cue features, such as cells that encoded cue modality tended to sustain the dis-
310 criminatory firing outside of cue presentation, while a subset cells that discriminated reward-available from
311 reward-unavailable cues maintained this discrimination until after reward receipt. Across all recorded cells,
312 a tiling of task structure was observed such that all points within our analyzed task space was accounted
313 for by the ordered peak firing rates of all cells, and this tiling differed between various conditions with a
314 cue feature, such as light versus sound blocks. Cells that discriminated across blocks were not simply due
315 to drifting of the signal across trials, as cells that showed a drift in firing between the first and second half
316 within a block were excluded from the analysis. Furthermore, even though actions were stereotyped during
317 correct trials, such that the rat always turned left at the decision point to approach for reward, and right to
318 skip the receptacle and initiate the next trial, cells that were modulated by the expected value of the cue
319 maintained their specific firing patterns even during error trials where the rat turned left after presentation of
320 the reward-unavailable cue, suggesting that these signals did not represent action values. Additionally, NAc
321 signals have been shown to be modulated by response vigor, to detangle this from our results we included
322 the trial length as a predictor in our GLMs, and found cells with correlates independent of trial length.

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

324 NAc has been shown to encode stimulus-outcome associations. In reinforcement learning, these associa-
325 tions are commonly framed as a value estimate tied to the reward-predicting stimulus. This view, while
326 informative, may not be complete in the sense that these value estimates may be more complex than just an
327 agnostic value signal, for instance, the NAc is tuned to the identity of a reward. Our goal was to expand
328 our understanding of the informational content of these outcome-predictive signals to include aspects of this
329 environmental signal expected to be encoded by the NAc based on its corticolimbic inputs (discrete cues -
330 amygdala input, spatial context - hpc input, task rules/set, behavioral strategy - pfc input). Having an en-
331 riched (multidimensional?) representation that includes details about the environmental context the animal
332 finds itself and is maintained online until the outcome is revealed could be useful for assigning credit of a
333 reward to the appropriate elements of an environment, to develop an accurate value function and optimize
334 long term acquisition of reward (2012 Lee - eligibility trace). From a motivated behavior standpoint, the
335 presence of these enriched representations can help inform action selection involving Pavlovian behaviors
336 (conditioned approach), and after a decision has been made the maintained representation can help to hold
337 the response in the face of competing alternatives while waiting for an outcome (Floresco 2015 summarizes
338 this literature).

339 If these representations function as an eligibility trace, then there should be a relationship between the con-
340 tent and robustness of these representations and the degree to which organisms form adaptive associations
341 between reinforcers and their environment, which could contribute to the presence of prediction errors in the
342 NAc in those that learned an experimental task in humans versus those that did not (Schonberg 2007). In
343 the present experiment, understanding what information about the environment the NAc is keeping track of
344 when reward arrives could be useful for potentially determining what representations are being reinforced.
345 Would be interesting to see if selectively reward an animal when certain representations are active (e.g. arm
346 1), would help shape their assignment of reward to that arm. Could be useful for altering a maladaptive pref-
347 erence back to a more adaptive one, as the loss of certain aspects of coding via BLA and PFC inputs leads to

348 the loss of outcome-predictive activity to discrete cues in the NAc. Furthermore, non-discriminatory coding
349 in the NAc could hypothetically be correlated to over-generalization of situations, in which the animal may
350 not recognize it is in a different situation and perform actions that are inefficient or maladaptive for reward
351 procurement.

352 **Cue identity:**

353 Our finding that NAc units can discriminate between outcome-predictive stimuli with similar biological
354 significance, in the current study cues from different sensory modalities, expands upon an extensive literature
355 examining neural correlates of conditioned stimuli. Perhaps the most comparable work in rodents comes
356 from a study that found distinct coding for an odor when it predicted separate but equally valued rewards
357 (Cooch). The present work is complementary to this as it shows that NAc cells have representations of
358 identifiable aspects of the cue itself, in addition to the reward it predicts. Another study paired separate
359 cues with appetitive or aversive outcomes, and found separate populations of cells that encode each cue,
360 with many switching selectivity after reversal of the associations between the cues and outcomes, providing
361 evidence that the NAc encodes the biological significance of stimuli. Once again, our study was different
362 as we recorded neural responses to distinct cues encoding the same anticipated outcome, suggesting that
363 even when the biological relevance of these stimuli is similar, the NAc dissociates their representations at
364 the level of the single-unit (Setlow), although this doesn't explain the difference in population averaged firing
365 rate for the epochs when the cue is not present. Another possibility is that these modality specific cells were
366 encoding the context, rule, or sequence within a session as some cells responded similarly for both reward-
367 available and reward-unavailable cues within a block. This interpretation is in alignment with a recent paper
368 from the non-human primate literature that recorded NAc responses during the Wisconsin Card Sorting Task
369 (WCST), a common set-shifting task used in both the laboratory and clinic, and found cells that preferred
370 firing to stimuli when a certain rule, or rule category was currently active (Sleazer). Indeed, an encoding of
371 the current strategy could be an explanation as to why a sustained difference in population averaged firing
372 was seen across stimulus blocks, as well as a potential explanation for the differentially tiling of task structure

373 across blocks in the current study. Further support for a modulation of NAc responses by strategy comes from
374 an fMRI study that examined BOLD levels during a set-shifting task (FitzGerald et al., 2014). In this task,
375 participants learned two sets of stimulus-outcome contingencies, a visual set and auditory set. During testing
376 they were presented with both simultaneously, and the stimulus dimension that was relevant was periodically
377 shifted between the two. Here, they found that bilateral NAc activity reflected value representations of
378 whatever the currently relevant stimulus dimension was, and not the irrelevant stimulus. The current finding
379 of separate, but overlapping, populations of cells encoding cue modality and expected value, suggests that
380 the fMRI finding is generated by the combined activity of several different functional cell types.

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

391 **Encoding of position:**

392 Our finding that cue-evoked activity was modulated by cue location sides with some of the literature (Lavoie,
393 1994; Tabuchi, 2000; Strait, 2016). An alternative explanation for a pure spatial representation, is that
394 these are task segmentation correlates, keeping track of where in the task the rat is. A previous non-human
395 primate paper has shown that when reward is contingent upon completion of a series of trials, separate
396 populations of NAc neurons signal the start of a schedule, subsequent trials in the schedule, and the first

397 trial in extended schedules (Shidara et al., 1998). This signalling of position within a sequence has been
398 observed in subsequent studies, and it is possible that the our rats were keeping track of which specific arm
399 they were in as part of a sequence of arms, and not just strictly a spatial representation (Mulder, 2004 and
400 2005; Khamassi et al., 2008; Berke, 2009). Also, given that our task is pseudo-random, it is possible that the
401 rats learned which cue to anticipate, and the neural activity could reflect this. However, this is unlikely as
402 including a previous trial variable in the analysis did not explain a significant amount of firing rate variance
403 in response to the cue for the vast majority of cells..

404 **Mixed selectivity:**

405 Several other papers have reported unit profiles that integrate different task-related variables. These papers
406 report integrated coding between expected value and subsequent motor responses, expected value and iden-
407 tity of a reward, and a combination of spatial-, movement-, and reward-related features (Roesch, Lavoie,
408 Cooch). However, our study is the first to show mixed selectivity among identifying features of a cue and
409 expected outcome or behavior. The presence of mixed selectivity responses confers a larger number of input-
410 output relationships that are available to a given neuron. A possible functional consequence of this attribute
411 of NAc units, is the combination and transformation of various motivationally relevant features into a sig-
412 nal informing downstream decoders such as the ventral pallidum about appropriate behaviors in obtaining
413 motivationally relevant goals and biasing action selection towards these behaviors.

414 Mixed selectivity in the NAc could be a consequence of synaptic integration from a variety of anatomically
415 distinct inputs, as seen in experiments examining the convergence of various NAc afferents at the level of
416 synaptic transmission and stimulation-induced firing (Goto and Grace 2008). In one such experiment it
417 was shown that NAc cells that responded to stimulation of either the fornix, amygdala, or PFC, typically
418 responded to stimulation from all inputs (ODonnell and Grace, 1995). Furthermore, an interaction between
419 these inputs was observed such that PFC stimulation failed to elicit spiking in the NAc neurons unless they
420 were in a depolarized UP-state, a state induced by hippocampal stimulation and was dependent on an intact

421 fornix. Hippocampal-induced suppression of other inputs has also been observed for the BLA (Mulder et al.,
422 1998). Recently, it has also been shown that train stimulation of PFC afferents reduces hippocampal-evoked
423 NAc responses, suggesting that there is competition between various inputs (Calhoon and O'Donnell, 2013).
424 These studies suggest that the integration of the variables we saw could be the result of this gating observed
425 in behaviorally-independent preparations. However, given that we did not systematically manipulate these
426 various limbic and cortical afferents, comments on the anatomical origins of the observed mixed selectivity
427 responses are speculative at this point.

428 Integrating cue identity and value, as seen modestly in the present study, could be one neural instantiation of
429 how value is associated with the appropriate predictive stimuli (credit assignment), keeping in mind that value
430 encoding is distributed, redundantly in some aspects, across various structures (Hayden Nat Neuro opinion).
431 Indeed, lesions of the NAc impair the ability to learn changes in reward value or identity in an unblocking
432 experiment, as well as disrupting dopamine RPEs generated by modification of timing of reward (McDannald
433 2011, Takahashi 2016). Would be interesting to see if uncoupling the integrated coding of stimulus features
434 and predictive properties of a cue has an effect on the ability of a rat to use reward-predictive cues to pursue
435 the associated reward.

436 **Tiling of task structure:**

437 Additionally, we found that the population of recorded units had a relatively uniform distribution of firing
438 fields within our task space, similar to what has been reported previously (Shidara, 1998; Berke, 2009;
439 Lansink, 2012). Uniquely, we found that this representation differed across conditions of a cue features,
440 such according to whether the rat was currently engaged in the light or sound block, suggesting that this
441 could be a possible neural correlate for encoding the currently relevant strategy in the NAc. It has been
442 previously shown that during progress through a predictable trial series, neurons represented state value
443 of cue (Shidara 1998), and that single-unit responses allowed the monkey to know how it was progressing
444 throughout the task. Likewise, the tiling we saw could be a consequence of upstream cortical or limbic inputs

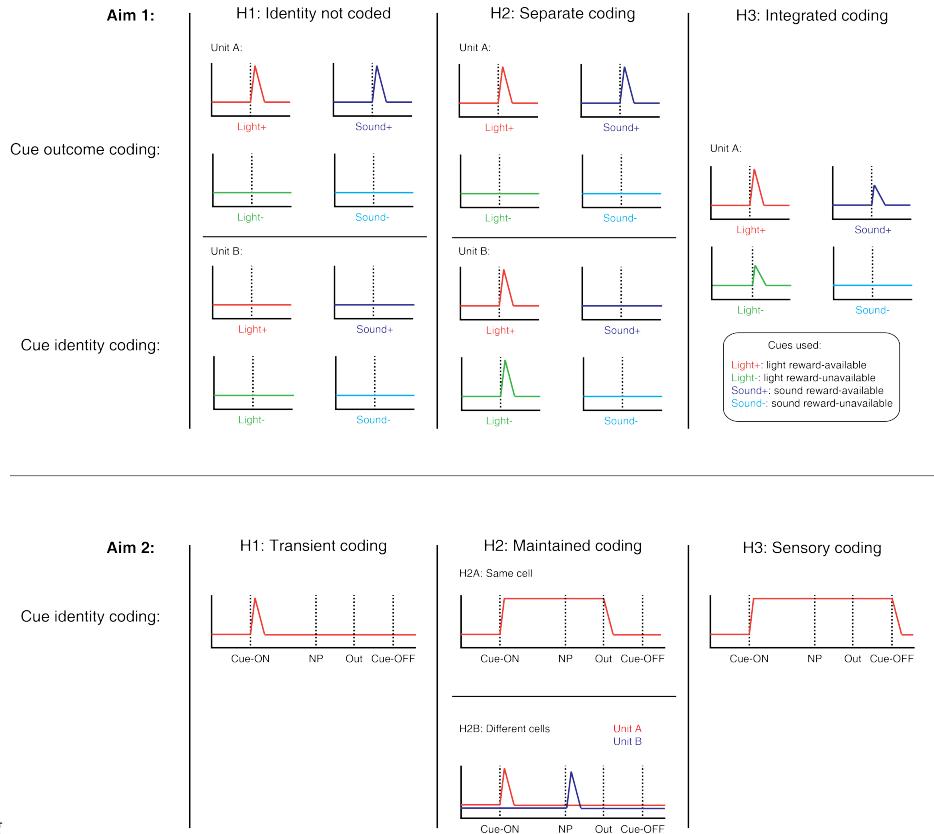
445 informing the striatum of the current task rules. Another possibility is that the NAc not only pays attention
446 to progress throughout a task within a trial, but also higher-order task information, like blocks. Furthermore,
447 dopamine levels in the NAc fluctuate through a trial, and it is possible that the observed tiling could be a
448 NAc-representation of state value related to this temporally evolving dopamine signal. Future experiments
449 should monitor this mapping of task structure during the application of dopamine antagonists. Finally, the
450 presence of functional correlates not evident when looking at single-unit responses time-locked to salient
451 task events emphasizes the need to employ ensemble level analyses across all aspects of a task.

452 References

- 453 Averbeck, B. B. and Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*,
454 20(4):505–512.
- 455 Berridge, K. C. (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. *The European
456 journal of neuroscience*, 35(7):1124–43.
- 457 Carelli, R. (2009). Drug Addiction: Behavioral Neurophysiology. In *Encyclopedia of Neuroscience*, pages 677–682. Elsevier.
- 458 Chang, S. E. and Holland, P. C. (2013). Effects of nucleus accumbens core and shell lesions on autoshaped lever-pressing. *Be-
459 havioural Brain Research*, 256:36–42.
- 460 Chang, S. E., Wheeler, D. S., and Holland, P. C. (2012). Roles of nucleus accumbens and basolateral amygdala in autoshaped lever
461 pressing. *Neurobiology of Learning and Memory*, 97(4):441–451.
- 462 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., and Wightman, R. M. (2007). Coordinated accumbal
463 dopamine release and neural activity drive goal-directed behavior. *Neuron*, 54(2):237–44.
- 464 Corbit, L. H. and Balleine, B. W. (2011). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are
465 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33):11786–11794.
- 466 Day, J. J., Roitman, M. F., Wightman, R. M., and Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine
467 signaling in the nucleus accumbens. *Nature neuroscience*, 10(8):1020–1028.
- 468 du Hoffmann, J. and Nicola, S. M. (2014). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the
469 Nucleus Accumbens. *Journal of Neuroscience*, 34(43):14349–14364.
- 470 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental
471 Psychology*, 32(2):150–155.

- 472 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,
473 H. (2011). A selective role for dopamine in stimulusreward learning. *Nature*, 469(7328):53–57.
- 474 Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of
475 Psychology*, 66(1):25–52.
- 476 Hauber, W., Bohn, I., and Giertler, C. (2000). NMDA, but not dopamine D(2), receptors in the rat nucleus accumbens areinvolved
477 in guidance of instrumental behavior by stimuli predicting reward magnitude. *The Journal of neuroscience : the official journal
478 of the Society for Neuroscience*, 20(16):6282–8.
- 479 Hearst, E. and Jenkins, H. M. (1974). *Sign-tracking: The stimulus-reinforcer relation and directed action*. Psychonomic Society.
- 480 Hyman, S. E., Malenka, R. C., and Nestler, E. J. (2006). NEURAL MECHANISMS OF ADDICTION: The Role of Reward-Related
481 Learning and Memory. *Annual Review of Neuroscience*, 29(1):565–598.
- 482 Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbensolfac-
483 tory tubercle complex. *Brain Research Reviews*, 56(1):27–78.
- 484 Joel, D., Niv, Y., and Ruppin, E. (2002). Actorcritic models of the basal ganglia: new anatomical and computational perspectives.
485 *Neural Networks*, 15(4-6):535–547.
- 486 Kalivas, P. W. and Volkow, N. D. (2005). The Neural Basis of Addiction: A Pathology of Motivation and Choice. *American Journal
487 of Psychiatry*, 162(8):1403–1413.
- 488 Khamassi, M. and Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and
489 model-free navigation strategies. *Frontiers in behavioral neuroscience*, 6:79.
- 490 Lee, D., Seo, H., and Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annu. Rev. Neurosci*,
491 35:287–308.
- 492 Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective, & Behav-
493 ioral Neuroscience*, 9(4):343–364.
- 494 Nicola, S. M. (2010). The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus
495 accumbens dopamine in the activation of reward-seeking behavior. *The Journal of neuroscience : the official journal of the
496 Society for Neuroscience*, 30(49):16585–16600.
- 497 Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psy-
498 chopharmacology*, 191(3):507–520.
- 499 Rescorla, R. A. and Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and
500 instrumental learning. *Psychological Review*, 74(3):151–182.
- 501 Robinson, T. E. and Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues
502 Through the Study of Individual Differences. *Biological Psychiatry*, 65(10):869–873.
- 503 Roitman, M. F., Wheeler, R. A., and Carelli, R. M. (2005). Nucleus Accumbens Neurons Are Innately Tuned for Rewarding and

- 504 Aversive Taste Stimuli, Encode Their Predictors, and Are Linked to Motor Output. *Neuron*, 45(4):587–597.
- 505 Salamone, J. and Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3):470–485.
- 506 Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience*,
- 507 17(3):183–195.
- 508 Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*,
- 509 275(5306):1593–9.
- 510 Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning.
- 511 *Neuron*, 38(4):625–36.
- 512 Takahashi, Y., Langdon, A., Niv, Y., and Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled by
- 513 Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1):182–193.



1 - Schematic neural.png

Figure 1: Schematic of potential coding strategies employed by single units in the NAc. Aim 1. 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 responds to a motivationally relevant variable, such as expected outcome, similarly across other cue features, such as stimulus modality or physical location. Hypothetical plot is firing rate across time. Light+ (red) signifies a reward-available light cue, Sound+ (navy blue) a reward-available sound cue, Light- (green) a reward-unavailable light cue, Sound- (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 like 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 stimulus modality, firing most for the reward-available light cue. Aim 2. Displayed are schematic PETHs illustrating potential ways in which cue identity is represented. 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. Same hypotheses applies 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

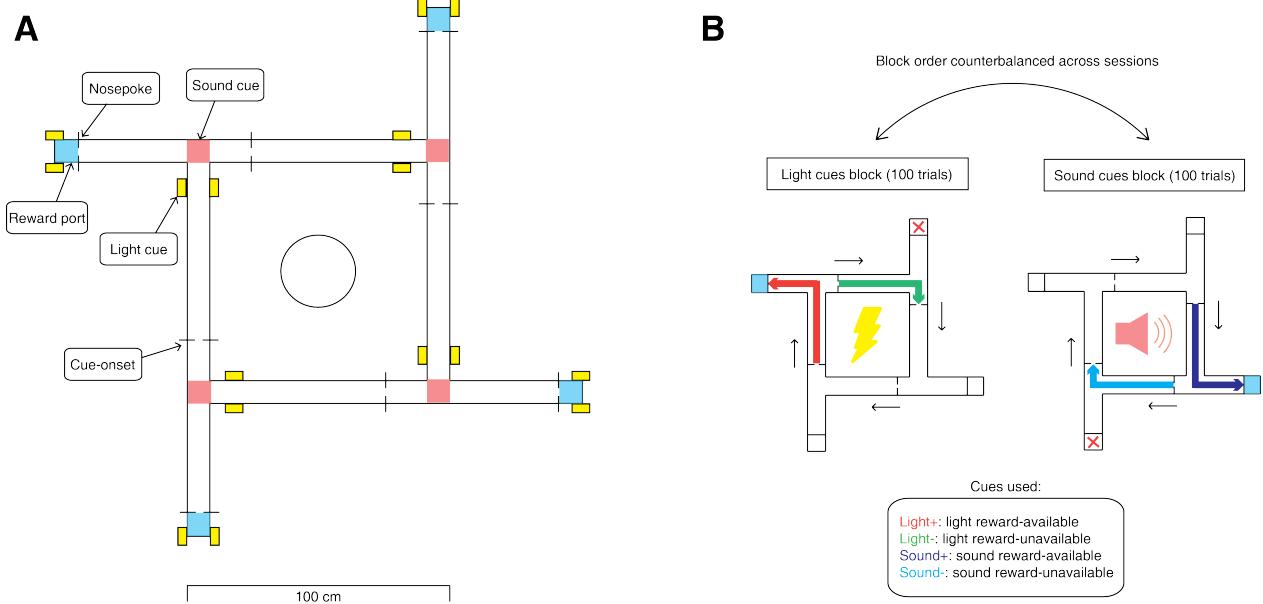


Figure 2: Schematic of behavioral task. A. To scale depiction of square track consisting of multiple identical T-choice points. Dashed lines indicate location of photobeams. Photobeams on the center, square portion of the track triggered the start of a trial and onset of a cue, whereas photobeams at 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

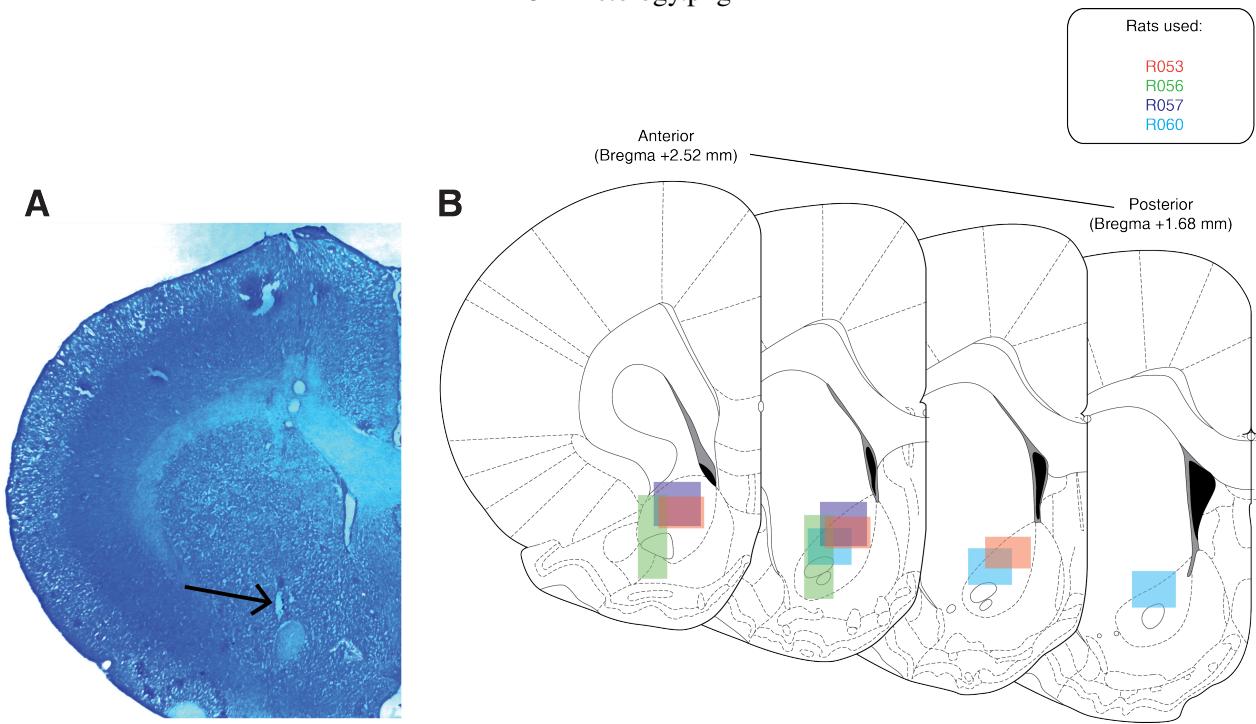


Figure 3: Histological verification of recording sites. Upon completion of experiments, brains were sectioned and tetrode placement was confirmed. A. Example section from R060 showing a recording site in the NAc core just dorsal to the anterior commissure (white circle). B. Schematic showing recording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

4 - Behavioral results.png

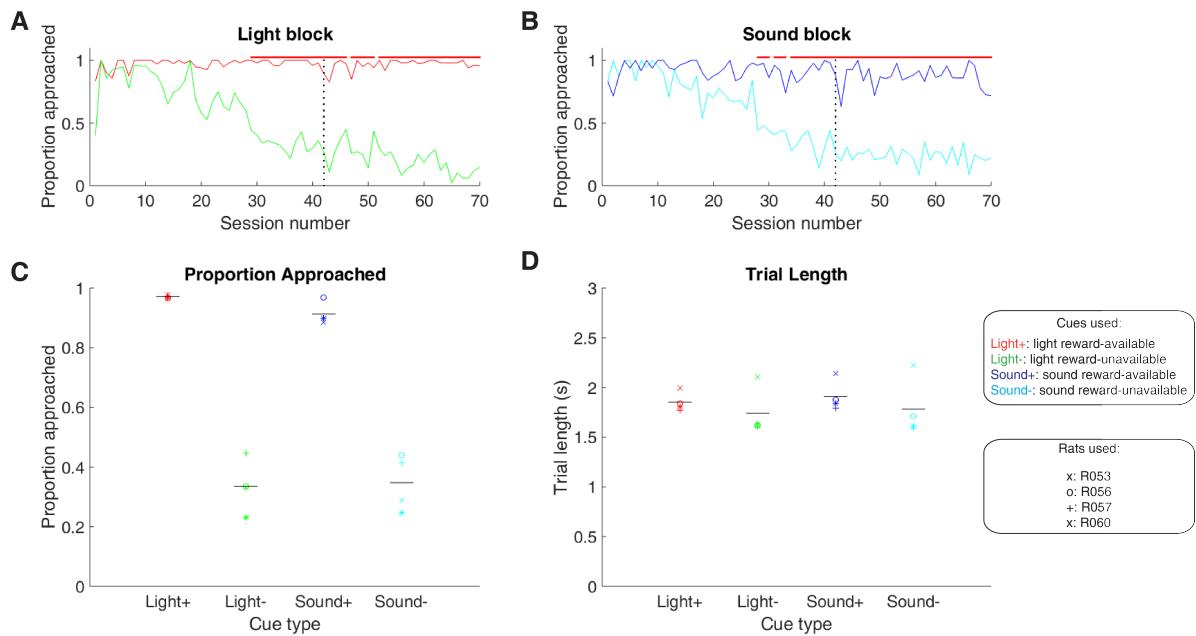


Figure 4: Performance on the behavioral task. A-B. Example learning curves across sessions from a single subject (R060) showing the proportion approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light block, light blue 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 start out approaching 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

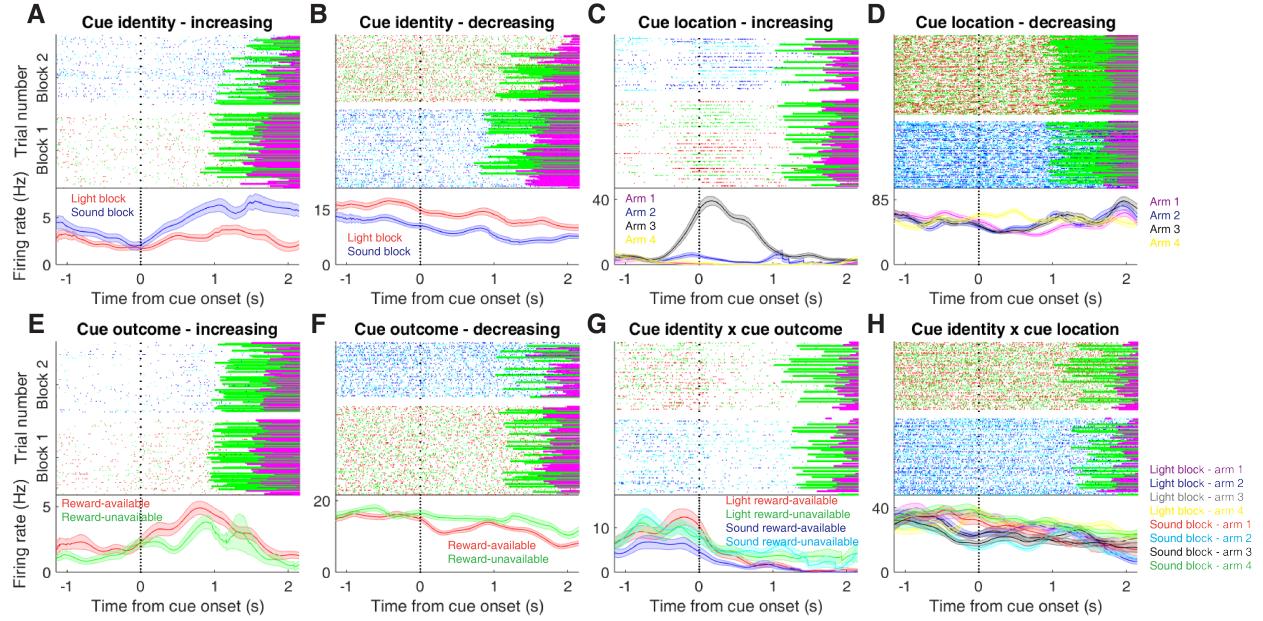


Figure 5: Examples of different cue-modulated NAc units influenced by various task parameters.

A. Example of a cue-modulated NAc unit that showed an increase in firing in response to 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. This 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

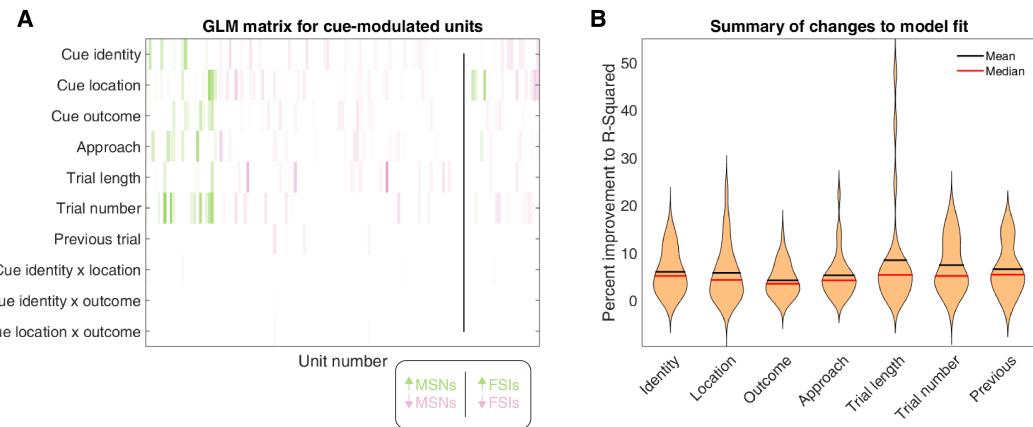


Figure 6: Summary of influence of various task parameters on cue-modulated NAc units after cue-onset. A. GLM matrix demonstrating impact of various task parameters on NAc firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank test. Each row 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 seven task parameters used in the GLM.

7 - Population averages.png

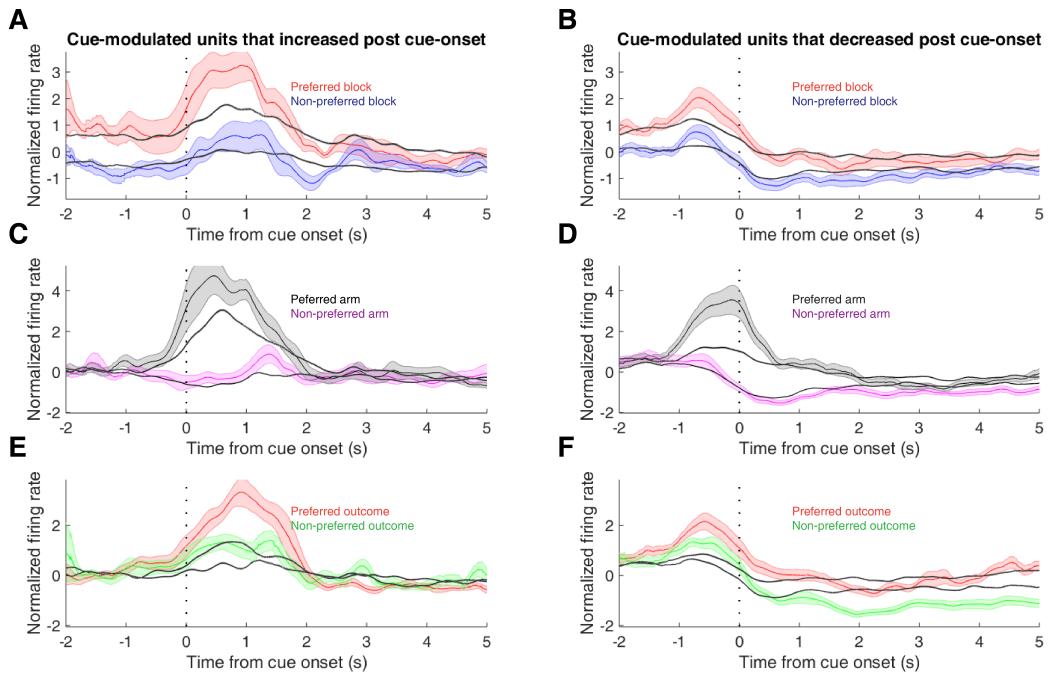
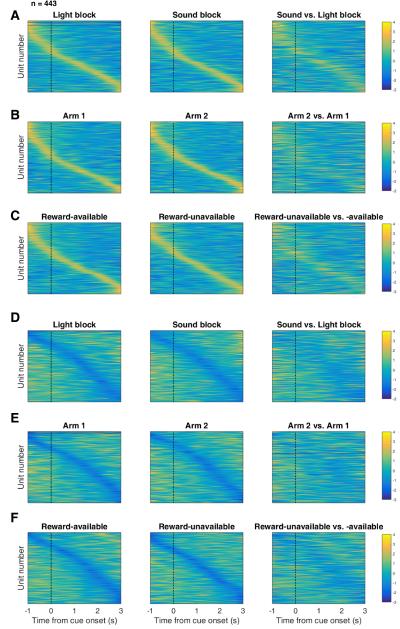


Figure 7: Population-level averages of cue feature sensitive NAc units. A. Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to cue-onset. Activity is plotted for preferred stimulus block (red) and nonpreferred stimulus 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

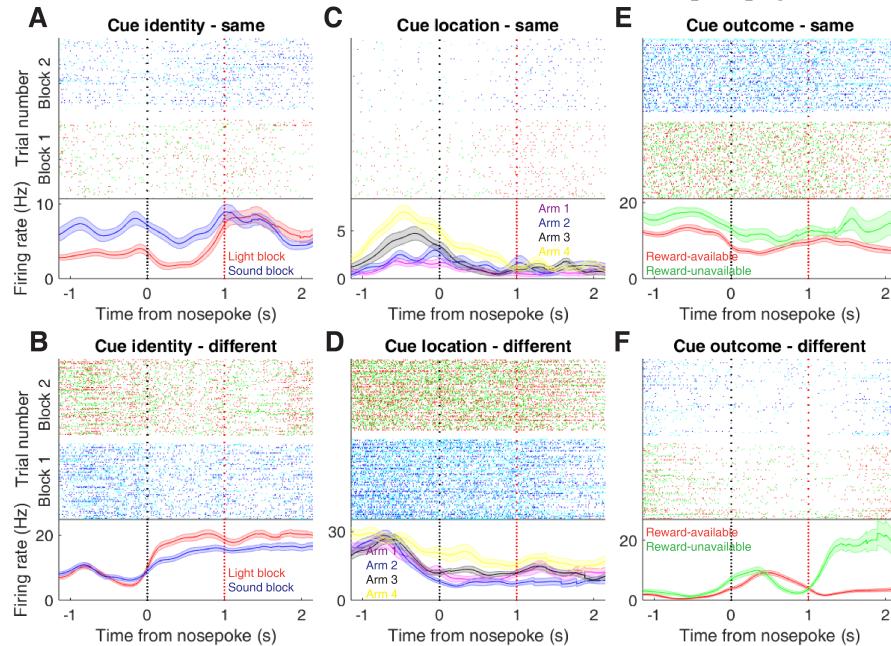


Figure 9: Examples of cue-modulated NAc units influenced by various task parameters at time of nosepoke. A. Example of a cue-modulated NAc unit that encoded cue identity at both cue-onset and during nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to nosepoke. Spikes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

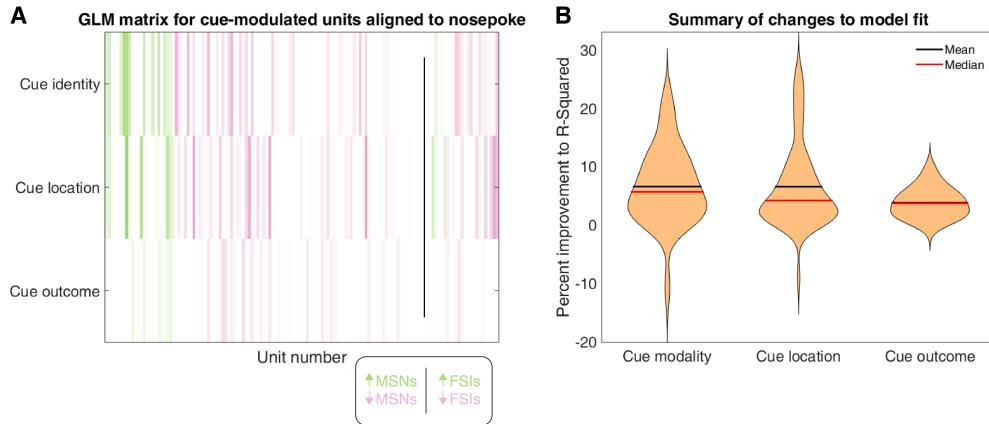


Figure 10: Summary of influence of various task parameters of cue-modulated NAc units during nosepoke. A. GLM matrix demonstrating impact of various task parameters on NAc firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank test. Each row 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

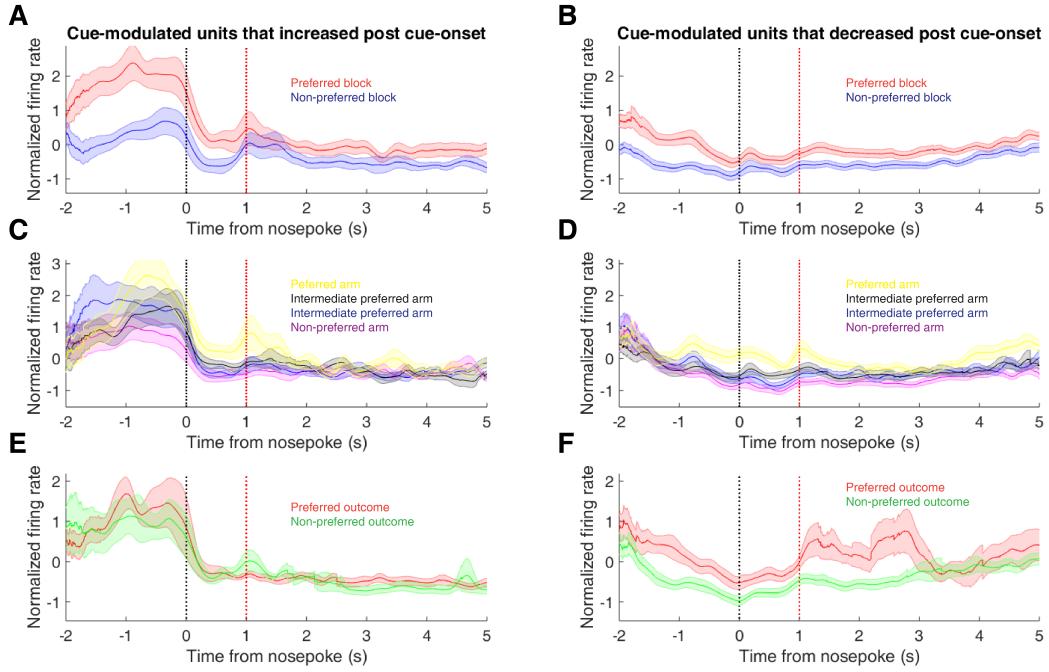
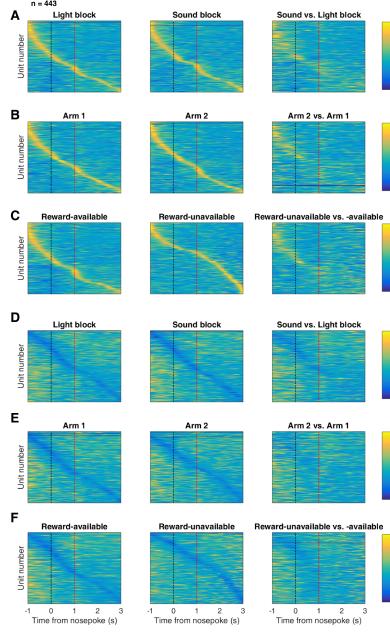


Figure 11: Population-level averages of cue feature sensitive NAc units during a nosepoke. A. Average smoothed normalized activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to nosepoke with reward delivery occurring 1 s after nosepoke. Activity is plotted 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
Rat ID					
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
GLM aligned to cue-onset					
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
GLM aligned to nosepoke					
Cue identity	66	14	36	2	14
Cue location	66	14	40	3	9
Cue outcome	42	8	29	0	5
GLM aligned to outcome					
Cue outcome	10	0	6	0	4

Table 1: Cells overview