

Persistent coding of outcome-predictive cue features in the rat 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 (eLife: 150 words)

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

16 Significance Statement (120 words)

17 to do-

18 Introduction

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

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

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

42 Niv et al., 2007), which tend to be affected by NAc manipulations (Chang et al., 2012; Corbit and Balleine,
43 2011; Flagel et al., 2011) (~~although not always straightforwardly (is Hauber appropriate?)~~; Chang and Holland, 2013; Hauber
44 (~~although not always straightforwardly~~; Giertler et al., 2004; Chang and Holland, 2013). Similarly, analysis
45 of RPEs typically proceeds by establishing an association between a cue and subsequent reward, with NAc
46 responses transferring from outcome to the cue with learning (Day et al., 2007; Roitman et al., 2005; Setlow
47 et al., 2003; Schultz et al., 1997)(~~Day 2007 is the only paper referenced here that looks at transferring of~~
48 ~~responses from reward to CS+, other papers are in VTA, or the development of CS+ responses~~).

49 Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues
50 (Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Goldstein et al., 2012;
51 Hollerman et al., 1998; Lansink et al., 2012; McGinty et al., 2013; Nicola, 2004; Roesch et al., 2009; Roit-
52 man et al., 2005; Saddoris et al., 2011; Setlow et al., 2003; Sugam et al., 2014; West and Carelli, 2016),
53 much less is known about how ~~reward-predictive outcome-predictive~~ cues themselves are encoded in the
54 NAc (Sleeker et al., 2016). This is an important issue for at least two reasons. First, in reinforcement
55 learning, motivationally relevant outcomes are typically temporally delayed relative to the cues that predict
56 them. In order to solve the problem of assigning credit (or blame) across such temporal gaps, some trace of
57 preceding activity needs to be maintained (Lee et al., 2012; Sutton and Barto, 1998). For example, if you
58 become ill after eating food X in restaurant A, depending on if you remember the identity of the restaurant
59 or the food at the time of illness, you may learn to avoid all restaurants, restaurant A only, food X only, or
60 the specific pairing of X-in-A. Therefore, a complete understanding of what is learned following feedback
61 requires understanding what trace is maintained. Since NAc is a primary target of DA signals interpretable
62 as RPEsreward prediction errors (RPEs), and NAc lesions impair RPEs related to timing, its activity trace
63 will help determine what can be learned when RPEs arrive (Hamid et al., 2015; Hart et al., 2014; Ikemoto,
64 2007; McDannald et al., 2011; Takahashi et al., 2016).

65 Second, for ongoing behavior, the relevance of cues typically depends on context. In experimental set-
66 tings, context may include the identity of a preceding cue(~~Holland, 1992~~), spatial or configural arrangements

67 ([Honey et al., 2014](#))[\(Holland, 1992; Bouton, 1993; Honey et al., 2014\)](#), and unsignaled rules as occurs in set
68 shifting and other cognitive control tasks ([Grant and Berg, 1948](#))[\(Grant and Berg, 1948; Cohen and Servan-Schreiber, 1992;](#)
69 . In such situations, the question arises how selective, context-dependent processing of [reward-predictive](#)
70 [outcome-predictive](#) cues is implemented. For instance, is there a gate prior to NAc, or are all cues rep-
71 resented in NAc but their current values dynamically updated ([FitzGerald et al., 2014; Sleezer et al., 2016](#)
72 [-](#)[\(Goto and Grace, 2008; FitzGerald et al., 2014; Sleezer et al., 2016\)](#). Representation of cue identity would
73 [allow for context-dependent mapping of outcomes predicted by specific cues.](#)

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

81 [Figure 1 about here.]

82 Methods

83 Subjects:

84 Adult male Long-Evans rats (n = 4, Charles River, Saint Constant, QC) were used as subjects. Rats were in-
85 dividually housed with a 12/12-h light-dark cycle, and tested during the light cycle. Rats were food deprived
86 to 85-90% of their free feeding weight (weight at time of implantation was 440 - 470 g), and water restricted

87 4-6 hours before testing. All experimental procedures were approved by the the University of Waterloo An-
88 imal Care Committee (protocol# 11-06) and carried out in accordance with Canadian Council for Animal
89 Care (CCAC) guidelines.

90 **Overall timeline:**

91 Each rat was first handled for seven days during which they were exposed to the experiment room, the su-
92 crose solution used as a reinforcer, and the click of the sucrose dispenser valves. Rats were then ~~shaped to~~
93 ~~run on the task for sucrose reward while being restricted to running in the clockwise direction by presenting~~
94 ~~a physical barrier to running counterclockwise. Cues signaling the availability and inavailability of reward~~
95 ~~were then introduced trained on the behavioral task~~ (described in ~~detail below, and rats were trained for~~
96 ~~200 trials per day until they discriminated between the cues~~^{the next section}) until they reached performance
97 ~~criterion~~. At this point they underwent hyperdrive implantation targeted at the NAc. Rats were allowed to
98 recover for a minimum of five days before being retrained on the task, and recording began once perfor-
99 mance returned to pre-surgery levels. Upon completion of recording, animals were gliosed, euthanized and
100 recording sites were histologically confirmed.

101 **Behavioral task and training:**

102 ~~Rats were trained to run clockwise on~~ The behavioral apparatus was an elevated, square-shaped track (100 x
103 100 cm, track width 10 cm) containing four possible reward locations at the end of track “arms” (Figure 2).
104 Rats initiated a *trial* by ~~running down the length of an arm, and~~ triggering a photobeam located 24 cm from
105 the start of each arm. Upon trial initiation, one of two possible light cues (L1, L2), or one of two possible
106 sound cues (S1, S2), was presented that signaled the presence (*reward-available trial*, L1+, S1+) or absence
107 (*reward-unavailable trial*, L2-, S2-) of a 12% sucrose water reward (0.1 mL) at the upcoming reward site. A
108 trial was classified as an *approach trial* if the rat turned left at the decision point and made a nosepoke at the
109 reward receptacle (40 cm from the decision point), while trials were classified as a *skip trial* if the rat instead

110 turned right at the decision point and triggered the photobeam to initiate the next trial. A trial is labeled
111 *correct* if the rat approached (i.e. nosepoked) on reward-available trials, and skipped (i.e. did not nosepoke)
112 on reward-unavailable trials. On reward-available trials there was a 1 second delay between a nosepoke and
113 subsequent reward delivery. *Trial length* was determined by measuring the length of time from cue onset
114 until nosepoke (for approach trials), or from cue onset until the start of the following trial (for skip trials).
115 Trials could only be initiated through clockwise progression through the series of arms, and each entry into
116 the subsequent arm on the track counted as a trial.

117 Each session consisted of both a *light block* and a *sound block* with 100 trials each. Within a block, one cue
118 signaled reward was available on that trial (L1+ or S1+), while the other signaled reward was not available
119 (L2- or S2-). Light block cues were a flashing white light, and a constant yellow light. Sound block cues
120 were a 2 kHz sine wave and a 8 kHz sine wave whose amplitude was modulated from 0 to maximum by
121 a 2 Hz sine wave. ~~Reward-cue~~ Outcome-cue associations were counterbalanced across rats, e.g. for some
122 rats L1+ was the flashing white light, and for others L1+ was the constant yellow light. The order of cue
123 presentation was pseudorandomized so that the same cue could not be presented more than twice in a row.
124 Block order within each day was also pseudorandomized, such that the rat could not begin a session with the
125 same block for more than two days in a row. Each session consisted of a 5 minute pre-session period on a
126 pedestal (a terracotta planter filled with towels), followed by the first block, then the second block, then a 5
127 minute post-session period on the pedestal. For approximately the first week of training, rats were restricted
128 to running in the clockwise direction by presenting a physical barrier to running counterclockwise. Cues
129 signaling the availability and unavailability of reward, as described above, were present from the start of
130 training. Rats were trained ~~daily until they could distinguish~~ for 200 trials per day (100 trials per block) until
131 they discriminated between the reward-available and reward-unavailable cues for both light and sound blocks
132 for three consecutive days, according to a chi-square test rejecting the null hypothesis of equal approaches
133 for reward-available and reward-unavailable trials, at which point they underwent ~~eletrode~~ electrode implant
134 surgery.

135

[Figure 2 about here.]

136 **Surgery:**

137 Surgical procedures were as described previously ([Malhotra et al., 2015](#)). ([Malhotra et al., 2015](#)). Briefly,
138 animals were administered analgesics and antibiotics, anesthetized with isoflurane, induced with 5% in med-
139 ical grade oxygen and maintained at 2% throughout the surgery (0.8 L/min). Rats were then chronically
140 implanted with a “hyperdrive” consisting of 16 independently drivable tetrodes, either all 16 targeted for the
141 right NAc (AP +1.4 mm and ML +1.6 mm relative to bregma; [Paxinos and Watson, 2005](#)[Paxinos and Watson 1998](#)
142), or 12 in the right NAc and 4 targeted at the mPFC (AP +3.0 mm and ML +0.6 mm, relative to bregma;
143 only data from NAc tetrodes was analyzed). Following surgery, all animals were given at least five days to
144 recover while receiving post-operative care, and tetrodes were lowered to the target (DV -6.0 mm) before
145 being reintroduced to the behavioral task.

146 **Data acquisition and preprocessing:**

147 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with a
148 RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex). Sig-
149 nals were referenced against a tetrode placed in the corpus callosum above the NAc. Candidate spikes for
150 sorting into putative single units were obtained by band-pass filtering the data between 600-9000 Hz, thresh-
151 olding and aligning the peaks ([UltraMegaSort2k, Hull et al., 2011](#)). ([UltraMegaSort2k, Hull et al., 2011](#))
152 Spike waveforms were then clustered with KlustaKwik using energy and the first derivative of energy as fea-
153 tures, and manually sorted into units (MClust 3.5, A.D. Redish et al., <http://redishlab.neuroscience.umn.edu/MClust/MClust.html>).
154 Isolated units containing a minimum of 200 spikes within a session were included for subsequent analysis.
155 Units were classified as fast spiking interneurons (FSIs) by an absence of interspike intervals (ISIs) > 2 s,
156 while medium spiny neurons (MSNs) had a combination of ISIs > 2 s and phasic activity with shorter ISIs
157 ([Barnes 2005, Atallah 2014](#))[\(Atallah et al., 2014; Barnes et al., 2005\)](#).

158 **Data analysis:**

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

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

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

182 dynamics of the influence of task parameters to unit activity, we then fit a sliding window GLM with the
183 same task parameters using 500 ms bins and 100 ms steps, starting 500 ms before cue-onset, up to 500 ms
184 after cue-onset, and measured the proportion of units and average R-squared value for a given time bin where
185 a particular predictor contributed significantly to the final model. To control for amount of units that would
186 be affected by a predictor by chance, we shuffled the order of firing rates used for a time bin, and took the
187 average of this value over 100 shuffles.

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

198 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units
199 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure
200 8). To compare representations of task space across cue conditions for a cue feature, the ordering of units
201 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other
202 condition (e.g. sound block). For control comparisons within cue conditions, half of the trials for a condition
203 were compared against the other half. To look at the correlation of firing rates of all eells-units within and
204 across various cue conditions, trials for each cue condition for a eell-unit were shuffled and divided into two
205 averages, and averages within and across cue conditions were correlated. A linear mixed effects model was
206 run for each cue condition to determine if correlations of firing rates within cue conditions were more similar

207 than correlations across cue conditions.

208 To identify the responsivity of units to different cue features at the time of a nosepoke into a reward re-
209 ceptacle, and subsequent reward delivery, the same cue-responsive units from the cue-onset analyses were
210 analyzed at the time of nosepoke and ~~reward~~outcome receipt using identical analysis techniques (Figures 9,
211 10, 11, and 12).

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

219 **Histology:**

220 Upon completion of the experiment, ~~rats were recording channels were gliosed by passing 150 μ A current for~~
221 ~~10 seconds and waiting 5 days before euthanasia, except for rat R057 whose implant detached prematurely.~~
222 Rats were anesthetized with 5% isoflurane, then asphyxiated with carbon dioxide. Transcardial perfusions
223 were performed, and brains were fixed and removed. Brains were sliced in 50 μ m coronal sections and
224 stained with thionin. Slices were visualized under light microscopy, tetrode placement was determined, and
225 electrodes with recording locations in the NAc were analyzed (Figure 3).

226 [Figure 3 about here.]

227 **Results**

228 **Behavior**

229 Rats were trained to discriminate between cues signaling the availability and absence of reward on a square
230 track with four identical arms for two distinct ~~sets~~set of cues (Figure 2). During each session, rats were pre-
231 sented sequentially with two behavioral blocks containing cues from different sensory modalities, a light and
232 a sound block, with each block containing a cue that signalled the availability of reward (reward-available),
233 and a cue that signalled the absence of reward (reward-unavailable). To maximize reward receipt, rats should
234 approach reward sites on reward-available trials, and skip reward sites on reward-unavailable trials (see Fig-
235 ure 4A for an example learning curve). All four rats learned to discriminate between the reward-available
236 and reward-unavailable cues for both the light and sound blocks as determined by reaching significance ($p <$
237 .05) on a daily chi-square test comparing approach behavior for reward-available and reward-unavailable
238 cues for each block, for at least three consecutive days (range for time to criterion: 22 - 57 days). Mainte-
239 nance of behavioral performance during recording sessions was assessed using linear mixed effects models
240 for both proportion of trials where the rat approached the receptacle, and trial length. Analyses revealed that
241 the likelihood of a rat to make an approach was influenced by ~~which of the cues was presented whether a~~
242 reward-available or reward-unavailable cue was presented, but was not significantly modulated by whether
243 the rat was presented with a light or sound cue (Percentage approached: light reward-available = 97%; light
244 reward-unavailable = 34%; sound reward-available = 91%; sound reward-unavailable 35%; ~~p=cue identity p~~
245 ~~= .115; cue outcome p< .001), but that cue identity did not significantly affect~~; Figure 4B). A similar trend
246 was seen with the length of time taken to complete a trial (Trial length: light reward-available = 1.85 s; light
247 reward-unavailable = 1.74 s; sound reward-available = 1.91 s; sound reward-unavailable 1.78 s; ~~cue identity~~
248 ~~p = .13; .115; cue outcome p < .001~~; Figure 4C,D). Thus, during recording, rats successfully discriminated
249 the cues according to whether or not they signaled the availability of reward at the reward receptacle.

[Figure 4 about here.]

251 **NAc ~~neurons~~ units encode behaviorally relevant and irrelevant cue features**

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

253 We sought to address which parameters of our task were encoded by NAc activity, specifically whether
254 the NAc encodes aspects of motivationally relevant cues not directly tied to reward, such as the identity
255 and location of the cue, and whether this coding is independent or integrated with coding of cue outcome.
256 To do this we recorded a total of 443 units with > 200 spikes in the NAc from 4 rats over 57 sessions
257 while they performed a cue discrimination task (Table 1). Units that exhibited a drift in firing rate over
258 the course of either block were excluded from further analysis, leaving 344 units for further analysis. The
259 activity of 133 (39%) of these 344 units were modulated by the cue, with more showing a decrease in firing
260 ($n = 103$) than an increase ($n = 30$) around the time of cue-onset (Table 1). Within this group, 24 were
261 classified as FSIs, while 109 were classified as SPNs. Upon visual inspection, we observed several patterns
262 of firing activity, including units that discriminated firing upon cue-onset across various cue conditions,
263 showed sustained differences in firing across cue conditions, had transient responses to the cue, showed a
264 ramping of activity starting at cue-onset, and showed elevated activity immediately preceding cue-onset, for
265 example (Figure 5). To characterize more formally whether these cue-evoked responses were modulated by
266 various aspects of the task, we fit a GLM to each cue-modulated unit. Fitting GLMs revealed that a variety of
267 task parameters accounted for a significant portion of firing rate variance in NAc cue-modulated units (Figure
268 6, Table 1). Notably, there were units that discriminated between whether the rat was performing in the light
269 or sound block (28% of cue-modulated units, accounting for 6% of variance on average), which arm the
270 rat was currently on (38% of cue-modulated units, accounting for 6% of variance on average), and whether
271 the rat was engaged in the common portion of a reward-available or reward-unavailable trial (26% of cue-
272 modulated units, accounting for 4% of variance on average), suggesting that the NAc ~~independently~~ encodes

273 features of reward-predictive cues separate from expected outcome (Figure 5A-F). Furthermore, overlap of
274 coding of cue features within units was not different than expected by chance according to chi-square tests,
275 suggesting for integrated coding across various aspects of a cue (Figure 5G,H). ~~Fitting a GLM to all recorded~~
276 ~~units revealed a similar pattern of results (data not shown)~~ Additionally, a sliding window GLM centered on
277 cue-onset revealed that cue identity and cue location contributed to the activity of a significant proportion
278 of cue-modulated units throughout this epoch, whereas an increase in units encoding cue outcome became
279 apparent after cue-onset (Figure 6D). Together, these findings show that various cue features are represented
280 in the NAc, and that this coding has some overlap, but is mostly distinct is both integrated and separate from
281 expected outcome (Figure 1; H2,H3).

282 [Table 1 about here.]

283 [Figure 5 about here.]

284 [Figure 6 about here.]

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

286 We observed a variety of single unit response profiles around the time of cue onset (Figure 5). To investigate
287 whether these firing rate patterns were related to what cue features were encoded, we plotted the population
288 level averages for units that were modulated by each feature. To do this, we normalized firing activity for
289 each unit that was modulated by a given cue feature, such as light block, then generated the cue-onset aligned
290 population average firing rate for each of the cue features (Figure 7). Overall, this analysis revealed that cells
291 that showed an increase upon cue presentation had stronger responses for the preferred cue condition (Figure
292 7A,C,E). Interestingly, units that were classified as decreasing in response to the cue showed a biphasic
293 response at the population level, with a small peak at a time in alignment with entry into the arm, followed by
294 a sustained dip after cue-onset (Figure 7B,D,F). Units that were modulated by cue identity showed a stronger

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

304 [Figure 7 about here.]

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

306 Given the varied time courses and response profiles of NAc units to various aspects of the cue, the NAc
307 may be computing a temporally evolving state value signal (Pennartz., 2011). If this is the case, then the
308 recruitment of NAc units should vary alongside changes in the environment. To look at the distribution of
309 responses throughout our task space and see if this distribution is modulated by cue features, we z-scored
310 the firing rate of each unit and plotted the normalized firing rates of all units aligned to cue-onset and sorted
311 them according to the time of peak firing rate (Figure 8). We did this separately for both the light and
312 sound blocks, and found a nearly uniform distribution of firing fields in task space that was not limited to
313 alignment to the cue (Figure 8A). Furthermore, to determine if this population level activity was similar
314 across blocks, we also organized firing during the sound blocks according to the ordering derived from the
315 light ~~blocks~~. (JIMMIE EDIT THIS SECTION!) Blocks. This revealed that while there was some preservation
316 of order, the overall firing was qualitatively different across the two blocks, implying that population activity
317 distinguishes between light and sound blocks. To control for the possibility that any comparison of trials
318 would produce this effect, we did a within block comparison, comparing half of the trials in the light block

against the other half. This comparison looked similar to our test comparison of sound block trials ordered by light block trials. Additionally, given that the majority of our units showed an inhibitory response to the cue, we also plotted the firing rates according to the lowest time in firing, and again found some maintenance of order, but largely different ordering across the two blocks, and the within block comparison (Figure 8B). To further test this, we divided each block into two halves and looked at the correlation of the average smoothed firing rates across various combinations of these halves across our cue-aligned centered epoch. A linear mixed effects model revealed that within block correlations (e.g. one half of light trials vs other half of light trials) were higher and more similar than across block correlations (e.g. half of light trials vs half of sound trials), suggesting that activity in the NAc discriminates across various cue conditions (within block correlations = .381, .383 (light), .379 (sound); across block correlations = .342, .343, .338, .337, .348; within block vs./ within block comparison = p = .934; within block vs./ across block comparisons = p < .001). This process was repeated for cue location (Figure 8C-D; within block correlations = .369 (arm 1), .350 (arm 2); across block correlations = .290, .286, .285, .291; within block vs./ within block comparison = p = .071; within block vs./ across block comparisons = p < .001) and cue outcome (Figure 8E-F; within block correlations = .429 (reward-available), .261 (reward-unavailable); across block correlations = .258, .253, .255, .249; within block vs./ within block comparison = p < .001; within block vs./ across block comparisons = p < .001), showing that NAc segmentation of the task is qualitatively different even during those parts of the task not immediately associated with a specific cue, action, or outcome, although the within condition comparison of reward-unavailable trials was less correlated than reward-available trials, and more similar to the across condition comparisons (Figure 8C-F), potentially due to the greater behavioral variability for the reward-unavailable trials.

[Figure 8 about here.]

Encoding of cue features is maintained persists until outcome:

In order to be useful for credit assignment in reinforcement learning, a trace of the cue must be maintained

343 until the outcome, so that information about the outcome can be associated with the outcome-predictive cue.
344 To test whether representations of cue features ~~were maintained~~persisted post-approach until the outcome
345 was revealed, we fit a GLM to the post-approach firing rates of cue-modulated units aligned to the time of
346 nosepoke into the reward receptacle. This analysis showed that a variety of units still discriminated firing
347 according to various cue features, but not other task parameters, showing that NAc activity discriminates
348 various cue conditions well into a trial (Table 1, Figures 9,10). Additionally, these units were a mix between
349 most of the units that encoded cue features at cue-onset (observed overlap greater than expected by chance
350 according to chi-square ~~tets~~tests), and those that did not previously have a cue feature as a predictor (29,
351 48, and 30 out of 133 cue-modulated units encoded both time points for cue identity, cue location, and cue
352 outcome, respectively). ~~Fitting a GLM to all recorded units revealed a similar pattern (data not shown).~~
353 Population level averages for units that increased to cue-onset showed a ramping up of activity that peaked
354 upon nosepoke, whereas units that decreased to cue-onset showed a gradual reduction of firing activity that
355 reached a minimum upon nosepoke (Figure 11). Additionally, a peak is seen for preferred cue outcome
356 in decreasing units at 1 second post cue-onset when reward was received, demonstrating an integration of
357 expected and received reward (Figure 11F). Furthermore, aligning normalized peak firing rates to nosepoke
358 onset, revealed a clustering of responses around ~~reward~~outcome receipt for all cue conditions where the
359 rat would have received reward (Figure 12). To determine whether coding of cue features ~~was maintained~~
360 persisted after the outcome was revealed, a GLM was fit to the firing rates of cue-modulated units at the
361 time of ~~reward~~outcome receipt, during which the cue was still present. Fitting a GLM revealed 10 units
362 (8%) where cue outcome accounted for an average of 32% of firing rate variance (Table 1, data not shown).
363 ~~However, an~~An absence of cue identity or cue location coding at this level of analysis was observed, but
364 looking at the data more closely with a sliding window GLM revealed that cue identity, cue location, and cue
365 outcome were encoded throughout time epochs surround cue-onset, nosepoke hold, and outcome receipt,
366 suggesting that the NAc ~~does not maintain~~maintains a representation of these cue features once the rat
367 receives behavioral feedback for its decision (Figure 10E-J).

368

[Figure 9 about here.]

369

[Figure 10 about here.]

370

[Figure 11 about here.]

371

[Figure 12 about here.]

372 **Discussion**

373 The present study found evidence for coding, by partially overlapping populations of NAc units, of multiple
374 identifying features of motivationally relevant stimuli in addition to expected outcome; the sensory modality
375 of the presented cue, as well as its physical location on the track (H2 and H3 in Figure 1A) main result of the
376 present study is that NAc units encode not only the expected outcome of outcome-predictive cues, but also
377 the identity of such cues. Importantly, this identity coding was maintained on approach trials during a delay
378 period where the rat held a nosepoke until receipt of outcome the outcome was received (H2 in Figure 1B).
379 Cells-Units coding for cue identity showed partial overlap with those coding for expected outcome (H3 in
380 Figure 1A). Units that coded different cue features (identity, outcome, location) exhibited different temporal
381 profiles as a whole, although across all recorded cells-units a tiling of task structure was observed such that
382 all points within our analyzed task space was accounted for by the ordered peak firing rates of all cellsunits.
383 Furthermore, this tiling differed between various conditions with a cue feature, such as light versus sound
384 blocks. We discuss these observations and their implications below.

385 **Cue identity:**

386 Our finding that NAc units can discriminate between different outcome-predictive stimuli with similar
387 biological significance, in the current study cues from different sensory modalities, motivational significance

388 (*i.e. encodes cue identity*) expands upon an extensive rodent literature examining NAc correlates of conditioned stimuli (Ambroggi et al., 2008; Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Day
389 et al., 2006; Dejean et al., 2017; Goldstein et al., 2012; Ishikawa et al., 2008; Lansink et al., 2012; McGinty
390 et al., 2013; Nicola, 2004; Roesch et al., 2009; Roitman et al., 2005; Saddoris et al., 2011; Setlow et al.,
391 2003; Sugam et al., 2014; West and Carelli, 2016; Yun et al., 2004). ~~Units that discriminated across blocks~~
392 ~~were not simply due to drifting of the signal across trials, as units that showed a drift in firing between the~~
393 ~~first and second half within a block were excluded from the analysis. Furthermore, even though actions~~
394 ~~were stereotyped during correct trials, such that the rat always turned left at the decision point to approach~~
395 ~~for reward, and right to skip the receptacle and initiate the next trial, units that were modulated by the~~
396 ~~expected value of the cue maintained their specific firing patterns even during error trials where the rat~~
397 ~~turned left after presentation of the reward-unavailable cue, suggesting that these signals did not represent~~
398 ~~action values. Additionally, NAc signals have been shown to be modulated by response vigor, to detangle~~
399 ~~this from our results we included the trial length as a predictor in our GLMs, and found units with cue feature~~
400 ~~correlates independent of trial length (McGinty et al., 2013).~~ Perhaps the most comparable work in rodents
401 comes from a study that found distinct coding for an odor when it predicted separate but equally valued
402 rewards (Cooch et al., 2015). The present work is complementary to ~~this such outcome identity coding~~ as
403 it shows that NAc units ~~have representations of identifiable aspects of the cue itself~~*cue identity*, both sepa-
404 rately and integrated with the reward it predicts (H2 and H3 in Figure 1A). ~~Another study paired separate~~
405 ~~Similarly, Setlow et al. 2003 paired distinct~~ cues with appetitive or aversive outcomes, and found separate
406 populations of units that encoded each cue, ~~with many switching selectivity after reversal of the associations~~
407 ~~between the cues and outcomes, providing evidence that the NAc encodes the biological significance of~~
408 ~~stimuli (Setlow et al., 2003)~~. Once again, our study was different ~~as we recorded neural responses to in~~
409 ~~asking how~~ distinct cues encoding the same anticipated outcome, ~~suggesting are encoded. Such cue identity~~
410 ~~encoding suggests~~ that even when the biological relevance of these stimuli is similar ~~the~~, NAc dissociates
411 their representations at the level of the ~~single-unit, although this doesnt explain the difference in population~~
412 ~~averaged firing rate for the epochs when the cue is not present~~*single-units*. A possible interpretation of this
413 coding of cue features alongside expected outcome is that these representations are used to associate reward
414 with relevant features of the environment, so-called credit assignment in the reinforcement learning litera-
415

ture (Sutton and Barto, 1998). A burgeoning body of human and non-human primate work has started to elucidated neural correlates of credit assignment in the PFC, particularly in the lateral orbitofrontal cortex (Akaishi et al., 2016; Asaad et al., 2017; Chau et al., 2015; Noonan et al., 2017). Given the importance of cortical inputs in NAc associative representations, it is possible that information related to credit assignment is relayed from the cortex to NAc (Coch et al., 2015; Ishikawa et al., 2008).

Another possibility is that these modality specific units were encoding the A different possible function for cue identity coding is to support contextual modulation of the motivational relevance of specific cues. A context can be understood as a particular mapping between specific cues and their outcomes: for instance, in context 1 cue A but not cue B is rewarded, whereas in context , rule, or sequence within a session as some units responded similarly for both reward available and reward unavailable cues within a block. This interpretation is in alignment with a recent paper from the non-human primate literature that 2 cue B but not cue A is rewarded. Successfully implementing such contextual mappings requires representation of the cue identities. Indeed, (Sleeker et al., 2016) recorded NAc responses during the Wisconsin Card Sorting Task (WCST), a common set-shifting task used in both the laboratory and clinic, and found units that preferred firing to stimuli when a certain rule, or rule category was currently active(Sleeker et al., 2016). Indeed, an encoding of the current strategy could be an explanation as to why a sustained difference in population averaged firing was seen across stimulus blocks, as well as a potential explanation for the differentially tiling of task structure across blocks in the current study. Further support for a modulation of NAc responses by strategy comes from an fMRI study that examined BOLD levels during a set-shifting task (FitzGerald et al., 2014). In this task, participants learned two sets of stimulus-outcome contingencies, a visual set and auditory set. During testing they were presented with both simultaneously, and the stimulus dimension that was relevant was periodically shifted between the two. Here, they found that bilateral NAc activity reflected value representations for the currently relevant stimulus dimension, and not the irrelevant stimulus. The current finding of separate, but overlapping, populations of units encoding cue modality and expected valueidentity and expected outcome, suggests that the fMRI finding is generated by the combined activity of several different functional cell types.

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

451 A different potential confound is that between outcome and action value coding. We discriminated between
452 these possibilities by analyzing error trials, where the rat approached reward (left turn) after presentation of
453 the reward-unavailable cue. Units that were modulated by the expected outcome of the cue maintained their
454 specific firing patterns even during error trials, as expected from outcome value coding but not action value
455 coding. Additionally, NAc signals have been shown to be modulated by response vigor (McGinty et al., 2013)
456 ; to detangle this from our results we included trial length (i.e. latency to arrival at the reward site) as a
457 predictor in our GLMs, and found units with cue feature correlates independent of trial length.

458 An overall limitation of the current study is that rats were never presented with both sets of cues simulta-
459 neously, and thus never had were not required to switch strategies , although extrapolating the data from
460 the primate study suggests that the activity of the cue identity units between multiple sets of cues. Thus,
461 it is unknown to what extent the cue identity encoding we observed is behaviorally relevant, although
462 extrapolating data from other work (Sleeker et al., 2016) suggests that cue identity coding would be mod-
463 ulted by relevance. Keeping along this theme, the current data set is unable to identify precisely what the
464 modality-sensitive neurons were encoding, that is were they tracking representations of stimulus identity, a
465 preferred context, or even a macroscale representation of progress through the session. Furthermore, their
466 relevance for ongoing behavior is also uncertain. NAc core lesions have been shown to impair shifting be-

467 tween different behavioral strategies (Floresco et al., 2006), and it is possible that selectively silencing the
468 units that prefer responding for a given modality or rule would impair performance when the animal is re-
469 quired to use that information, or artificial enhancement of those units would cause them to use the rule when
470 it is the inappropriate strategy.

471 **Encoding of position:**

472 Our finding that cue-evoked activity was modulated by cue location is in alignment with several previ-
473 ous reports (Lavoie and Mizumori, 1994; Mulder et al., 2005; Strait et al., 2016; Wiener et al., 2003).
474 The NAc receives ~~strong~~ inputs from the hippocampus, and the communication of place-reward informa-
475 tion across the two structures suggests that the NAc ~~might be keeping track of tracks~~ locations associ-
476 ated with reward (Lansink et al., 2008, 2009, 2016; Pennartz, 2004; Sjulson et al., 2017; Tabuchi et al.,
477 2000; van der Meer and Redish, 2011). ~~Alternatively, it is possible that this positional modulation is part~~
478 ~~of a representation of task space, keeping track of where in the task the rat is. A previous non-human~~
479 ~~primate experiment showed that when reward is contingent upon completion of a series of trials, separate~~
480 ~~populations of NAc neurons signal the start of a schedule, subsequent trials in the schedule, and the first trial~~
481 ~~in extended schedules (Shidara et al., 1998). This signalling of position within a sequence has been observed~~
482 ~~in subsequent studies, and it is possible that the rats were keeping track of which specific arm they were in as~~
483 ~~part of a sequence of arms, especially since they started from the same arm during the start of each session~~
484 ~~(Atallah et al., 2014; Berke et al., 2009; Khamassi et al., 2008; Lansink et al., 2012; Mulder et al., 2004) NAc~~
485 ~~units can also signal progress through a sequence of cues and/or actions (Shidara et al., 1998; Atallah et al., 2014; Berke et al.~~
486 . Given that the current task was pseudo-random, it is possible that the rats learned the structure of sequential
487 cue presentation, and the neural activity could reflect this. However, this is unlikely as including a previous
488 trial variable in the analysis did not explain a significant amount of firing rate variance in response to the
489 cue for the vast majority of ~~cells. A limitation of the current study is that the behavioral relevance of cue~~
490 ~~location was not manipulated in the sense that varying cue location did not change the outcome-predictive~~
491 ~~nature of the cue units. In any case, NAc units on the present task continued to distinguish between different~~

492 locations, even though location, and progress through a sequence, were explicitly irrelevant in predicting
493 reward. We speculate that this persistent coding of location in NAc may represent a bias in credit assignment,
494 and associated tendency for rodents to associate motivationally relevant events with the locations where they
495 occur.

496 **Potential functional consequences of maintained coding****Implications:**

497 In the current study we found that the coding of cue features persisted after the choice point, during a
498 delay period while the rat waited at the receptacle for the outcome (H2 in Figure 1B). Having an enriched
499 representation that includes details about the environmental context the animal finds itself in alongside
500 expected outcome, and is maintained online until the outcome is revealed could be useful for assigning
501 credit of a reward to the appropriate elements of an environment, to develop an accurate value function and
502 optimize long term acquisition of reward (Lee et al., 2012). A recent non-human primate experiment found
503 evidence for simultaneous coding of outcome and outcome-predictive stimuli at the time of feedback in the
504 dorsolateral PFC, and that this coding of stimulus information was strongest before behavioral performance
505 stabilized (Asaad et al., 2017). Interestingly, we found coding of cue identity and location after presentation
506 Maladaptive decision making, as occurs in schizophrenia, addiction, Parkinsons, among others, can result
507 from dysfunctional RPE and value signals (Frank et al., 2004; Gradin et al., 2011; Maia and Frank, 2011)
508 . This view has been successful in explaining both positive and negative symptoms in schizophrenia,
509 and deficits in learning from feedback in Parkinsons (Frank et al., 2004; Gradin et al., 2011). However, the
510 effects of RPE and value updating are contingent upon encoding of preceding action and cue features,
511 the eligibility trace (Lee et al., 2012; Sutton and Barto, 1998). Value updates can only be performed on
512 these aspects of preceding experience that are encoded when the update occurs. Therefore, maladaptive
513 learning and decision making can result from not only aberrant RPEs but also from altered cue feature
514 encoding. For instance, on this task the environmental stimulus that signaled the availability of reward
515 was conveyed by two distinct cues that were presented in four locations. While in our current study, the
516 location and identity of the cue , and during a delay period between nosepoke and outcome, but not after

517 the outcome was received. The absence of coding during feedback, and that these recordings were done
518 after behavioral performance had stabilized suggests the possibility that these NAc representations represent
519 an integrated value estimate that is did not require any adjustments in the animals behavior, we found
520 coding of these features alongside the expected outcome of the cue that could be the outcome of earlier
521 upstream credit assignment in cortical inputs. From a motivated behavior standpoint, the presence of these
522 enriched representations can help inform action selection involving Pavlovian behaviors such as conditioned
523 approach (Ciano et al., 2001; Parkinson et al., 2000; Saunders and Robinson, 2012), and after a decision has
524 been made the maintained representation can help to hold the response in the face of competing alternatives
525 while waiting for an outcome (Di Ciano et al., 2008; Floresco et al., 2008; Floresco, 2015; Peters et al., 2008)
526 .credit assignment computations computed upstream. Identifying neural coding related to an aspect of
527 credit assignment is important as inappropriate credit assignment could be a contributor to conditioned fear
528 overgeneralization seen in disorders with pathological anxiety such as generalized anxiety disorder, post
529 traumatic stress disorder, and obsessive-compulsive disorder (Shmuel Lissek, 2013; Kaczkurkin et al., 2017; Lissek et al., 2017
530 , and delusions observed in disorders such as schizophrenia, Alzheimers and Parkinsons (Corlett et al., 2010; Kapur, 2003)
531 . Thus, our results provide a neural window into the process of credit assignment, such that the extent and
532 specific manner in which this process fails in e.g. syndromes such as schizophrenia, obsessive-compulsive
533 disorder, etc. can be experimentally accessed.

534 OUTTAKES

535 **Tiling of task structure:**

536 Additionally, we We found that the population of recorded units had a relatively uniform distribution of firing
537 fields within our task space, similar to what has been reported previously (Berke et al., 2009; Lansink et al.,
538 2012; Shidara et al., 1998). Uniquely, we found that this representation was statistically different across
539 conditions of a cue features, such as according to whether the rat was currently engaged in the light or sound
540 block, suggesting that this could be a possible neural correlate for encoding the currently relevant strategy

541 in the NAc. It has been previously shown that during progress through a predictable trial series, **neurons**
542 **units** represented state value of cue, and that single-unit responses allowed the monkey to know how it was
543 progressing throughout the task (Shidara et al., 1998). Likewise, the tiling we saw could be a consequence
544 of upstream cortical or limbic inputs informing the striatum of the current task rules. Another possibility is
545 that the NAc not only pays attention to progress throughout a task within a trial, but also higher-order task
546 information, like blocks. Furthermore, dopamine levels in the NAc fluctuate through a trial, and it is possible
547 that the observed tiling could be a NAc-representation of state value related to this temporally evolving
548 dopamine signal. Future experiments should monitor this mapping of task structure during the application of
549 dopamine antagonists. Finally, the presence of functional correlates not evident when looking at single-unit
550 responses time-locked to salient task events emphasizes the need to employ ensemble level analyses across
551 all aspects of a task.

552 **Implications****Potential functional consequences of persistent coding:**

553 ~~Maladaptive decision making, as occurs in schizophrenia, addiction, Parkinsons, among others, can result~~
554 ~~from dysfunctional RPE and value signals (Frank et al., 2004; Gradin et al., 2011; Maia and Frank, 2011)~~
555 ~~. This view has been successful in explaining both positive and negative symptoms in schizophrenia,~~
556 ~~and deficits in learning from feedback in Parkinsons (Frank et al., 2004; Gradin et al., 2011). However, the~~
557 ~~effects of RPE and value updating are contingent upon encoding of preceding action and cue features, the~~
558 ~~eligibility trace (Lee et al., 2012; Sutton and Barto, 1998). Value updates can only be performed on these~~
559 ~~aspects of preceding experience that are encoded when the update occurs. Therefore, maladaptive learning~~
560 ~~and decision making can result from not only aberrant RPEs but also from altered cue feature encoding.~~
561 ~~For instance, on this task the environmental stimulus that signaled the availability of reward was conveyed~~
562 ~~by two distinct cues that were presented in four locations. While in our current study, the location and~~
563 ~~identity~~ In the current study we found that the coding of cue features persisted after the choice point, during
564 a delay period while the rat waited at the receptacle for the outcome (H2 in Figure 1B). Having an enriched
565 representation that includes details about the environmental context the animal finds itself in alongside

566 expected outcome, and is maintained online until the outcome is revealed could be useful for assigning
567 credit of a reward to the appropriate elements of an environment, to develop an accurate value function and
568 optimize long term acquisition of reward (Lee et al., 2012). A recent non-human primate experiment found
569 evidence for simultaneous coding of outcome and outcome-predictive stimuli at the time of feedback in the
570 dorsolateral PFC, and that this coding of stimulus information was strongest before behavioral performance
571 stabilized (Asaad et al., 2017). Interestingly, we found coding of cue identity and location after presentation
572 of the cue~~did not require any adjustments in the animals behavior, we found coding of these features~~
573 alongside the expected outcome of the cue that could be~~, and during a delay period between nosepoke~~
574 and outcome, but not after the outcome was received. The absence of coding during feedback, and that these
575 recordings were done after behavioral performance had stabilized suggests the possibility that these NAc
576 representations represent an integrated value estimate that is~~the outcome of credit assignment computations~~
577 computed upstream. Identifying neural coding related to an aspect of credit assignment is important
578 as inappropriate credit assignment could be a contributor to conditioned fear overgeneralization seen in
579 disorders with pathological anxiety such as generalized anxiety disorder, post traumatic stress disorder,
580 and obsessive-compulsive disorder (Shmuel Lissek, 2013; Kaeckurkin et al., 2017; Lissek et al., 2014), and
581 delusions observed in disorders such as schizophrenia, Alzheimers and Parkinsons (Corlett et al., 2010; Kapur, 2003)
582 . Thus, our results provide a neural window into the process of credit assignment, such that the extent and
583 specific manner in which this process fails in e. g. syndromes such as schizophrenia, obsessive-compulsive
584 disorder, etc. can be experimentally accessed. earlier upstream credit assignment in cortical inputs. From
585 a motivated behavior standpoint, the presence of these enriched representations can help inform action
586 selection involving Pavlovian behaviors such as conditioned approach (Ciano et al., 2001; Parkinson et al., 2000; Saunders and
587 , and after a decision has been made the maintained representation can help to hold the response in the face of
588 competing alternatives while waiting for an outcome (Di Ciano et al., 2008; Floresco et al., 2008; Floresco, 2015; Peters et al.
589 ~

590 Orphan paragraph? (also removed mixed selectivity paragraphs)

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

604 A fundamental problem faced by all reinforcement learning agents is which features to credit and blame for
605 particular outcomes (Sutton and Barto, 1998); adaptive behavior requires associating motivationally relevant
606 outcomes with the cues that predict them through learning from feedback. Much work has focused on value
607 signals such as reward prediction errors (RPEs), state values and action values (Lee et al., 2012; Maia, 2009)
608 . However, successful learning requires not only a RPE but also a trace of the preceding actions and/or cues.

609 References

- 610 Akaishi, R., Kolling, N., Brown, J. W., and Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue
611 Environment. *Journal of Neuroscience*, 36(4):1096–1112.
612 Ambroggi, F., Ishikawa, A., Fields, H. L., and Nicola, S. M. (2008). Basolateral Amygdala Neurons Facilitate Reward-Seeking
613 Behavior by Exciting Nucleus Accumbens Neurons. *Neuron*, 59(4):648–661.
614 Asaad, W. F., Lauro, P. M., Perge, J. A., and Eskandar, E. N. (2017). Prefrontal Neurons Encode a Solution to the Credit-Assignment

- 615 Problem. *The Journal of Neuroscience*, 37(29):6995–7007.
- 616 Atallah, H. E., McCool, A. D., Howe, M. W., and Graybiel, A. M. (2014). Neurons in the ventral striatum exhibit cell-type-specific
617 representations of outcome during learning. *Neuron*, 82(5):1145–1156.
- 618 Averbeck, B. B. and Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*,
619 20(4):505–512.
- 620 Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., and Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and
621 recoding of procedural memories. *Nature*, 437(7062):1158–1161.
- 622 Berke, J. D., Breck, J. T., and Eichenbaum, H. (2009). Striatal Versus Hippocampal Representations During Win-Stay Maze
623 Performance. *Journal of Neurophysiology*, 101(3):1575–1587.
- 624 Berridge, K. C. (2012). From prediction error to incentive salience: Mesolimbic computation of reward motivation. *European
625 Journal of Neuroscience*, 35(7):1124–1143.
- 626 Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., Kashtelyan, V., and Roesch, M. R.
627 (2013). Separate Populations of Neurons in Ventral Striatum Encode Value and Motivation. *PLoS ONE*, 8(5):e64673.
- 628 Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological
629 bulletin*, 114(1):80–99.
- 630 Carelli, R. M. (2010). Drug Addiction: Behavioral Neurophysiology. In *Encyclopedia of Neuroscience*, pages 677–682. Elsevier.
- 631 Chang, S. E. and Holland, P. C. (2013). Effects of nucleus accumbens core and shell lesions on autoshaped lever-pressing. *Be-
632 havioural Brain Research*, 256:36–42.
- 633 Chang, S. E., Wheeler, D. S., and Holland, P. C. (2012). Roles of nucleus accumbens and basolateral amygdala in autoshaped lever
634 pressing. *Neurobiology of Learning and Memory*, 97(4):441–451.
- 635 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).
636 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5):1106–
637 1118.
- 638 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., and Wightman, R. M. (2007). Coordinated Accumbal
639 Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2):237–244.
- 640 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.,
641 and Everitt, B. J. (2001). Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens
642 core in the acquisition and performance of pavlovian approach behavior. *The Journal of neuroscience*, 21(23):9471–9477.
- 643 Cohen, J. D. and Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in
644 schizophrenia. *Psychological review*, 99(1):45.
- 645 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., and Schoenbaum, G. (2015).
646 Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature Communi-*

- 647 cations, 6:7195.
- 648 Corbit, L. H. and Balleine, B. W. (2011). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are
649 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33):11786–11794.
- 650 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., and Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in
651 Neurobiology*, 92:345–369.
- 652 Day, J. J., Roitman, M. F., Wightman, R. M., and Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine
653 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8):1020–1028.
- 654 Day, J. J., Wheeler, R. A., Roitman, M. F., and Carelli, R. M. (2006). Nucleus accumbens neurons encode Pavlovian approach
655 behaviors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5):1341–1351.
- 656 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., Boraud, T., and Le Moine, C. (2017). Memories of Opiate
657 Withdrawal Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychopharmacology*,
658 42(5):1157–1168.
- 659 Di Ciano, P., Robbins, T. W., and Everitt, B. J. (2008). Differential effects of nucleus accumbens core, shell, or dorsal striatal
660 inactivations on the persistence, reacquisition, or reinstatement of responding for a drug-paired conditioned reinforcer. *Neu-
661 ropsychopharmacology*, 33(6):1413–1425.
- 662 du Hoffmann, J. and Nicola, S. M. (2014). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the
663 Nucleus Accumbens. *Journal of Neuroscience*, 34(43):14349–14364.
- 664 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental
665 Psychology*, 32(2):150–155.
- 666 Fitzgerald, T. H. B., Schwartenbeck, P., and Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent
667 and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4):1271–1279.
- 668 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,
669 H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, 469(7328):53–59.
- 670 Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of
671 Psychology*, 66(1):25–52.
- 672 Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., and Magyar, O. (2006). Dissociable roles for the nucleus accumbens core and
673 shell in regulating set shifting. *J Neurosci*, 26(9):2449–2457.
- 674 Floresco, S. B., McLaughlin, R. J., and Haluk, D. M. (2008). Opposing roles for the nucleus accumbens core and shell in cue-induced
675 reinstatement of food-seeking behavior. *Neuroscience*, 154(3):877–884.
- 676 Frank, M. J., Seeberger, L. C., and O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism.
677 *Science*, 306(5703):1940–1943.
- 678 Giertler, C., Bohn, I., and Hauber, W. (2004). Transient inactivation of the rat nucleus accumbens does not impair guidance of

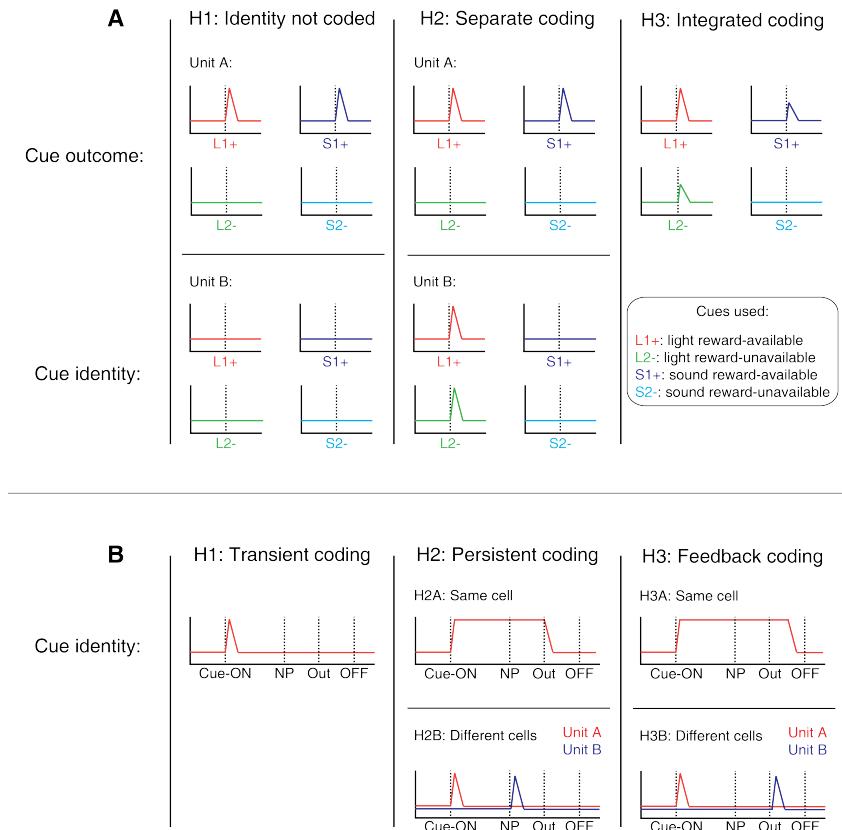
- 679 instrumental behaviour by stimuli predicting reward magnitude. *Behavioural Pharmacology*, 15(1):55–63.
- 680 Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., Bryden, D. W., and Roesch, M. R.
681 (2012). Ventral Striatum Encodes Past and Predicted Value Independent of Motor Contingencies. *Journal of Neuroscience*,
682 32(6):2027–2036.
- 683 Goto, Y. and Grace, A. A. (2008). Limbic and cortical information processing in the nucleus accumbens. *Trends in Neurosciences*,
684 31(11):552–558.
- 685 Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., and Steele, J. D. (2011). Expected value
686 and prediction error abnormalities in depression and schizophrenia. *Brain*, 134(6):1751–1764.
- 687 Grant, D. A. and Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a
688 Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4):404–411.
- 689 Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J.,
690 and Berke, J. D. (2015). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1):117–126.
- 691 Hart, A. S., Rutledge, R. B., Glimcher, P. W., and Phillips, P. E. M. (2014). Phasic Dopamine Release in the Rat Nucleus Accumbens
692 Symmetrically Encodes a Reward Prediction Error Term. *The Journal of Neuroscience*, 34(3):698–704.
- 693 Hauber, W., Bohn, I., and Giertler, C. (2000). NMDA, but not dopamine D(2), receptors in the rat nucleus accumbens are involved
694 in guidance of instrumental behavior by stimuli predicting reward magnitude. *J Neurosci*, 20(16):6282–6288.
- 695 Hearst, E. and Jenkins, H. M. (1974). *Sign-tracking: the stimulus-reinforcer relation and directed action*. Psychonomic Society.
- 696 Hill, D. N., Mehta, S. B., and Kleinfeld, D. (2011). Quality Metrics to Accompany Spike Sorting of Extracellular Signals. *Journal
697 of Neuroscience*, 31(24):8699–8705.
- 698 Holland, P. C. (1992). Occasion setting in pavlovian conditioning. *Psychology of Learning and Motivation - Advances in Research
699 and Theory*, 28(C):69–125.
- 700 Hollerman, J. R., Tremblay, L., and Schultz, W. (1998). Influence of Reward Expectation on Behavior-Related Neuronal Activity in
701 Primate Striatum. *Journal of Neurophysiology*, 80(2):947–963.
- 702 Honey, R. C., Iordanova, M. D., and Good, M. (2014). Associative structures in animal learning: Dissociating elemental and
703 configural processes. *Neurobiology of Learning and Memory*, 108:96–103.
- 704 Hyman, S. E., Malenka, R. C., and Nestler, E. J. (2006). NEURAL MECHANISMS OF ADDICTION: The Role of Reward-Related
705 Learning and Memory. *Annual Review of Neuroscience*, 29(1):565–598.
- 706 Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-
707 olfactory tubercle complex. *Brain Research Reviews*, 56(1):27–78.
- 708 Ishikawa, A., Ambroggi, F., Nicola, S. M., and Fields, H. L. (2008). Dorsomedial Prefrontal Cortex Contribution to Behavioral and
709 Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, 28(19):5088–5098.
- 710 Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives.

- 711 *Neural Networks*, 15(4-6):535–547.
- 712 Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., Sponheim, S. R., and Lissek,
713 S. (2017). Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174(2):125–134.
- 714 Kalivas, P. W. and Volkow, N. D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal
715 of Psychiatry*, 162(8):1403–1413.
- 716 Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in
717 schizophrenia. *American Journal of Psychiatry*, 160(1):13–23.
- 718 Khamassi, M. and Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and
719 model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, 6:79.
- 720 Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., and Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal
721 neurons of behaving rats. *European Journal of Neuroscience*, 28(9):1849–1866.
- 722 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., and Pennartz, C. M. A. (2008).
723 Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, 28(25):6372–
724 6382.
- 725 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., and Pennartz, C. M. (2009). Hippocampus leads ventral
726 striatum in replay of place-reward information. *PLoS Biology*, 7(8):e1000173.
- 727 Lansink, C. S., Jackson, J. C., Lankelma, J. V., Ito, R., Robbins, T. W., Everitt, B. J., and Pennartz, C. M. A. (2012). Reward
728 Cues in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles. *Journal of
729 Neuroscience*, 32(36):12444–12459.
- 730 Lansink, C. S., Meijer, G. T., Lankelma, J. V., Vinck, M. A., Jackson, J. C., and Pennartz, C. M. A. (2016). Reward Expectancy
731 Strengthens CA1 Theta and Beta Band Synchronization and Hippocampal-Ventral Striatal Coupling. *Journal of Neuroscience*,
732 36(41):10598–10610.
- 733 Lavoie, A. M. and Mizumori, S. J. (1994). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons
734 of rats. *Brain Research*, 638(1-2):157–168.
- 735 Lee, D., Seo, H., and Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of
736 Neuroscience*, 35(1):287–308.
- 737 Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., and Grillon, C. (2014). Generalized anxiety disorder is associated
738 with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75(11):909–915.
- 739 Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and
740 Behavioral Neuroscience*, 9(4):343–364.
- 741 Maia, T. V. and Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature
742 Neuroscience*, 14(2):154–162.

- 743 Malhotra, S., Cross, R. W., Zhang, A., and Van Der Meer, M. A. A. (2015). Ventral striatal gamma oscillations are highly variable
744 from trial to trial, and are dominated by behavioural state, and only weakly influenced by outcome value. *European Journal of*
745 *Neuroscience*, 42(10):2818–2832.
- 746 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., and Schoenbaum, G. (2011). Ventral Striatum and Orbitofrontal Cortex
747 Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, 31(7):2700–2705.
- 748 McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., and Nicola, S. M. (2013). Invigoration of reward seeking by cue and proximity
749 encoding in the nucleus accumbens. *Neuron*, 78(5):910–922.
- 750 Mulder, A. B., Shibata, R., Trullier, O., and Wiener, S. I. (2005). Spatially selective reward site responses in tonically active neurons
751 of the nucleus accumbens in behaving rats. *Experimental Brain Research*, 163(1):32–43.
- 752 Mulder, A. B., Tabuchi, E., and Wiener, S. I. (2004). Neurons in hippocampal afferent zones of rat striatum parse routes into
753 multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7):1923–1932.
- 754 Nicola, S. M. (2004). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a Discrimi-
755 native Stimulus Task. *Journal of Neurophysiology*, 91(4):1840–1865.
- 756 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of
757 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49):16585–16600.
- 758 Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor.
759 *Psychopharmacology*, 191(3):507–520.
- 760 Noonan, M. P., Chau, B. K., Rushworth, M. F., and Fellows, L. K. (2017). Contrasting Effects of Medial and Lateral Orbitofrontal
761 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29):7023–7035.
- 762 Parkinson, J. A., Willoughby, P. J., Robbins, T. W., and Everitt, B. J. (2000). Disconnection of the anterior cingulate cortex
763 and nucleus accumbens core impairs pavlovian approach behavior: Further evidence for limbic cortical-ventral striatopallidal
764 systems. *Behavioral Neuroscience*, 114(1):42–63.
- 765 Paxinos, G. and Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates*. Academic Press, San Diego, 4 edition.
- 766 Pennartz, C. M. A. (2004). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation by
767 Hippocampal Ripples. *Journal of Neuroscience*, 24(29):6446–6456.
- 768 Peters, J., LaLumiere, R. T., and Kalivas, P. W. (2008). Infralimbic Prefrontal Cortex Is Responsible for Inhibiting Cocaine Seeking
769 in Extinguished Rats. *Journal of Neuroscience*, 28(23):6046–6053.
- 770 Rescorla, R. A. and Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and
771 Instrumental Learning. *Psychological Review*, 74(3):151–182.
- 772 Robinson, T. E. and Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues
773 Through the Study of Individual Differences. *Biological Psychiatry*, 65(10):869–873.
- 774 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., and Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the

- 775 Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42):13365–13376.
- 776 Roitman, M. F., Wheeler, R. A., and Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for rewarding and
777 aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4):587–597.
- 778 Saddoris, M. P., Stamatakis, A., and Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nu-
779 cleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*,
780 33(12):2274–2287.
- 781 Salamone, J. D. and Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3):470–485.
- 782 Saunders, B. T. and Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of pavlovian-conditioned
783 responses. *European Journal of Neuroscience*, 36(4):2521–2532.
- 784 Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1):23–32.
- 785 Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.
- 786 Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning.
787 *Neuron*, 38(4):625–636.
- 788 Shidara, M., Aigner, T. G., and Richmond, B. J. (1998). Neuronal signals in the monkey ventral striatum related to progress through
789 a predictable series of trials. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18(7):2613–25.
- 790 Shmuel Lissek, A. N. K. (2013). Generalization of Conditioned Fear and Obsessive-Compulsive Traits. *Journal of Psychology &*
791 *Psychotherapy*, 7:3.
- 792 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., and Buzsáki, G. (2017). Cocaine place conditioning strengthens location-
793 specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, pages 1–10.
- 794 Sleezer, B. J., Castagno, M. D., and Hayden, B. Y. (2016). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.
795 *Journal of Neuroscience*, 36(44):11223–11237.
- 796 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., and Hayden, B. Y. (2016). Neuronal selectivity for spatial
797 positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3):1098–1111.
- 798 Sugam, J. A., Saddoris, M. P., and Carelli, R. M. (2014). Nucleus accumbens neurons track behavioral preferences and reward
799 outcomes during risky decision making. *Biological Psychiatry*, 75(10):807–816.
- 800 Sutton, R. and Barto, A. (1998). Reinforcement Learning: An Introduction.
- 801 Tabuchi, E. T., Mulder, A. B., and Wiener, S. I. (2000). Position and behavioral modulation of synchronization of hippocampal and
802 accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6):717–728.
- 803 Takahashi, Y. K., Langdon, A. J., Niv, Y., and Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled
804 by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1):182–193.
- 805 van der Meer, M. A. A. and Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward
806 Information. *Journal of Neuroscience*, 31(8):2843–2854.

- 807 West, E. A. and Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after
808 Reinforcer Devaluation. *Journal of Neuroscience*, 36(4):1128–1139.
- 809 Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., and Mulder, A. B. (2003). Spatial and behavioral correlates
810 in nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,
811 1250(C):275–292.
- 812 Yun, I. A., Wakabayashi, K. T., Fields, H. L., and Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral
813 and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, 24(12):2923–2933.



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 across different units (A) and in time (B).

A: Displayed are schematic PETHs illustrating putative responses to different cues under different hypotheses of how cue identity and value-outcome are coded.

- H1 (left panel):** Coding of cue identity is absent in the NAc. Top: Unit A encodes a motivationally relevant variable, such as expected outcome, similarly across other cue features, such as cue identity or physical location. Hypothetical plot is firing rate across time. L1+ (red) signifies a reward-available light cue, S1+ (navy blue) a reward-available sound cue, L2- (green) a reward-unavailable light cue, S2- (light blue) a reward-unavailable sound cue. Dashed line indicates onset of cue. Bottom: No units within the NAc discriminate their firing according to cue identity.
- H2 (middle panel):** Coding of cue identity occurs independently of encoding of motivationally relevant variables such as expected outcome or subsequent vigor. Top: Same as H1, with unit A discriminating between reward-available and reward-unavailable cues. Bottom: Unit B discriminates firing across stimulus modalities, depicted here as firing to light cues but not sound cues.
- H3 (right panel):** Coding of cue identity is integrated with coding of other motivationally relevant variables. Hypothetical example demonstrating a unit that responds to reward-predictive-outcome-predictive cues, but firing rate is also modulated by cue identity, firing most for the reward-available light cue.

B: Displayed are schematic PETHs illustrating potential ways in which cue identity signals may persist over time.

- H1 (left panel):** Cue-onset triggers a transient response to a unit that codes for cue identity. Dashed lines indicate time of a behavioral or environmental event. 'Cue-ON' signifies onset of cue, 'NP' signifies when the rat holds a nosepoke at a reward receptacle, 'Out' signifies when the outcome is revealed, 'Cue-OFFOFF' signifies when the cue turns off.
- H2 (middle panel):** Coding of cue identity is maintained persists during a nosepoke hold period until outcome is revealed. Coding can either be maintained by the same unit as during cue-onset (H2A) or by a sequence of units (H2B).
- H3 (right panel):** Coding of cue identity is maintained during entire duration of cue, including persists after the outcome is revealed received when the rat gets feedback about his decision, by either the same unit as during cue-onset (H3A) or by a sequence of units (H3B).

The same hypotheses apply to other information-containing aspects of the environment when the cue is presented, such as the physical location of the cue.

2 - Schematic task.png

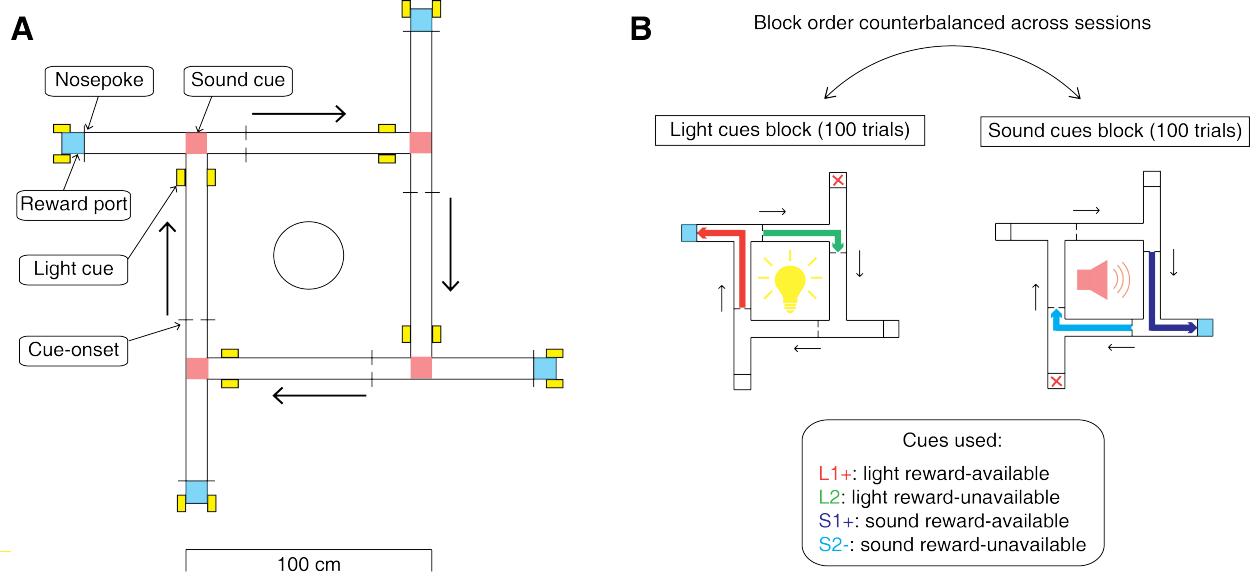


Figure 2: Schematic of behavioral task. **A:** To scale depiction of square track consisting of multiple identical T-choice points. At each choice point, the availability of 12% sucrose reward at the nearest reward receptacle (light blue fill) was signaled by one of four possible cues, presented when the rat initiated a trial by crossing a photobeam on the track (dashed lines). Photobeams at the ends of the arms by the receptacles registered nosepokes (solid lines). Rectangular boxes with yellow fill indicate location of LEDs used for light cues. Speakers for tone cues were placed underneath the choice points, indicated by magenta fill on track. Light blue fill on ends of arms track indicate location of receptacles where sucrose was dispensed on rewarded trials correct running direction. Circle in the center indicates location of pedestal during pre- and post-records. Scale bar is located beneath the track. **B:** Progression of a recording session. A session was started with a 5 minute recording period on a pedestal placed in the center of the apparatus. Rats then performed two blocks 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 followed by another 5 minute recording period 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 (approach trial; red) and reward-unavailable (skip trial; green) trial. Right in figure depicts a sound block, with a reward-available (approach trial; navy blue) and reward-unavailable (skip trial; light blue) trial. Ordering of the light and sound blocks was counterbalanced across sessions. Dashed line within track indicates photobeam. 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. Solid line within track indicates photobeam that registered onset. Receptacles containing sucrose solution are present within each arm. Location of the squares at cue on the end track was irrelevant for behavior, all cue locations contained an equal amount of each arm. Light blue fill signifies reward receipt on a correct reward-available trial and reward-unavailable trials. Arrows outside of track indicate correct running direction. A session ended with another 5 minute recording period on the pedestal.

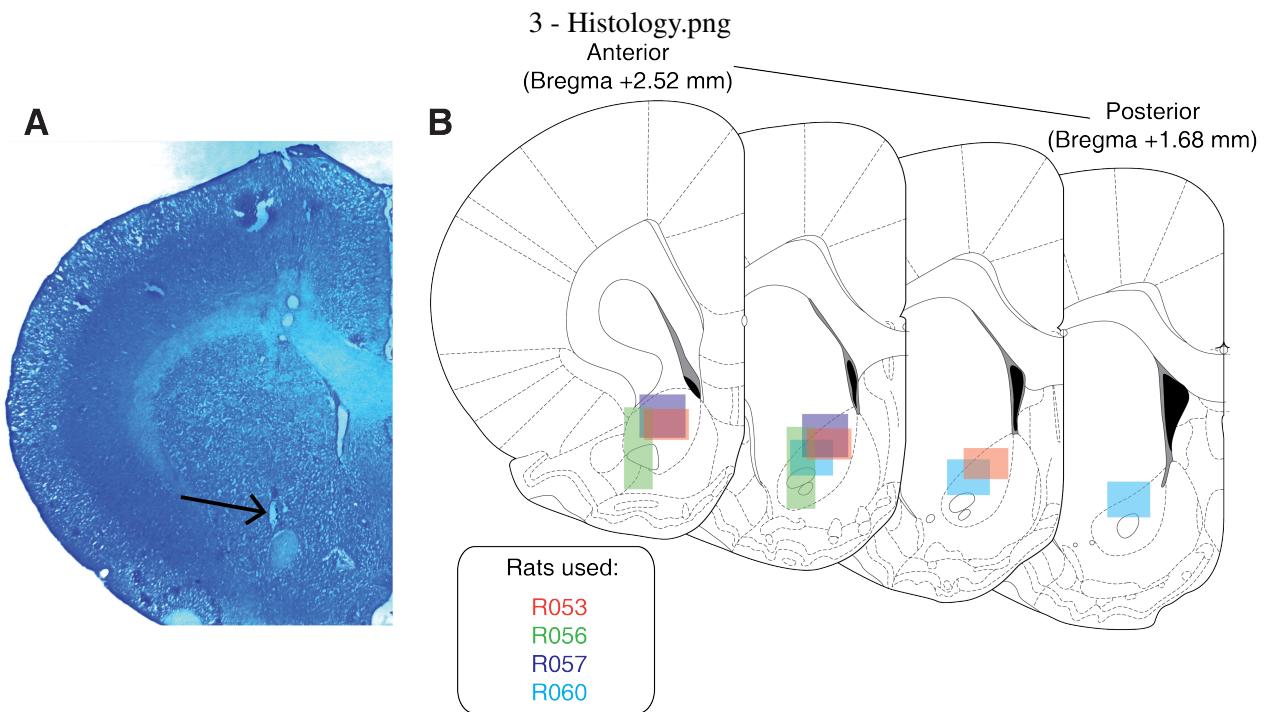


Figure 3: Histological verification of recording sites. Upon completion of experiments, brains were sectioned and tetrode placement was confirmed. **A:** Example section from R060 showing a recording site in the NAc core just dorsal to the anterior commissure ([white circle arrow](#)). **B:** Schematic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

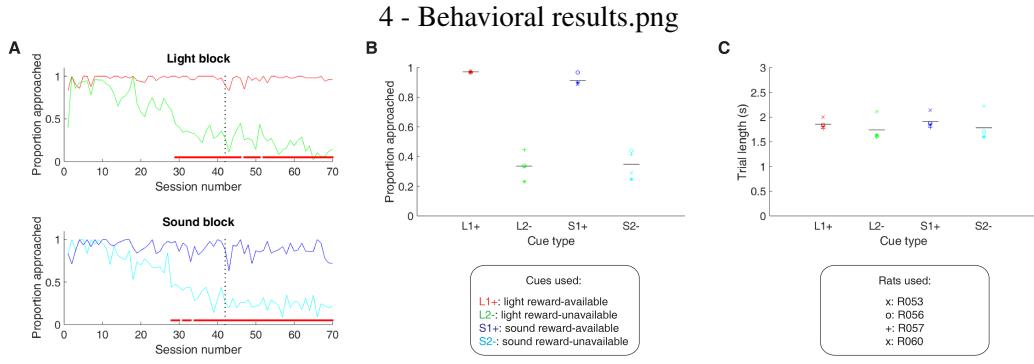


Figure 4: Performance on the behavioral task. **A-BA**: Example learning curves across sessions from a single subject (R060) showing the proportion approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light block, light blue line for sound block) **for light (top) and sound (bottom) blocks**. Fully correct performance corresponds to an approach proportion of 1 for reward-available trials and 0 for reward-unavailable trials. Rats initially approach on both reward-available and reward-unavailable trials, and learn with experience to skip non-rewarded trials. Red bars indicate days in which a rat statistically discriminated between reward-available and reward-unavailable cues, determined by a chi square test. Dashed line indicates time of electrode implant surgery. **C-B-C**: Summary of performance during recording sessions for each rat. **CB**: 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). **DC**: Average trial length for each cue. Note that the time to complete a trial was comparable for the different cues.

5 - Neural examples.png

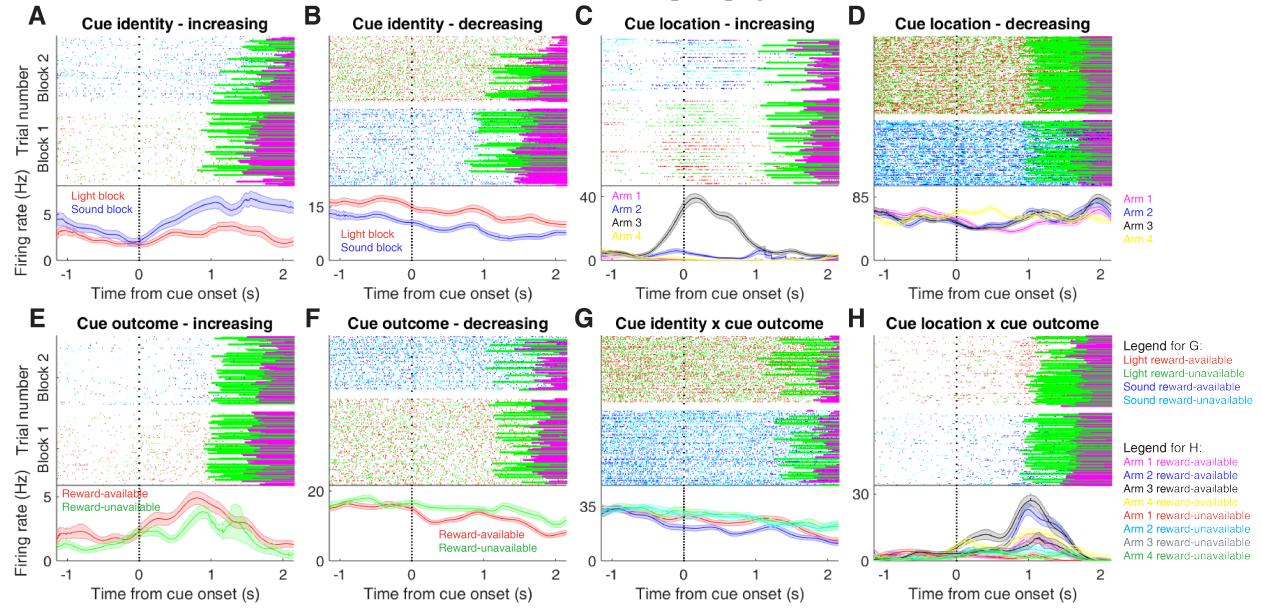


Figure 5: Examples of different cue-modulated NAc units influenced by various task parameters. **A:** Example of a cue-modulated NAc unit that showed an increase in firing following the cue, and encoded cue identity. Top: rasterplot showing the spiking activity across all trials aligned to cue-onset.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

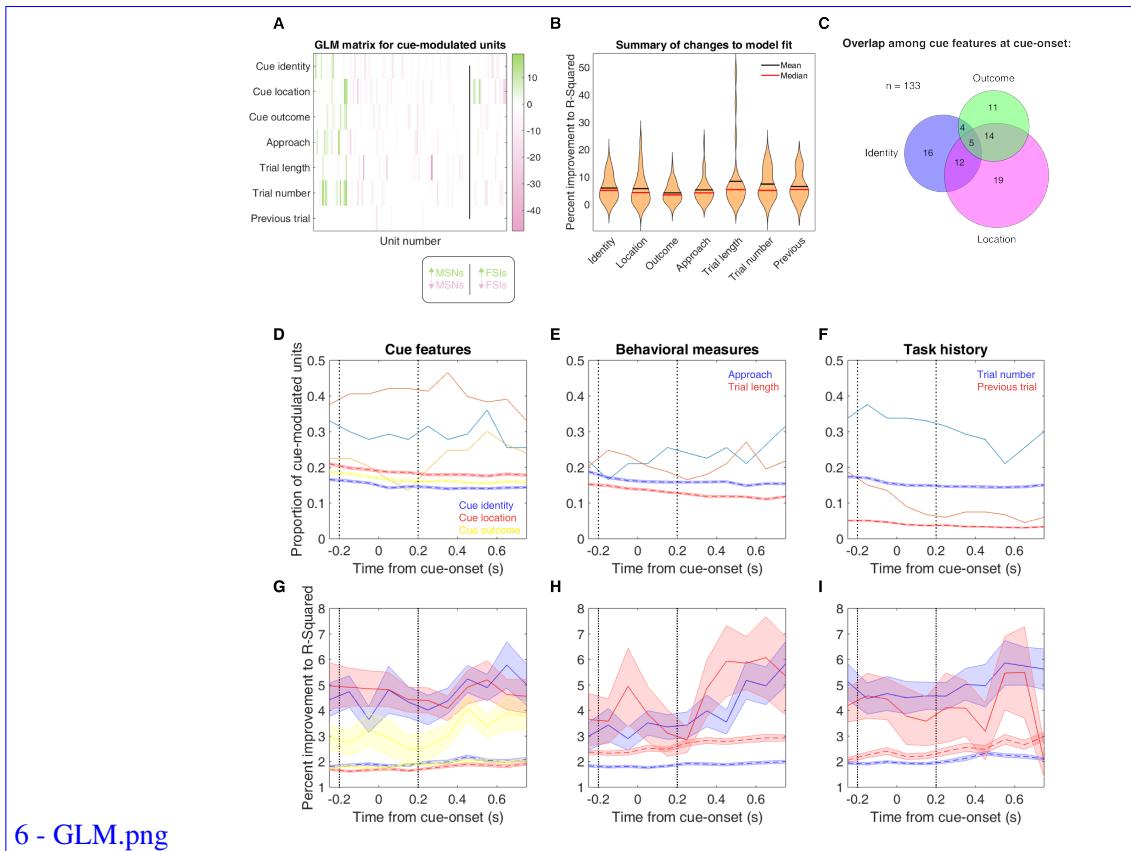


Figure 6: Summary of influence of various task parameters on cue-modulated NAc units after cue-onset. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank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

7 - Population averages.png

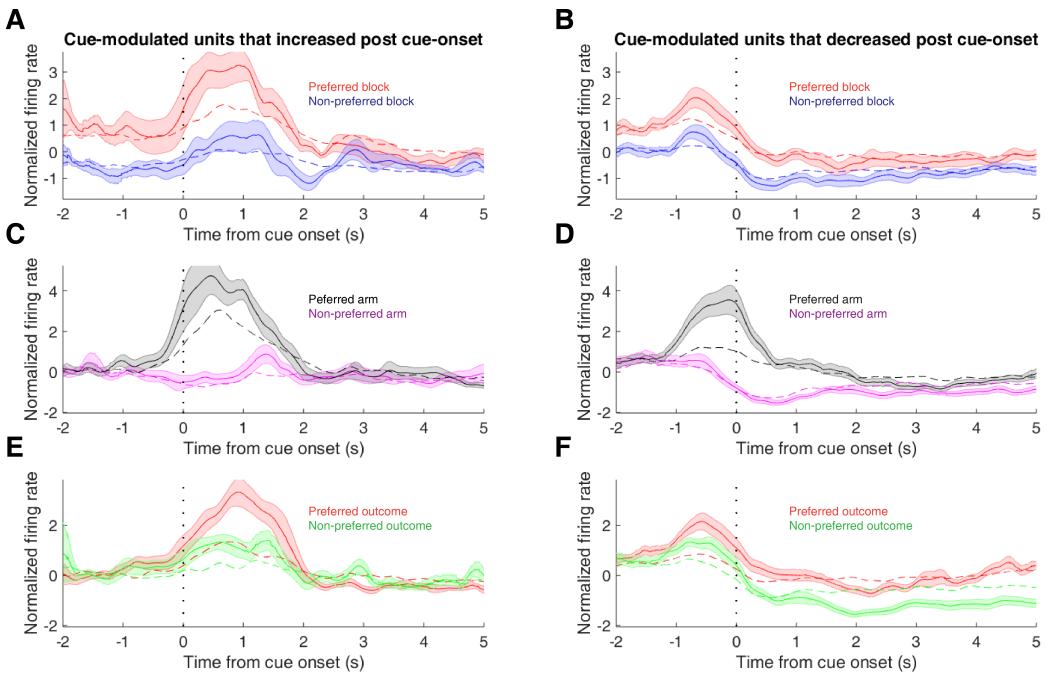
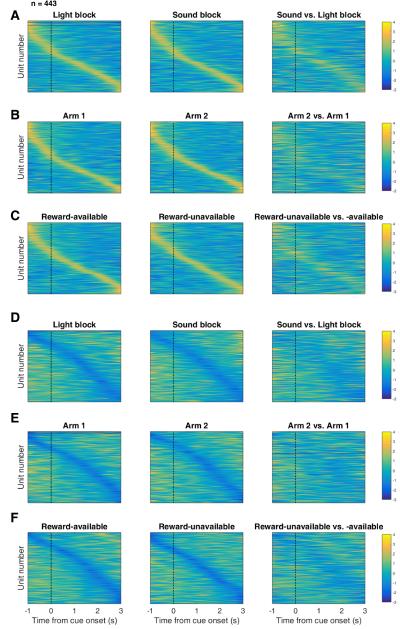


Figure 7: Population-level averages of cue feature sensitive NAc units. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to cue-onset. Activity is plotted for preferred stimulus block (red) and non-preferred stimulus block (blue). Dashed vertical line indicates onset of cue. **B:** Same as A but for units that decreased in firing. Note population level activity reveals cells-units classified as decreasing in response to cue show a biphasic response at the population level, with a transient increase around the time the rat starts on the arm, followed by a minimum after cue onset. Also, note the sustained difference in firing between the two blocks. **C-D:** Same as A-B for cue location. Activity is plotted from-for most preferred arm (yellow), in-decreasing order to-and 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 cellunits 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 exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text). **B:** Same layout as in A, except that the panels now compare two different locations on the track instead of two cue modalities. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C:** Same layout as in A, except that panels now compare reward-available and reward-unavailable trials. **D-F:** Heat plots aligned to normalized minimum firing rates. **D**: Responses during different stimulus blocks as in A, but with units ordered according to the time of their minimum firing rate. **E**: Responses during trials on different arms as in B, but with units ordered by their minimum firing rate. **F**: Responses during cues signalling different outcomes as in C, but with units ordered by their minimum firing rate. Overall, NAc units "tiled" experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

9 - NP Neural examples.png

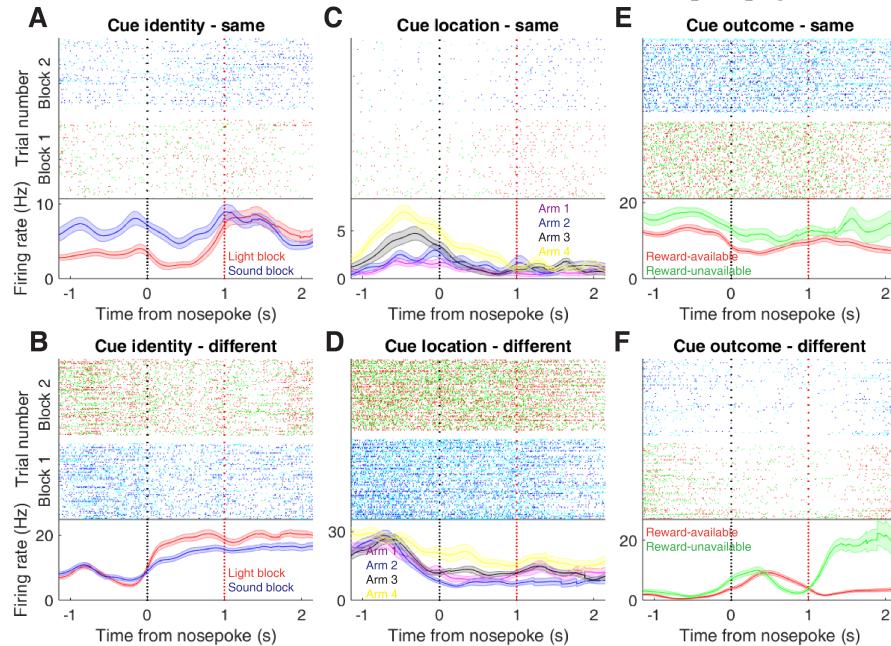


Figure 9: Examples of cue-modulated NAc units influenced by various task parameters at time of nosepoke. **A:** Example of a cue-modulated NAc unit that encoded cue identity at both cue-onset and during nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to nosepoke.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

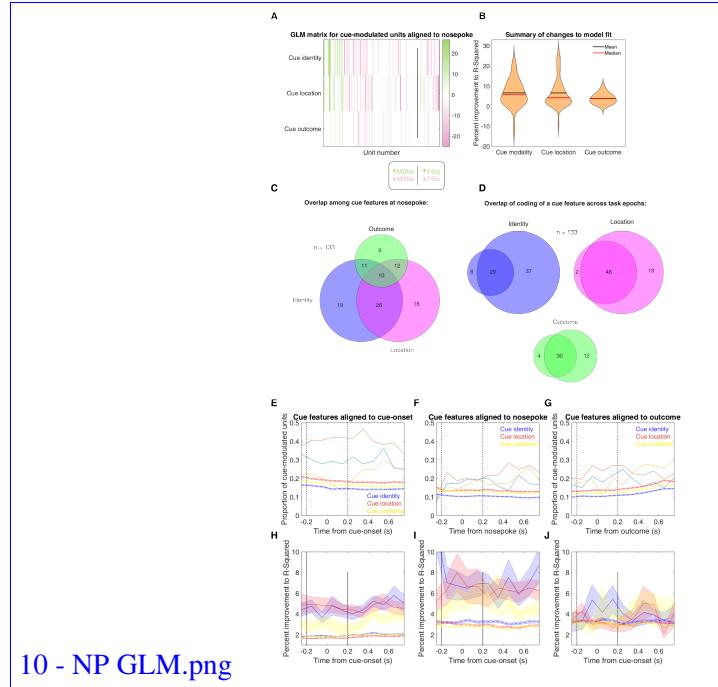


Figure 10: Summary of influence of various task parameters of cue-modulated NAc units during nosepoke. **A:** GLM matrix demonstrating impact illustrating the contribution of various task parameters on NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by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

11 - NP population averages.png

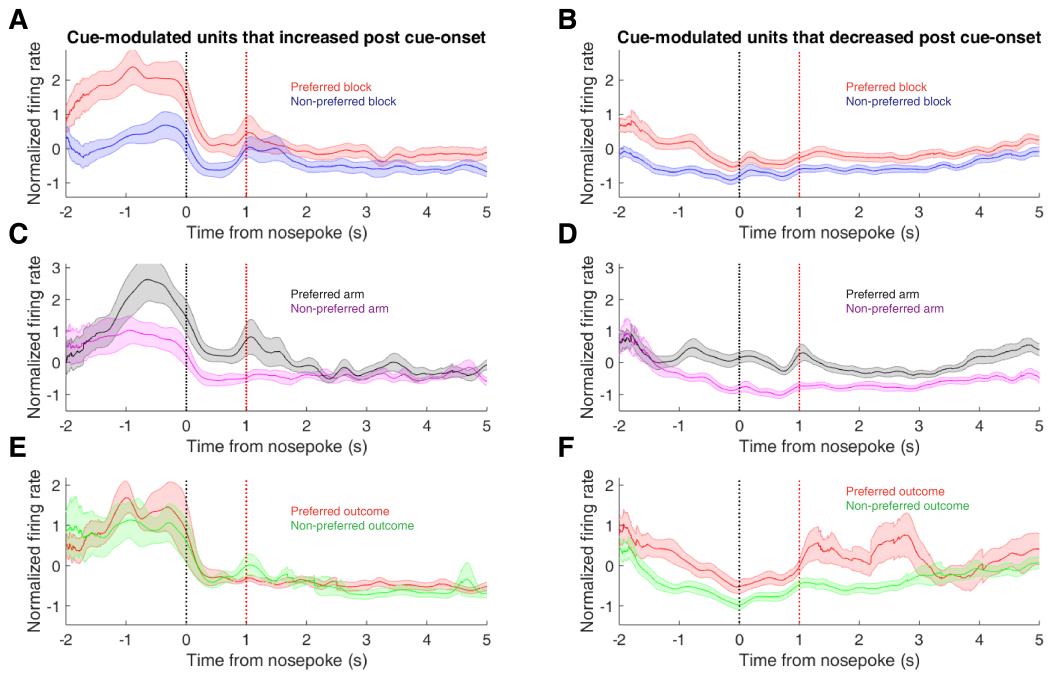
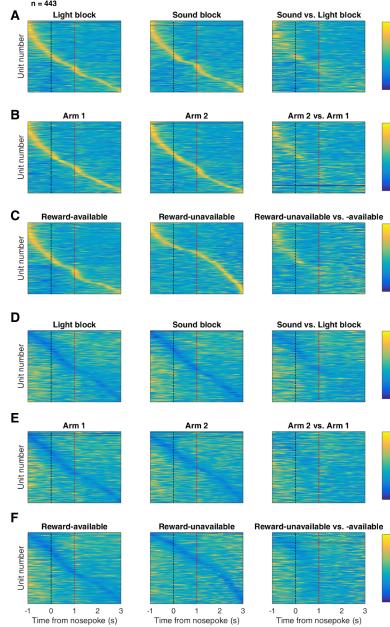


Figure 11: Population-level averages of cue feature sensitive NAc units during a nosepoke. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to nosepoke with reward delivery occurring 1 s after nosepoke. Activity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



12 - NP task tiling.png

Figure 12: Distribution of NAc firing rates across ~~task space during time surrounding nosepoke for approach trials~~. Each panel shows normalized ~~(z-score)~~ firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates nosepoke), averaged across all approach trials for a specific cue type, indicated by text labels. **A-C:** ~~Heat plots aligned to normalized peak firing rates~~. **A**, far left: Heat plot showing smoothed normalized firing activity of all recorded NAc units ordered according to the time of their peak firing rate during the light block. Each row is a units average activity across time to the light block. Black dashed line indicates nosepoke. Red dashed line indicates reward delivery occurring 1 s after nosepoke for reward-available trials. Notice the yellow band across time, indicating all aspects of visualized task space were captured by the peak firing rates of various units. **A, middle 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 outcome~~ 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 exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text).~~ **B:** Same layout as in A, ~~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-C:** ~~As for C, but with units ordered by their minimum firing rate~~. **E:** Same layout as in A, except that panels now compare ~~correct~~ reward-available and ~~incorrect~~ reward-unavailable trials. ~~Notice the absence of a cluster of The disproportionate tiling around reward outcome receipt for reward-available, but not reward-unavailable trials suggests encoding of reward receipt by NAc units.~~ **F-D-F:** ~~As for 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

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

Table 1: Cells-Units overview