

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: 2150

Introduction Word Count: ?

Discussion Word Count: ?

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

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

1 Abstract(eLife: 150 words)

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

12 Introduction

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

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

35 Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues

36 (Atallah, McCool, Howe, & Graybiel, 2014; Bissonette et al., 2013; Cooch et al., 2015; Day, Wheeler, Roitman, & Carelli, 2012; Hollerman et al., 1998).
37 (Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Goldstein et al., 2012; Hollerman et al., 1998).
38 , much less is known about how outcome-predictive cues themselves are encoded in the NAc (Sleezer, Castagno, & Hayden, 2016;
39 (but see: Sleezer et al., 2016). This is an important issue for at least two reasons. First, in reinforcement
40 learning, motivationally relevant outcomes are typically temporally delayed relative to the cues that predict
41 them. In order to solve the problem of assigning credit (or blame) across such temporal gaps, some trace of
42 preceding activity needs to be maintained (Lee et al., 2012; Sutton & Barto, 1998)(Lee et al., 2012; Sutton & Barto, 1998)
43 . For example, if you become ill after eating food X in restaurant A, depending on if you remember
44 the identity of the restaurant or the food at the time of illness, you may learn to avoid all restaurants,
45 restaurant A only, food X only, or the specific pairing of X-in-A. Therefore, a complete understanding
46 of what is learned following feedback requires understanding what trace is maintained. Since NAc is a
47 primary target of DA signals interpretable as reward prediction errors (RPEs)RPEs, and NAc lesions im-
48 pair RPEs related to timing, its activity trace will help determine what can be learned when RPEs arrive
49 (Hamid et al., 2015; Hart, Rutledge, Glimcher, & Phillips, 2014; Ikemoto, 2007; McDannald, Lucantonio, Burke, Niv, & Sealfon,
50 (Hamid et al., 2015; Hart et al., 2014; Ikemoto, 2007; McDannald et al., 2011; Takahashi et al., 2016).

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

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

68 [Figure 1 about here.]

69 **Methods**

70 **Subjects:**

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

77 **Overall timeline:**

78 Each rat was first handled for seven days during which they were exposed to the experiment room, the
79 sucrose solution used as a reinforcer, and the click of the sucrose dispenser valves. Rats were then trained on

80 the behavioral task (described in the next section) until they reached performance criterion. At this point they
81 underwent hyperdrive implantation targeted at the NAc. Rats were allowed to recover for a minimum of five
82 days before being retrained on the task, and recording began once performance returned to pre-surgery levels.
83 Upon completion of recording, animals were gliosed, euthanized and recording sites were histologically
84 confirmed.

85 Behavioral task and training:

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

100 Each session consisted of both a light block and a sound block with 100 trials each. Within a block, one cue
101 signaled reward was available on that trial (L1+ or S1+), while the other signaled reward was not available
102 (L2- or S2-). Light block cues were a flashing white light, and a constant yellow light. Sound block cues
103 were a 2 kHz sine wave and a 8 kHz sine wave whose amplitude was modulated from 0 to maximum by a 2

104 Hz sine wave. Outcome-cue associations were counterbalanced across rats, e.g. for some rats L1+ was the
105 flashing white light, and for others L1+ was the constant yellow light. The order of cue presentation was
106 pseudorandomized so that the same cue could not be presented more than twice in a row. Block order within
107 each day was also pseudorandomized, such that the rat could not begin a session with the same block for more
108 than two days in a row. Each session consisted of a 5 minute pre-session period on a pedestal (a terracotta
109 planter filled with towels), followed by the first block, then the second block, then a 5 minute post-session
110 period on the pedestal. For approximately the first week of training, rats were restricted to running in
111 the clockwise direction by presenting a physical barrier to running counterclockwise. Cues signaling the
112 availability and unavailability of reward, as described above, were present from the start of training. Rats
113 were trained for 200 trials per day (100 trials per block) until they discriminated between the reward-available
114 and reward-unavailable cues for both light and sound blocks for three consecutive days, according to a
115 chi-square test rejecting the null hypothesis of equal approaches for reward-available and reward-unavailable
116 trials, at which point they underwent electrode implant surgery.

117 Schematic of behavioral task. **A:** Top scale depiction of square track consisting of multiple identical T-choice
118 points. At each choice point, the availability of 12% sucrose reward at the nearest reward receptacle (light
119 blue fill) was signaled by one of four possible cues, presented when the rat initiated a trial by crossing a
120 photobeam on the track (dashed lines). Photobeams at the ends of the arms by the receptacles registered
121 Nosepokes (solid lines). Rectangular boxes with yellow fill indicate location of LEDs used for light cues.
122 Speakers for tone cues were placed underneath the choice points, indicated by magenta fill on track. Arrows
123 outside of track indicate correct running direction. Circle in the center indicates location of pedestal during
124 pre- and post-records. Scale bar is located beneath the track. **B:** Progression of a recording session. A
125 session was started with a 5 minute recording period on a pedestal placed in the center of the apparatus. Rats
126 then performed the light and sound blocks of the cue discrimination task in succession for 100 trials each,
127 followed by another 5 minute recording period on the pedestal. Left in figure depicts a light block, showing
128 an example trajectory for a correct reward-available (approach trial; red) and reward-unavailable (skip trial;
129 green) trial. Right in figure depicts a sound block, with a reward-available (approach trial; navy blue) and

130 reward-unavailable (skip trial; light blue) trial. Ordering of the light and sound blocks was counterbalanced
131 across sessions. Reward-available and reward-unavailable cues were presented pseudo-randomly, such that
132 not more than two of the same type of cue could be presented in a row. Location of the cue on the
133 track was irrelevant for behavior, all cue locations contained an equal amount of reward-available and
134 reward-unavailable trials.

135 **Surgery:**

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

145 **Data acquisition and preprocessing:**

146 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with
147 a RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex).
148 Signals were referenced against a tetrode placed in the corpus callosum above the NAc. Candidate spikes
149 for sorting into putative single units were obtained by band-pass filtering the data between 600-9000 Hz,
150 thresholding and aligning the peaks (UltraMegaSort2k, Hill, Mehta, & Kleinfeld, 2011). Spike waveforms
151 were then clustered with KlustaKwik using energy and the first derivative of energy as features, and manually
152 sorted into units (MClust 3.5, A.D. Redish et al., <http://redishlab.neuroscience.umn.edu/MClust/MClust.html>).

153 Isolated units containing a minimum of 200 spikes within a session were included for subsequent analysis.
154 Units were classified as fast spiking interneurons (FSIs) by an absence of interspike intervals (ISIs) > 2 s,
155 while medium spiny neurons (MSNs) had a combination of ISIs > 2 s and phasic activity with shorter ISIs
156 (Atallah et al., 2014; Barnes, Kubota, Hu, Jin, & Graybiel, 2005).

157 Data analysis:

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

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

172 To determine the relative contribution of different task parameters to firing rate variance (as in Figures 4–5),
173 a forward selection stepwise general linear model (GLM) was fit to each cue-modulated unit. Cue identity
174 (light block, sound block), cue location (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available,
175 reward-unavailable), behavior (approach, skip), trial length, trial number, and trial history (reward availability

176 on the previous 2 trials) were used as predictors, and the 1 s post-cue firing rate as the response variable.
177 Units were classified as being modulated by a given task parameter if addition of the parameter significantly
178 improved model fit using deviance as the criterion ($p < .01$). A comparison of the R-squared value between
179 the final model and the final model minus the predictor of interest was used to determine the amount of
180 firing rate variance explained by the addition of that predictor for a given unit. To investigate more finely the
181 temporal dynamics of the influence of task parameters to unit activity, we then fit a sliding window GLM
182 with the same task parameters using 500 ms bins and 100 ms steps, starting 500 ms before cue-onset, up to
183 500 ms after cue-onset, and measured the proportion of units and average R-squared value for a given time
184 bin where a particular predictor contributed significantly to the final model. To control for amount of units
185 that would be affected by a predictor by chance, we shuffled the order of firing rates used for a time bin, and
186 took the average of this value over 100 shuffles.

187 To better visualize responses to cues and enable subsequent population level analyses (as in Figures 4, 6,
188 and 7), spike trains were convolved with a Gaussian kernel ($\sigma = 100$ ms), and peri-event time histograms
189 (PETHs) were generated by taking the average of the convolved spike trains across all trials for a given
190 task condition. For analysis of population level responses for cue features (Figure 6), convolved spike trains
191 for all units where cue identity, cue location, or cue outcome explained a significant portion of firing rate
192 variance were z-scored. Within a given cue feature, normalized spike trains were then separated according
193 to the preferred and non-preferred cue condition (e.g. light vs. sound block), and averaged across units to
194 generate population-level averages. To account for separation that would result from any random selection
195 of units, unit identity was shuffled and the shuffled average for preferred and non-preferred cue conditions
196 was generated for 1000 shuffles.

197 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units
198 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure
199 7). To compare representations of task space across cue conditions for a cue feature, the ordering of units
200 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other

201 condition (e.g. sound block). For control comparisons within cue conditions, half of the trials for a condition
202 were compared against the other half. To look at the correlation of firing rates of all units within and across
203 various cue conditions, trials for each cue condition for a unit were shuffled and divided into two averages,
204 and averages within and across cue conditions were correlated. A linear mixed effects model was run for
205 each cue condition to determine if correlations of firing rates within cue conditions were more similar than
206 correlations across cue conditions.

207 To identify the responsivity of units to different cue features at the time of a nosepoke into a reward
208 receptacle, and subsequent reward delivery, the same cue-responsive units from the cue-onset analyses were
209 analyzed at the time of nosepoke and outcome receipt using identical analysis techniques (Figures 8, 9, 10,
210 and 11).

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

218 Histology:

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

224 electrodes with recording locations in the NAc were analyzed (Figure 12).-

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

228 Results

229 Behavior

230 Rats were trained to discriminate between cues signaling the availability and absence of reward on a square
231 track with four identical arms for two distinct set of cues (Figure 2). During each session, rats were pre-
232 sented sequentially with two behavioral blocks containing cues from different sensory modalities, a light and
233 a sound block, with each block containing a cue that signalled the availability of reward (reward-available),
234 and a cue that signalled the absence of reward (reward-unavailable). To maximize reward receipt, rats should
235 approach reward sites on reward-available trials, and skip reward sites on reward-unavailable trials (see Fig-
236 ure 3A for an example learning curve). All four rats learned to discriminate between the reward-available and
237 reward-unavailable cues for both the light and sound blocks as determined by reaching significance ($p < .05$)
238 on a daily chi-square test comparing approach behavior for reward-available and reward-unavailable cues for
239 each block, for at least three consecutive days (range for time to criterion: 22 - 57 days). Maintenance of
240 behavioral performance during recording sessions was assessed using linear mixed effects models for both
241 proportion of trials where the rat approached the receptacle, and trial length. Analyses revealed that the like-
242 lihood of a rat to make an approach was influenced by whether a reward-available or reward-unavailable cue
243 was presented, but was not significantly modulated by whether the rat was presented with a light or sound
244 cue (Percentage approached: light reward-available = 97%; light reward-unavailable = 34%; sound reward-

245 available = 91%; sound reward-unavailable 35%; cue identity $p = .115$; cue outcome $p < .001$; Figure 3B). A
246 similar trend was seen with the length of time taken to complete a trial (Trial length: light reward-available
247 = 1.85 s; light reward-unavailable = 1.74 s; sound reward-available = 1.91 s; sound reward-unavailable 1.78
248 s; cue identity $p = \underline{.445}$,.106; cue outcome $p < .001$; Figure 3C). Thus, during recording, rats successfully
249 discriminated the cues according to whether or not they signaled the availability of reward at the reward
250 receptacle.

251 [Figure 2 about here.]

252 [Figure 3 about here.]

253 **NAc units encode behaviorally relevant and irrelevant cue features**

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

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

266 ramping of activity starting at cue-onset, and showed elevated activity immediately preceding cue-onset, for
267 example (Figure 4). To characterize more formally whether these cue-evoked responses were modulated by
268 various aspects of the task, we fit a GLM to each cue-modulated unit. Fitting GLMs revealed that a variety
269 of task parameters accounted for a significant portion of firing rate variance in NAc cue-modulated units
270 (Figure 5, Table 1). Notably, there were units that discriminated between whether the rat was performing
271 in the light or sound block (28% of cue-modulated units, accounting for 6% of variance on average), which
272 arm the rat was currently on (38% of cue-modulated units, accounting for 6% of variance on average), and
273 whether the rat was engaged in the common portion of a reward-available or reward-unavailable trial (26%
274 of cue-modulated units, accounting for 4% of variance on average), suggesting that the NAc encodes features
275 of reward-predictive cues separate from expected outcome (Figure 4A-F). Furthermore, overlap of coding of
276 cue features within units was not different than expected by chance according to chi-square tests, suggesting
277 for integrated coding across various aspects of a cue (Figure 4G,H). Additionally, a sliding window GLM
278 centered on cue-onset revealed that cue identity and cue location contributed to the activity of a significant
279 proportion of cue-modulated units throughout this epoch, whereas an increase in units encoding cue outcome
280 became apparent after cue-onset (Figure 5D). Together, these findings show that various cue features are
281 represented in the NAc, and that this coding is both integrated and separate from expected outcome (Figure
282 1; H2,H3).

283 [Table 1 about here.]

284 [Figure 4 about here.]

285 [Figure 5 about here.]

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

287 We observed a variety of single unit response profiles around the time of cue onset (Figure 4). To investigate

whether these firing rate patterns were related to what cue features were encoded, we plotted the population level averages for units that were modulated by each feature. To do this, we normalized firing activity for each unit that was modulated by a given cue feature, such as light block, then generated the cue-onset aligned population average firing rate for each of the cue features (Figure 6). Overall, this analysis revealed that cells that showed an increase upon cue presentation had stronger responses for the preferred cue condition (Figure 6A,C,E). Interestingly, units that were classified as decreasing in response to the cue showed a biphasic response at the population level, with a small peak at a time in alignment with entry into the arm, followed by a sustained dip after cue-onset (Figure 6B,D,F). Units that were modulated by cue identity showed a stronger increase in response to the preferred task block, as well as a higher tonic firing rate to the preferred task block, most notably in units that decreased in firing rate to the cue (Figure 6A,B). Units that were modulated by cue location showed a graded response to locations of decreasing preference, with peak firing occurring around cue-onset (Figure 6C,D). Units that were modulated by cue outcome showed a ramping of activity after cue-onset for their preferred cue type. Additionally, units that exhibited a decrease in firing in response to the cue and whose activity was modulated by cue outcome, showed a sustained discriminatory response to reward-available and reward-unavailable cues that extended beyond cue-onset (Figure 6F). Together, these visualizations of the averaged population responses revealed nuanced differences in the way NAc units are modulated by cue conditions across cue features.

[Figure 6 about here.]

NAc units dynamically segment the task:

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

them according to the time of peak firing rate (Figure 7). We did this separately for both the light and sound blocks, and found a nearly uniform distribution of firing fields in task space that was not limited to alignment to the cue (Figure 7A). Furthermore, to determine if this population level activity was similar across blocks, we also organized firing during the sound blocks according to the ordering derived from the light **Blocks**. This revealed that while there was some preservation of order, the overall firing was qualitatively different across the two blocks, implying that population activity distinguishes between light and sound blocks. To control for the possibility that any comparison of trials would produce this effect, we did a within block comparison, comparing half of the trials in the light block against the other half. This comparison looked similar to our test comparison of sound block trials ordered by light block trials. Additionally, given that the majority of our units showed an inhibitory response to the cue, we also plotted the firing rates according to the lowest time in firing, and again found some maintenance of order, but largely different ordering across the two blocks, and the within block comparison (Figure 7BD). To further test this, we divided each block into two halves and looked at the correlation of the average smoothed firing rates across various combinations of these halves across our cue-aligned centered epoch. A linear mixed effects model revealed that within block correlations (e.g. one half of light trials vs other half of light trials) were higher and more similar than across block correlations (e.g. half of light trials vs half of sound trials) suggesting that activity in the NAc discriminates across various cue conditions (within block correlations = .383 (light), .379 (sound); across block correlations = .343, .338, .337, .348; within block vs.~within block comparison = $p = .934$; within block vs.~across block comparisons = $p < .001$). This process was repeated for cue location (Figure 7C-DB,E; 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 7E-F_{CQ}F; within block correlations = .429 (reward-available), .261 (reward-unavailable); across block correlations = .258, .253, .255, .249; within block vs.~within block comparison = $p < .001$; within block vs.~across block comparisons = $p < .001$), showing that NAc segmentation of the task is qualitatively different even during those parts of the task not immediately associated with a specific cue, action, or outcome, although the within condition comparison of reward-unavailable trials was less correlated than reward-available trials, and more similar to the across condition comparisons, potentially due to the greater behavioral variability for the reward-unavailable trials.

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

342 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
344 cue. To test whether representations of cue features persisted post-approach until the outcome was revealed,
345 we fit a GLM to the post-approach firing rates of cue-modulated units aligned to the time of nosepoke into
346 the reward receptacle. This analysis showed that a variety of units still discriminated firing according to
347 various cue features, but not other task parameters, showing that NAc activity discriminates various cue
348 conditions well into a trial (Table 1, Figures 8,9). Additionally, these units were a mix between most of the
349 units that encoded cue features at cue-onset (observed overlap greater than expected by chance according
350 to chi-square tests), and those that did not previously have a cue feature as a predictor (29, 48, and 30
351 out of 133 cue-modulated units encoded both time points for cue identity, cue location, and cue outcome,
352 respectively). Population level averages for units that increased to cue-onset showed a ramping up of activity
353 that peaked upon nosepoke, whereas units that decreased to cue-onset showed a gradual reduction of firing
354 activity that reached a minimum upon nosepoke (Figure 10). Additionally, a peak is seen for preferred
355 cue outcome in decreasing units at 1 second post cue-onset when reward was received, demonstrating an
356 integration of expected and received reward (Figure 10F). Furthermore, aligning normalized peak firing rates
357 to nosepoke onset, revealed a clustering of responses around outcome receipt for all cue conditions where the
358 rat would have received reward (Figure 11)–
~~, in addition to the same trend of higher within- vs across-block~~
359 correlations for cue identity (Figure 11A,C; within block correlations = .560 (light), .541 (sound); across
360 block correlations = .487, .481, .483, .486; within block vs. within block comparison = p = .112; within
361 block vs. across block comparisons = p < .001) and cue location (Figure 11B,E; within block correlations
362 = .474 (arm 1), .461 (arm 2); across block correlations = .416, .402, .416, .415; within block vs. within
363 block comparison = p = .810; within block vs. across block comparisons = p < .001), but not cue outcome
364 (Figure 11C,F; within block correlations = .620 (reward-available), .401 (reward-unavailable); across block

365 correlations = .418, .414, .390, .408; within block vs. within block comparison = $p < .001$; within block
366 vs. across block comparisons = $p < .001$). To determine whether coding of cue features persisted after
367 the outcome was revealed, a GLM was fit to the firing rates of cue-modulated units at the time of outcome
368 receipt, during which the cue was still present. Fitting a GLM revealed 10 units (8%) where cue outcome
369 accounted for an average of 32% of firing rate variance (Table 1, data not shown). An absence of cue identity
370 or cue location coding at this level of analysis was observed, but looking at the data more closely with a
371 sliding window GLM revealed that cue identity, cue location, and cue outcome were encoded throughout
372 time epochs surround cue-onset, nosepoke hold, and outcome receipt, suggesting that the NAc maintains a
373 representation of these cue features once the rat receives behavioral feedback for its decision (Figure 9E-J).

374 [Figure 8 about here.]

375 [Figure 9 about here.]

376 [Figure 10 about here.]

377 [Figure 11 about here.]

378 Discussion

379 The main result of the present study is that NAc units encode not only the expected outcome of outcome-
380 predictive cues, but also the identity of such cues. Importantly, this identity coding was maintained on
381 approach trials both during a delay period where the rat held a nosepoke until the outcome was received,
382 after immediately after outcome receipt (H2 and H3 in Figure 1B). Units coding for cue identity showed
383 partial overlap with those coding for expected outcome (H3 in Figure 1A). Units that coded different cue

384 features (identity, outcome, location) exhibited different temporal profiles as a whole, although across all
385 recorded units a tiling of task structure was observed such that all points within our analyzed task space was
386 accounted for by the ordered peak firing rates of all units. Furthermore, this tiling differed between various
387 conditions with a cue feature, such as light versus sound blocks. We discuss these observations and their
388 implications below.

389 **Cue identity:**

390 Our finding that NAc units can discriminate between different outcome-predictive stimuli with similar moti-
391 vational significance (i.e. encodes cue identity) expands upon an extensive rodent literature examining NAc
392 correlates of conditioned stimuli ([Ambroggi, Ishikawa, Fields, & Nicola, 2008](#); [Atallah et al., 2014](#); [Bissonette et al., 2013](#); [Cooch et al., 2015](#); [Day et al., 2006](#); [Dejean et al., 2017](#)).
393 [\(Ambroggi et al., 2008; Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Dejean et al., 2017\)](#). Perhaps the most comparable work in rodents comes from a study that found distinct coding for an odor
394 when it predicted separate but equally valued rewards (Cooch et al., 2015). The present work is complemen-
395 tary to such *outcome identity* coding as it shows that NAc units *cue identity*, both separately and integrated
396 with the reward it predicts (H2 and H3 in Figure 1A). Similarly, Setlow et al. 2003 paired distinct cues
397 with appetitive or aversive outcomes, and found separate populations of units that encoded each cue. Once
398 again, our study was different in asking how distinct cues encoding the same anticipated outcome are en-
399 coded. Such cue identity encoding suggests that even when the biological relevance of these stimuli is
400 similar, NAc dissociates their representations at the level of the single-units. A possible interpretation of
401 this coding of cue features alongside expected outcome is that these representations are used to associate
402 reward with relevant features of the environment, so-called credit assignment in the reinforcement learning
403 literature (Sutton & Barto, 1998). A burgeoning body of human and non-human primate work has started to
404 elucidated neural correlates of credit assignment in the PFC, particularly in the lateral orbitofrontal cortex
405 ([Akaishi, Kolling, Brown, & Rushworth, 2016](#); [Asaad, Lauro, Perge, & Eskandar, 2017](#); [Chau et al., 2015](#); [Noonan, Chau, R](#)).
406 [\(Akaishi et al., 2016; Asaad et al., 2017; Chau et al., 2015; Noonan et al., 2017\)](#). Given the importance of
407 cortical inputs in NAc associative representations, it is possible that information related to credit assignment
408

409 is relayed from the cortex to NAc ([Cooch et al., 2015; Ishikawa et al., 2008](#))[\(Cooch et al., 2015; Ishikawa et al., 2008\)](#)

410 .

411 A different possible function for cue identity coding is to support contextual modulation of the motivational
412 relevance of specific cues. A context can be understood as a particular mapping between specific cues and
413 their outcomes: for instance, in context 1 cue A but not cue B is rewarded, whereas in context 2 cue B but not
414 cue A is rewarded. Successfully implementing such contextual mappings requires representation of the cue
415 identities. Indeed, ([Sleazeer et al., 2016](#))[Sleazeer et al. 2016](#) recorded NAc responses during the Wisconsin
416 Card Sorting Task (WCST), a common set-shifting task used in both the laboratory and clinic, and found
417 units that preferred firing to stimuli when a certain rule, or rule category was currently active. Further support
418 for a modulation of NAc responses by strategy comes from an fMRI study that examined BOLD levels during
419 a set-shifting task (FitzGerald et al., 2014). In this task, participants learned two sets of stimulus-outcome
420 contingencies, a visual set and auditory set. During testing they were presented with both simultaneously,
421 and the stimulus dimension that was relevant was periodically shifted between the two. Here, they found
422 that bilateral NAc activity reflected value representations for the currently relevant stimulus dimension, and
423 not the irrelevant stimulus. The current finding of separate, but overlapping, populations of units encoding
424 cue identity and expected outcome, suggests that the fMRI finding is generated by the combined activity of
425 several different functional cell types.

426 Our analyses were designed to eliminate several potential alternative interpretations to cue identity coding.
427 Because the different cues were separated into different blocks, units that discriminated between cue identi-
428 ties could instead be encoding time or other slowly-changing quantities. We excluded this possible confound
429 by excluding units that showed a drift in firing between the first and second half within a block. However, the
430 possibility remains that instead of or in addition to stimulus identity, these units encode a preferred context,
431 or even a macroscale representation of progress through the session. Indeed, encoding of the current strategy
432 could be an explanation for the sustained difference in population averaged firing across stimulus blocks
433 (Figure 6), as well as a potential explanation for the differentially tiling of task structure across blocks in the

434 current study (Figure 7).

435 A different potential confound is that between outcome and action value coding. We discriminated between
436 these possibilities by analyzing error trials, where the rat approached reward (left turn) after presentation of
437 the reward-unavailable cue. Units that were modulated by the expected outcome of the cue maintained their
438 specific firing patterns even during error trials, as expected from outcome value coding but not action value
439 coding. Additionally, NAc signals have been shown to be modulated by response vigor (McGinty et al.,
440 2013); to detangle this from our results we included trial length (i.e. latency to arrival at the reward site) as a
441 predictor in our GLMs, and found units with cue feature correlates independent of trial length.

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

450 **Encoding of position:**

451 Our finding that cue-evoked activity was modulated by cue location is in alignment with several previous re-
452 ports ([Lavoie & Mizumori, 1994](#); [Mulder, Shibata, Trullier, & Wiener, 2005](#); [Strait et al., 2016](#); [Wiener et al., 2003](#))
453 ([Lavoie & Mizumori, 1994](#); [Mulder et al., 2005](#); [Strait et al., 2016](#); [Wiener et al., 2003](#)). The NAc receives
454 inputs from the hippocampus, and the communication of place-reward information across the two structures
455 suggests that the NAc tracks locations associated with reward ([Lansink et al., 2008](#); [Lansink, Goltstein, Lankelma, McNaughton, & Redish, 2011](#))
456 ([Lansink et al., 2008, 2009, 2016](#); [Pennartz, 2004](#); [Sjulson et al., 2017](#); [Tabuchi et al., 2000](#); [van der Meer & Redish, 2011](#))

457 . NAc units can also signal progress through a sequence of cues and/or actions (Atallah et al., 2014; Berke, Breck, & Eichenba
458 (Atallah et al., 2014; Berke et al., 2009; Khamassi et al., 2008; Lansink et al., 2012; Mulder et al., 2004; Shidara et al., 1998
459 . Given that the current task was pseudo-random, it is possible that the rats learned the structure of sequential
460 cue presentation, and the neural activity could reflect this. However, this is unlikely as including a previous
461 trial variable in the analysis did not explain a significant amount of firing rate variance in response to the
462 cue for the vast majority of units. In any case, NAc units on the present task continued to distinguish be-
463 tween different locations, even though location, and progress through a sequence, were explicitly irrelevant
464 in predicting reward. We speculate that this persistent coding of location in NAc may represent a bias in
465 credit assignment, and associated tendency for rodents to associate motivationally relevant events with the
466 locations where they occur.

467 **Implications:**

468 Maladaptive decision making, as occurs in schizophrenia, addiction, Parkinson's, among others, can re-
469 sult from dysfunctional RPE and value signals (Frank, Seeberger, & O'Reilly, 2004; Gradin et al., 2011;
470 Maia & Frank, 2011). This view has been successful in explaining both positive and negative symptoms in
471 schizophrenia, and deficits in learning from feedback in Parkinson's (Frank et al., 2004; Gradin et al., 2011).
472 However, the effects of RPE and value updating are contingent upon encoding of preceding action and cue
473 features, the eligibility trace (Lee et al., 2012; Sutton & Barto, 1998)(Lee et al., 2012; Sutton & Barto, 1998)
474 . Value updates can only be performed on these aspects of preceding experience that are encoded when
475 the update occurs. Therefore, maladaptive learning and decision making can result from not only aber-
476 rant RPEs but also from altered cue feature encoding. For instance, on this task the environmental stim-
477 ulus that signaled the availability of reward was conveyed by two distinct cues that were presented in
478 four locations. While in our current study, the location and identity of the cue did not require any ad-
479 justments in the animals behavior, we found coding of these features alongside the expected outcome of
480 the cue that could be the outcome of credit assignment computations computed upstream. Identifying
481 neural coding related to an aspect of credit assignment is important as inappropriate credit assignment

482 could be a contributor to conditioned fear overgeneralization seen in disorders with pathological anxiety
483 such as generalized anxiety disorder, post traumatic stress disorder, and obsessive-compulsive disorder
484 (Kaczkurkin et al., 2017; Kaczkurkin & Lissek, 2013; Lissek et al., 2014)(Kaczkurkin et al., 2017; Kaczkurkin & Lissek, 2014),
485 , and delusions observed in disorders such as schizophrenia, Alzheimer's and Parkinson's (Corlett, Taylor, Wang, Fletcher, & H
486 's (Corlett et al., 2010; Kapur, 2003). Thus, our results provide a neural window into the process of credit
487 assignment, such that the extent and specific manner in which this process fails in e.g.- syndromes such as
488 schizophrenia, obsessive-compulsive disorder, etc. can be experimentally accessed.

489 **OUTTAKES**

490 **Methods**

491 **Tiling of task structure Subjects:**

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

498 **Overall timeline:**

499 We found that the population of recorded units had a relatively uniform distribution of firing fields within our
500 taskspace, similar to what has been reported previously (Berke et al., 2009; Lansink et al., 2012; Shidara et al., 1998)

501 . Uniquely, we found that this representation was statistically different across conditions of a cue features,
502 such as according to whether the rat was currently engaged in the light or sound block, suggesting that this
503 could be a possible neural correlate for encoding the currently relevant strategy in the NAc. It has been
504 previously shown that during progress through a predictable trial series, units represented state value of
505 cue, and that single-unit responses allowed the monkey to know how it was progressing throughout the task
506 (Shidara et al., 1998). Likewise, the tiling we saw could be a consequence of upstream cortical or limbic
507 inputs informing the striatum of the current task rules. Another possibility is that the NAc not only pays
508 attention to progress throughout a task within a Each rat was first handled for seven days during which they
509 were exposed to the experiment room, the sucrose solution used as a reinforcer, and the click of the sucrose
510 dispenser valves. Rats were then trained on the behavioral task (described in the next section) until they
511 reached performance criterion. At this point they underwent hyperdrive implantation targeted at the NAc.
512 Rats were allowed to recover for a minimum of five days before being retrained on the task, and recording
513 began once performance returned to pre-surgery levels. Upon completion of recording, animals were glosed,
514 euthanized and recording sites were histologically confirmed.

515 **Behavioral task and training:**

516 The behavioral apparatus was an elevated, square-shaped track (100 x 100 cm, track width 10 cm) containing
517 four possible reward locations at the end of track “arms” (Figure 2). Rats initiated a trial by triggering
518 a photobeam located 24 cm from the start of each arm. Upon trial initiation, one of two possible light
519 cues (L1, but also higher-order task information, like blocks. Furthermore, dopamine levels in the NAc
520 fluctuate through a trial, and it is possible that the observed tiling could be L2), or one of two possible
521 sound cues (S1, S2), was presented that signaled the presence (*reward-available trial*, L1+, S1+) or absence
522 (*reward-unavailable trial*, L2-, S2-) of a 12% sucrose water reward (0.1 mL) at the upcoming reward site. A
523 trial was classified as an *approach trial* if the rat turned left at the decision point and made a nosepoke at the
524 reward receptacle (40 cm from the decision point), while trials were classified as a *skip trial* if the rat instead
525 turned right at the decision point and triggered the photobeam to initiate the next trial. A trial is labeled

526 *correct* if the rat approached (i.e. nosepoked) on reward-available trials, and skipped (i.e. did not nosepoke)
527 on reward-unavailable trials. On reward-available trials there was a 1 second delay between a nosepoke and
528 subsequent reward delivery. *Trial length* was determined by measuring the length of time from cue onset
529 until nosepoke (for approach trials), or from cue onset until the start of the following trial (for skip trials).
530 Trials could only be initiated through clockwise progression through the series of arms, and each entry into
531 the subsequent arm on the track counted as a trial.

532 Each session consisted of both a *light block* and a *sound block* with 100 trials each. Within a block, one cue
533 signaled reward was available on that trial (L1+ or S1+), while the other signaled reward was not available
534 (L2- or S2-). Light block cues were a flashing white light, and a constant yellow light. Sound block cues
535 were a 2 kHz sine wave and a 8 kHz sine wave whose amplitude was modulated from 0 to maximum by
536 a 2 Hz sine wave. Outcome-cue associations were counterbalanced across rats, e.g. for some rats L1+ was
537 the flashing white light, and for others L1+ was the constant yellow light. The order of cue presentation
538 was pseudorandomized so that the same cue could not be presented more than twice in a row. Block order
539 within each day was also pseudorandomized, such that the rat could not begin a session with the same
540 block for more than two days in a row. Each session consisted of a ~~NAc representation of state value~~
541 ~~related to this temporally evolving dopamine signal. Future experiments should monitor this mapping of~~
542 ~~task structure during the application of dopamine antagonists. Finally, the presence of functional correlates~~
543 ~~not evident when looking at single-unit responses time-locked to salient task events emphasizes the need to~~
544 ~~employ ensemble level analyses across all aspects of a 5 minute pre-session period on a pedestal (a terracotta~~
545 ~~planter filled with towels), followed by the first block, then the second block, then a 5 minute post-session~~
546 ~~period on the pedestal. For approximately the first week of training, rats were restricted to running in~~
547 ~~the clockwise direction by presenting a physical barrier to running counterclockwise. Cues signaling the~~
548 ~~availability and unavailability of reward, as described above, were present from the start of training. Rats~~
549 ~~were trained for 200 trials per day (100 trials per block) until they discriminated between the reward-available~~
550 ~~and reward-unavailable cues for both light and sound blocks for three consecutive days, according to a~~
551 ~~chi-square test rejecting the null hypothesis of equal approaches for reward-available and reward-unavailable~~

552 trials, at which point they underwent electrode implant surgery.

553 **Surgery:**

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

562 **Potential functional consequences of persistent coding****Data acquisition and preprocessing:**

563 In the current study we found that the coding of cue features persisted after the choice point, during a
564 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. After recovery, rats were placed back on the task for recording. NAc signals were
567 acquired at 20 kHz with a RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition
568 system (Amplipex). Signals were referenced against a tetrode placed in the corpus callosum above the NAc.
569 Candidate spikes for sorting into putative single units were obtained by band-pass filtering the data between
570 600-9000 Hz, thresholding and aligning the peaks (UltraMegaSort2k, Hill et al., 2011). Spike waveforms
571 were then clustered with KlustaKwik using energy and the first derivative of energy as features, and manually
572 sorted into units (MClust 3.5, A.D. Redish et al., <http://redishlab.neuroscience.umn.edu/MClust/MClust.html>).
573 Isolated units containing a minimum of 200 spikes within a session were included for subsequent analysis.
574 Units were classified as fast spiking interneurons (FSIs) by an absence of interspike intervals (ISIs) > 2 s,

575 while medium spiny neurons (MSNs) had a combination of ISIs > 2 s and phasic activity with shorter ISIs
576 (Atallah et al., 2014; Barnes et al., 2005).

577 **Data analysis:**

578 *Behavior.* To determine if rats distinguished behaviorally between the reward-available and is-maintained
579 online until the outcome is revealed could be useful for assigning credit of a reward to reward-unavailable
580 cues (*cue outcome*), we generated linear mixed effects models to investigate the relationships between cue
581 type and our behavioral variables, with *cue outcome* (reward available or not) and *cue identity* (light or sound)
582 as fixed effects, and the appropriate elements of an environment, to develop an accurate value function and
583 optimize long-term acquisition of reward (Lee et al., 2012). A recent non-human primate experiment found
584 evidence for simultaneous coding of outcome and outcome-predictive stimuli at the time of feedback in the
585 dorsolateral PFC, and that this coding of stimulus information was strongest before behavioral performance
586 stabilized (Asaad et al., 2017). Interestingly, we found coding of cue identity and location after presentation
587 of the cue, addition of an intercept for rat identity as a random effect. For each cue, the average proportion of
588 trials approached and trial length for a session were used as response variables. Contribution of cue outcome
589 to behavior was determined by comparing the full model to a model with cue outcome removed for each
590 behavioral variable.

591 *Neural data.* To investigate the contribution of different cue features (cue identity and cue outcome) on the
592 firing rates of NAc single units, we first determined whether firing rates for a unit were modulated by the
593 onset of a cue by collapsing across all cues and comparing the firing rates for the 1 s preceding cue-onset with
594 the 1 s following cue-onset. Single units were considered to be *cue-modulated* if a Wilcoxon signed-rank
595 test comparing pre- and during a delay period between nosepoke and outcome, but not after the outcome
596 was received. The absence of coding during feedback, and that these recordings were done after behavioral
597 performance had stabilized suggests the possibility that these NAc representations represent an integrated
598 value estimate that is the outcome of earlier upstream credit assignment in cortical inputs. From a motivated

599 behavior standpoint, post-cue firing was significant at $p < .01$. Cue-modulated units were then classified
600 as either increasing or decreasing if the post-cue activity was higher or lower than the pre-cue activity,
601 respectively.

602 To determine the relative contribution of different task parameters to firing rate variance (as in Figures 4-5),
603 a forward selection stepwise general linear model (GLM) was fit to each cue-modulated unit. Cue identity
604 (light block, sound block), cue location (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available,
605 reward-unavailable), behavior (approach, skip), trial length, trial number, and trial history (reward availability
606 on the previous 2 trials) were used as predictors, and the 1 s post-cue firing rate as the response variable.
607 Units were classified as being modulated by a given task parameter if addition of the parameter significantly
608 improved model fit using deviance as the criterion ($p < .01$). A comparison of the R-squared value between
609 the final model and the presence of these enriched representations can help inform action selection involving
610 Pavlovian behaviors such as conditioned approach (Ciano et al., 2001; Parkinson, Willoughby, Robbins, & Everitt, 2000; Sau
611 , and after a decision has been made the maintained representation can help to hold the response in the face of
612 competing alternatives while waiting for an outcome (Di Ciano, Robbins, & Everitt, 2008; Floresco, 2015; Floresco, McLaugh
613 .—final model minus the predictor of interest was used to determine the amount of firing rate variance
614 explained by the addition of that predictor for a given unit. To investigate more finely the temporal dynamics
615 of the influence of task parameters to unit activity, we then fit a sliding window GLM with the same task
616 parameters using 500 ms bins and 100 ms steps, starting 500 ms before cue-onset, up to 500 ms after
617 cue-onset, and measured the proportion of units and average R-squared value for a given time bin where
618 a particular predictor contributed significantly to the final model. To control for the amount of units that
619 would be affected by a predictor by chance, we shuffled the trial order of firing rates for a particular unit
620 within a time bin, and took the average of this value over 100 shuffles.

621 If these representations function as an eligibility trace, then there should be a relationship between the
622 content and robustness of these representations and the degree to which organisms form adaptive associations
623 between reinforcers and their environment, which could contribute to the presence of prediction errors in the

624 NAc in those that learned an experimental task in humans versus those that did not (?). In the present
625 experiment, understanding what information about the environment the NAc is keeping track of when
626 reward arrives could be useful for potentially determining what representations are being reinforced. Would
627 be interesting to see if selectively reward an animal when certain representations are active To better visualize
628 responses to cues and enable subsequent population level analyses (as in Figures 4, 6, and 7), spike trains
629 were convolved with a Gaussian kernel ($\sigma = 100$ ms), and peri-event time histograms (PETHs) were generated
630 by taking the average of the convolved spike trains across all trials for a given task condition. For analysis of
631 population-level responses for cue features (Figure 6), convolved spike trains for all units where cue identity,
632 cue location, or cue outcome explained a significant portion of firing rate variance were z-scored. Within a
633 given cue feature, normalized spike trains were then separated according to the preferred and non-preferred
634 cue condition (e.g. light vs. sound block), and averaged across units to generate population-level averages.
635 To account for separation that would result from any random selection of units, unit identity was shuffled
636 and the shuffled average for preferred and non-preferred cue conditions was generated for 1000 shuffles.

637 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units
638 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure
639 7). To compare representations of task space across cue conditions for a cue feature, the ordering of units
640 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other
641 condition (e.g. arm 1), would help shape their assignment of reward to that arm. Could be useful for altering
642 a maladaptive preference back to a more adaptive one, as the loss of certain aspects of coding via BLA
643 and PFC inputs leads to the loss of outcome predictive activity to discrete cues in the NAc. Furthermore,
644 non-discriminatory coding in the NAc could hypothetically be correlated to over-generalization of situations,
645 in which the animal may not recognize it is in a different situation and perform actions that are inefficient or
646 maladaptive for reward procurement sound block). For control comparisons within cue conditions, half of
647 the trials for a condition were compared against the other half. To look at the correlation of firing rates of
648 all units within and across various cue conditions, trials for each cue condition for a unit were shuffled and
649 divided into two averages, and averages within and across cue conditions were correlated. A linear mixed

650 effects model was run for each cue condition to determine if correlations of firing rates within cue conditions
651 were more similar than correlations across cue conditions.

652 A fundamental problem faced by all reinforcement learning agents is which features to credit and blame for
653 particular outcomes (Sutton & Barto, 1998); adaptive behavior requires associating motivationally relevant
654 outcomes with the cues that predict them through learning from feedback. Much work has focused on value
655 signals such as reward prediction errors (RPEs), state values and action values (Lee et al., 2012; Maia, 2009)
656 . However, successful learning requires not only a RPE but also a trace of the preceding actions and To
657 identify the responsivity of units to different cue features at the time of a nosepoke into a reward receptacle,
658 and subsequent reward delivery, the same cue-responsive units from the cue-onset analyses were analyzed at
659 the time of nosepoke and outcome receipt using identical analysis techniques (Figures 8, 9, 10, and 11).

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

667 **Histology:**

668 Upon completion of the experiment, recording channels were glosed by passing 10 μ A current for 10
669 seconds and waiting 5 days before euthanasia, except for rat R057 whose implant detached prematurely.
670 Rats were anesthetized with 5% isoflurane, then asphyxiated with carbon dioxide. Transcardial perfusions
671 were performed, and brains were fixed and removed. Brains were sliced in 50 μ m coronal sections and
672 stained with thionin. Slices were visualized under light microscopy, tetrode placement was determined, and

673 electrodes with recording locations in the NAc were analyzed (Figure 12).

674 [Figure 12 about here.]

675 **References**

- 676 Akaiishi, R., Kolling, N., Brown, J. W., & Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue
677 Environment. *Journal of Neuroscience*, 36(4), 1096–1112. doi: 10.1523/JNEUROSCI.3159-15.2016
- 678 Ambroggi, F., Ishikawa, A., Fields, H. L., & Nicola, S. M. (2008). Basolateral Amygdala Neurons Facilitate Reward-Seeking
679 Behavior by Exciting Nucleus Accumbens Neurons. *Neuron*, 59(4), 648–661. doi: 10.1016/j.neuron.2008.07.004
- 680 Asaad, W. F., Lauro, P. M., Perge, J. A., & Eskandar, E. N. (2017). Prefrontal Neurons Encode a Solution to the Credit-Assignment
681 Problem. *The Journal of Neuroscience*, 37(29), 6995–7007. doi: 10.1523/JNEUROSCI.3311-16.2017
- 682 Atallah, H. E., McCool, A. D., Howe, M. W., & Graybiel, A. M. (2014). Neurons in the ventral striatum exhibit cell-type-specific
683 representations of outcome during learning. *Neuron*, 82(5), 1145–1156. doi: 10.1016/j.neuron.2014.04.021
- 684 Averbeck, B. B., & Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*,
685 20(4), 505–512. doi: 10.1038/nn.4506
- 686 Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and
687 recoding of procedural memories. *Nature*, 437(7062), 1158–1161. doi: 10.1038/nature04053
- 688 Berke, J. D., Breck, J. T., & Eichenbaum, H. (2009). Striatal Versus Hippocampal Representations During Win-Stay Maze Perform-
689 ance. *Journal of Neurophysiology*, 101(3), 1575–1587. doi: 10.1152/jn.91106.2008
- 690 Berridge, K. C. (2012). From prediction error to incentive salience: Mesolimbic computation of reward motivation. *European
691 Journal of Neuroscience*, 35(7), 1124–1143. doi: 10.1111/j.1460-9568.2012.07990.x
- 692 Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., ... Roesch, M. R. (2013). Separate
693 Populations of Neurons in Ventral Striatum Encode Value and Motivation. *PLoS ONE*, 8(5), e64673. doi: 10.1371/jour-
694 nal.pone.0064673
- 695 Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological
696 Bulletin*, 114(1), 80–99. doi: 10.1371/journal.pone.0064673
- 697 Carelli, R. M. (2010). Drug Addiction: Behavioral Neurophysiology. In *Encyclopedia of neuroscience* (pp. 677–682). Elsevier.
698 doi: 10.1016/B978-008045046-9.01546-1
- 699 Chang, S. E., & Holland, P. C. (2013). Effects of nucleus accumbens core and shell lesions on autoshaped lever-pressing. *Be-*

- 700 *havioural Brain Research*, 256, 36–42. doi: 10.1016/j.bbr.2013.07.046
- 701 Chang, S. E., Wheeler, D. S., & Holland, P. C. (2012). Roles of nucleus accumbens and basolateral amygdala in autoshaped lever
702 pressing. *Neurobiology of Learning and Memory*, 97(4), 441–451. doi: 10.1016/j.nlm.2012.03.008
- 703 Chau, B. K. H., Sallet, J., Papageorgiou, G. K., Noonan, M. A. P., Bell, A. H., Walton, M. E., & Rushworth, M. F. S. (2015).
704 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5),
705 1106–1118. doi: 10.1016/j.neuron.2015.08.018
- 706 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., & Wightman, R. M. (2007). Coordinated
707 Accumbal Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2), 237–244. doi:
708 10.1016/j.neuron.2007.03.021
- 709 Ciano, P. D., Cardinal, R. N., Cowell, R. A., Little, S. J., Everitt, B. J., Di Ciano, P., ... Everitt, B. J. (2001). Differential involvement
710 of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of
711 pavlovian approach behavior. *The Journal of neuroscience*, 21(23), 9471–9477. doi: 21/23/9471 [pii]
- 712 Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in
713 schizophrenia. *Psychological Review*, 99(1), 45. doi: 21/23/9471 [pii]
- 714 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., & Schoenbaum, G. (2015).
715 Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature Com-
716 munications*, 6, 7195. doi: 10.1038/ncomms8195
- 717 Corbit, L. H., & Balleine, B. W. (2011). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are
718 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33), 11786–11794. doi:
719 10.1523/JNEUROSCI.2711-11.2011
- 720 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in
721 Neurobiology*, 92, 345–369. doi: 10.1016/j.pneurobio.2010.06.007
- 722 Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine
723 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8), 1020–1028. doi: 10.1038/nn1923
- 724 Day, J. J., Wheeler, R. A., Roitman, M. F., & Carelli, R. M. (2006). Nucleus accumbens neurons encode Pavlovian approach behav-
725 iors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5), 1341–1351. doi: 10.1111/j.1460-
726 9568.2006.04654.x
- 727 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., ... Le Moine, C. (2017). Memories of Opiate Withdrawal
728 Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychopharmacology*,
729 42(5), 1157–1168. doi: 10.1038/npp.2016.272
- 730 Di Ciano, P., Robbins, T. W., & Everitt, B. J. (2008). Differential effects of nucleus accumbens core, shell, or dorsal striatal
731 inactivations on the persistence, reacquisition, or reinstatement of responding for a drug-paired conditioned reinforcer. *Neu-
732 ropsychopharmacology*, 33(6), 1413–1425. doi: 10.1038/sj.npp.1301522

- 733 du Hoffmann, J., & Nicola, S. M. (2014). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the
734 Nucleus Accumbens. *Journal of Neuroscience*, 34(43), 14349–14364. doi: 10.1523/JNEUROSCI.3492-14.2014
- 735 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental
736 Psychology*, 32(2), 150–155. doi: 10.1037/h0058316
- 737 FitzGerald, T. H. B., Schwartenbeck, P., & Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent
738 and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4), 1271–1279. doi: 10.1523/JNEUROSCI.4389-
739 13.2014
- 740 Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., ... Akil, H. (2011). A selective role for dopamine in
741 stimulus-reward learning. *Nature*, 469(7328), 53–59. doi: 10.1038/nature09588
- 742 Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of
743 Psychology*, 66(1), 25–52. doi: 10.1146/annurev-psych-010213-115159
- 744 Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., & Magyar, O. (2006). Dissociable roles for the nucleus accumbens core and shell
745 in regulating set shifting. *Journal of Neuroscience*, 26(9), 2449–2457. doi: 10.1523/JNEUROSCI.4431-05.2006
- 746 Floresco, S. B., McLaughlin, R. J., & Haluk, D. M. (2008). Opposing roles for the nucleus accumbens core and shell in cue-induced
747 reinstatement of food-seeking behavior. *Neuroscience*, 154(3), 877–884. doi: 10.1016/j.neuroscience.2008.04.004
- 748 Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism.
749 *Science*, 306(5703), 1940–1943. doi: 10.1126/science.1102941
- 750 Giertler, C., Bohn, I., & Hauber, W. (2004). Transient inactivation of the rat nucleus accumbens does not impair guidance of
751 instrumental behaviour by stimuli predicting reward magnitude. *Behavioural Pharmacology*, 15(1), 55–63. doi: 10.1126/sci-
752 ence.1102941
- 753 Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., ... Roesch, M. R. (2012). Ventral Striatum
754 Encodes Past and Predicted Value Independent of Motor Contingencies. *Journal of Neuroscience*, 32(6), 2027–2036. doi:
755 10.1523/JNEUROSCI.5349-11.2012
- 756 Goto, Y., & Grace, A. A. (2008). Limbic and cortical information processing in the nucleus accumbens. *Trends in Neurosciences*,
757 31(11), 552–558. doi: 10.1016/j.tins.2008.08.002
- 758 Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., ... Steele, J. D. (2011). Expected value and prediction
759 error abnormalities in depression and schizophrenia. *Brain*, 134(6), 1751–1764. doi: 10.1093/brain/awr059
- 760 Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a
761 Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. doi: 10.1037/h0059831
- 762 Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., ... Berke, J. D. (2015).
763 Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1), 117–126. doi: 10.1038/nn.4173
- 764 Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014). Phasic Dopamine Release in the Rat Nucleus Ac-
765 cumbens Symmetrically Encodes a Reward Prediction Error Term. *The Journal of Neuroscience*, 34(3), 698–704. doi:

- 766 10.1523/JNEUROSCI.2489-13.2014
- 767 Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: the stimulus-reinforcer relation and directed action*. Psychonomic Society.doi:
768 10.1523/JNEUROSCI.2489-13.2014
- 769 Hill, D. N., Mehta, S. B., & Kleinfeld, D. (2011). Quality Metrics to Accompany Spike Sorting of Extracellular Signals. *Journal of*
770 *Neuroscience*, 31(24), 8699–8705. doi: 10.1523/JNEUROSCI.0971-11.2011
- 771 Holland, P. C. (1992). Occasion setting in pavlovian conditioning. *Psychology of Learning and Motivation*, 28(C), 69–125. doi:
772 10.1016/S0079-7421(08)60488-0
- 773 Hollerman, J. R., Tremblay, L., & Schultz, W. (1998). Influence of Reward Expectation on Behavior-Related Neuronal Activity in
774 Primate Striatum. *Journal of Neurophysiology*, 80(2), 947–963. doi: 10.1152/jn.1998.80.2.947
- 775 Honey, R. C., Iordanova, M. D., & Good, M. (2014). Associative structures in animal learning: Dissociating elemental and
776 configural processes. *Neurobiology of Learning and Memory*, 108, 96–103. doi: 10.1016/j.nlm.2013.06.002
- 777 Hyman, S. E., Malenka, R. C., & Nestler, E. J. (2006). NEURAL MECHANISMS OF ADDICTION: The Role of Reward-Related
778 Learning and Memory. *Annual Review of Neuroscience*, 29(1), 565–598. doi: 10.1146/annurev.neuro.29.051605.113009
- 779 Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-
780 olfactory tubercle complex. *Brain Research Reviews*, 56(1), 27–78. doi: 10.1016/j.brainresrev.2007.05.004
- 781 Ishikawa, A., Ambroggi, F., Nicola, S. M., & Fields, H. L. (2008). Dorsomedial Prefrontal Cortex Contribution to Behav-
782 ioral and Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, 28(19), 5088–5098. doi:
783 10.1523/JNEUROSCI.0253-08.2008
- 784 Joel, D., Niv, Y., & Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives.
785 *Neural Networks*, 15(4-6), 535–547. doi: 10.1016/S0893-6080(02)00047-3
- 786 Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., ... Lissek, S. (2017).
787 Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174(2), 125–134. doi:
788 10.1176/appi.ajp.2016.15121549
- 789 Kaczkurkin, A. N., & Lissek, S. (2013). Generalization of Conditioned Fear and Obsessive-Compulsive Traits. *Journal of Psychol-*
790 *ogy & Psychotherapy*, 7, 3. doi: 10.4172/2161-0487.S7-003
- 791 Kalivas, P. W., & Volkow, N. D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal of*
792 *Psychiatry*, 162(8), 1403–1413. doi: 10.1176/appi.ajp.162.8.1403
- 793 Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in
794 schizophrenia. *American Journal of Psychiatry*, 160(1), 13–23. doi: 10.1176/appi.ajp.160.1.13
- 795 Khamassi, M., & Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and
796 model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, 6, 79. doi: 10.3389/fnbeh.2012.00079
- 797 Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., & Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal

- 798 neurons of behaving rats. *European Journal of Neuroscience*, 28(9), 1849–1866. doi: 10.1111/j.1460-9568.2008.06480.x
- 799 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., & Pennartz, C. M. A. (2008).
- 800 Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, 28(25),
- 801 6372–6382. doi: 10.1523/JNEUROSCI.1054-08.2008
- 802 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., & Pennartz, C. M. (2009). Hippocampus leads ventral
- 803 striatum in replay of place-reward information. *PLoS Biology*, 7(8), e1000173. doi: 10.1371/journal.pbio.1000173
- 804 Lansink, C. S., Jackson, J. C., Lankelma, J. V., Ito, R., Robbins, T. W., Everitt, B. J., & Pennartz, C. M. A. (2012). Reward Cues
- 805 in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles. *Journal of Neuro-*
- 806 *science*, 32(36), 12444–12459. doi: 10.1523/JNEUROSCI.0593-12.2012
- 807 Lansink, C. S., Meijer, G. T., Lankelma, J. V., Vinck, M. A., Jackson, J. C., & Pennartz, C. M. A. (2016). Reward Expectancy
- 808 Strengthens CA1 Theta and Beta Band Synchronization and Hippocampal-Ventral Striatal Coupling. *Journal of Neuro-*
- 809 *science*, 36(41), 10598–10610. doi: 10.1523/JNEUROSCI.0682-16.2016
- 810 Lavoie, A. M., & Mizumori, S. J. (1994). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of
- 811 rats. *Brain Research*, 638(1-2), 157–168. doi: 10.1016/0006-8993(94)90645-9
- 812 Lee, D., Seo, H., & Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of*
- 813 *Neuroscience*, 35(1), 287–308. doi: 10.1146/annurev-neuro-062111-150512
- 814 Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized anxiety disor-
- 815 der is associated with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75(11), 909–915. doi:
- 816 10.1016/j.biopsych.2013.07.025
- 817 Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and*
- 818 *Behavioral Neuroscience*, 9(4), 343–364. doi: 10.3758/CABN.9.4.343
- 819 Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature*
- 820 *Neuroscience*, 14(2), 154–162. doi: 10.1038/nn.2723
- 821 Malhotra, S., Cross, R. W., Zhang, A., & van der Meer, M. A. A. (2015). Ventral striatal gamma oscillations are highly variable
- 822 from trial to trial, and are dominated by behavioural state, and only weakly influenced by outcome value. *European Journal*
- 823 *of Neuroscience*, 42(10), 2818–2832. doi: 10.1038/nn.2723
- 824 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011). Ventral Striatum and Orbitofrontal Cortex
- 825 Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, 31(7), 2700–
- 826 2705. doi: 10.1523/JNEUROSCI.5499-10.2011
- 827 McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., & Nicola, S. M. (2013). Invigoration of reward seeking by cue and proximity
- 828 encoding in the nucleus accumbens. *Neuron*, 78(5), 910–922. doi: 10.1016/j.neuron.2013.04.010
- 829 Mulder, A. B., Shibata, R., Trullier, O., & Wiener, S. I. (2005). Spatially selective reward site responses in tonically active neurons
- 830 of the nucleus accumbens in behaving rats. *Experimental Brain Research*, 163(1), 32–43. doi: 10.1007/s00221-004-2135-3

- 831 Mulder, A. B., Tabuchi, E., & Wiener, S. I. (2004). Neurons in hippocampal afferent zones of rat striatum parse routes into
832 multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7), 1923–1932. doi: 10.1111/j.1460-
833 9568.2004.03301.x
- 834 Nicola, S. M. (2004). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a Discrimi-
835 native Stimulus Task. *Journal of Neurophysiology*, 91(4), 1840–1865. doi: 10.1152/jn.00657.2003
- 836 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of
837 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49), 16585–
838 16600. doi: 10.1523/JNEUROSCI.3958-10.2010
- 839 Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psy-
840 chopharmacology*, 191(3), 507–520. doi: 10.1007/s00213-006-0502-4
- 841 Noonan, M. P., Chau, B. K., Rushworth, M. F., & Fellows, L. K. (2017). Contrasting Effects of Medial and Lateral Orbitofrontal
842 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29), 7023–7035.
843 doi: 10.1523/JNEUROSCI.0692-17.2017
- 844 Parkinson, J. A., Willoughby, P. J., Robbins, T. W., & Everitt, B. J. (2000). Disconnection of the anterior cingulate cortex and
845 nucleus accumbens core impairs pavlovian approach behavior: Further evidence for limbic cortical-ventral striatopallidal
846 systems. *Behavioral Neuroscience*, 114(1), 42–63. doi: 10.1037//0735-7044.114.1.42
- 847 Paxinos, G., & Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates* (4th ed.). San Diego: Academic Press. doi:
848 10.1037//0735-7044.114.1.42
- 849 Pennartz, C. M. A. (2004). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation by
850 Hippocampal Ripples. *Journal of Neuroscience*, 24(29), 6446–6456. doi: 10.1523/JNEUROSCI.0575-04.2004
- 851 Peters, J., LaLumiere, R. T., & Kalivas, P. W. (2008). Infralimbic Prefrontal Cortex Is Responsible for Inhibiting Cocaine Seeking
852 in Extinguished Rats. *Journal of Neuroscience*, 28(23), 6046–6053. doi: 10.1523/JNEUROSCI.1045-08.2008
- 853 Rescorla, R. A., & Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and
854 Instrumental Learning. *Psychological Review*, 74(3), 151–182. doi: 10.1037/h0024475
- 855 Robinson, T. E., & Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues
856 Through the Study of Individual Differences. *Biological Psychiatry*, 65(10), 869–873. doi: 10.1016/j.biopsych.2008.09.006
- 857 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value
858 of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42),
859 13365–13376. doi: 10.1523/JNEUROSCI.2572-09.2009
- 860 Roitman, M. F., Wheeler, R. A., & Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for reward-
861 ing and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4), 587–597. doi:
862 10.1016/j.neuron.2004.12.055
- 863 Saddoris, M. P., Stamatakis, A., & Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nucleus ac-

- 864 cumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*, 33(12),
865 2274–2287. doi: 10.1111/j.1460-9568.2011.07683.x
- 866 Salamone, J. D., & Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3), 470–485.
867 doi: 10.1016/j.neuron.2012.10.021
- 868 Saunders, B. T., & Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of pavlovian-conditioned
869 responses. *European Journal of Neuroscience*, 36(4), 2521–2532. doi: 10.1111/j.1460-9568.2012.08217.x
- 870 Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32. doi:
871 10.1038/nrn.2015.26
- 872 Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.
873 doi: 10.1126/science.275.5306.1593
- 874 Setlow, B., Schoenbaum, G., & Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning.
875 *Neuron*, 38(4), 625–636. doi: 10.1016/S0896-6273(03)00264-2
- 876 Shidara, M., Aigner, T. G., & Richmond, B. J. (1998). Neuronal signals in the monkey ventral striatum related to progress through
877 a predictable series of trials. *Journal of Neuroscience*, 18(7), 2613–25. doi: 10.1016/S0896-6273(03)00264-2
- 878 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., & Buzsáki, G. (2017). Cocaine place conditioning strengthens location-
879 specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, 1–10. doi: 10.1101/105890
- 880 Sleezer, B. J., Castagno, M. D., & Hayden, B. Y. (2016). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.
881 *Journal of Neuroscience*, 36(44), 11223–11237. doi: 10.1523/JNEUROSCI.1766-16.2016
- 882 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., & Hayden, B. Y. (2016). Neuronal selectivity
883 for spatial positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3), 1098–1111. doi:
884 10.1152/jn.00325.2015
- 885 Sugam, J. A., Saddoris, M. P., & Carelli, R. M. (2014). Nucleus accumbens neurons track behavioral preferences and reward
886 outcomes during risky decision making. *Biological Psychiatry*, 75(10), 807–816. doi: 10.1016/j.biopsych.2013.09.010
- 887 Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction* (Vol. 9) (No. 5). MIT Press, Cambridge, MA. doi:
888 10.1109/TNN.1998.712192
- 889 Tabuchi, E. T., Mulder, A. B., & Wiener, S. I. (2000). Position and behavioral modulation of synchronization of hippocam-
890 pal and accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6), 717–728. doi: 10.1002/1098-
891 1063(2000)10:6;717::AID-HIPO1009;3.0.CO;2-3
- 892 Takahashi, Y. K., Langdon, A. J., Niv, Y., & Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors
893 Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1), 182–193. doi:
894 10.1016/j.neuron.2016.05.015
- 895 van der Meer, M. A. A., & Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward
896 Information. *Journal of Neuroscience*, 31(8), 2843–2854. doi: 10.1523/JNEUROSCI.4869-10.2011

- 897 West, E. A., & Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after
898 Reinforcer Devaluation. *Journal of Neuroscience*, 36(4), 1128–1139. doi: 10.1523/JNEUROSCI.2976-15.2016
- 899 Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., & Mulder, A. B. (2003). Spatial and behavioral correlates in
900 nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,
901 1250(C), 275–292. doi: 10.1016/S0531-5131(03)00978-6
- 902 Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral
903 and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, 24(12), 2923–2933. doi:
904 10.1523/JNEUROSCI.5282-03.2004

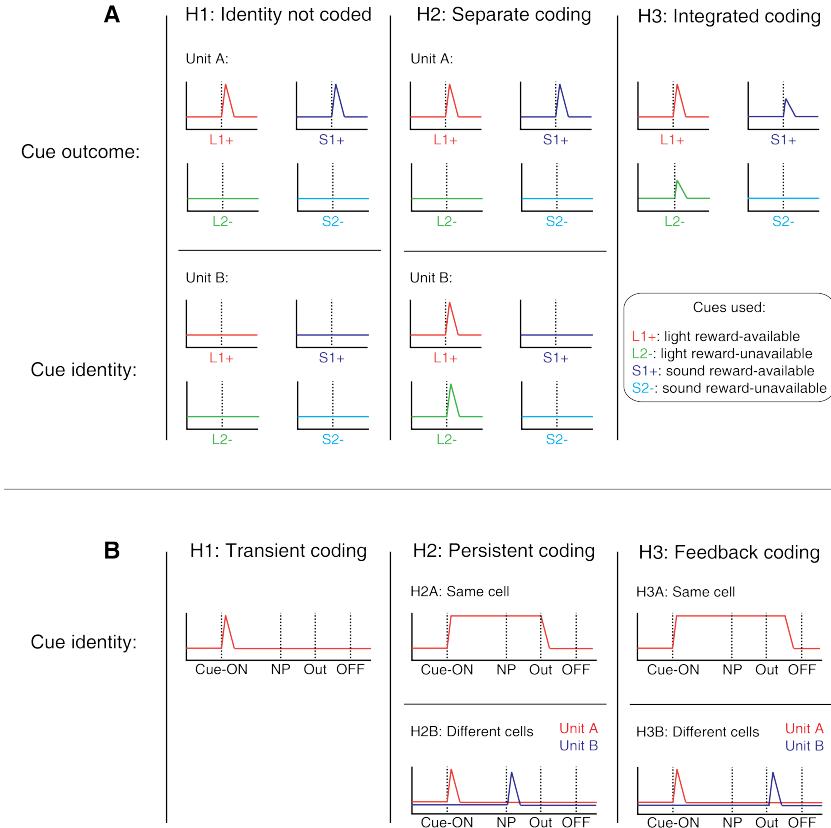


Figure 1: Schematic of potential coding strategies for cue identity (light, sound) and cue outcome (reward-available, reward-unavailable) employed by single units in the NAc across different units (A) and in time (B). **A:** Displayed are schematic PETHs illustrating putative responses to different cues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

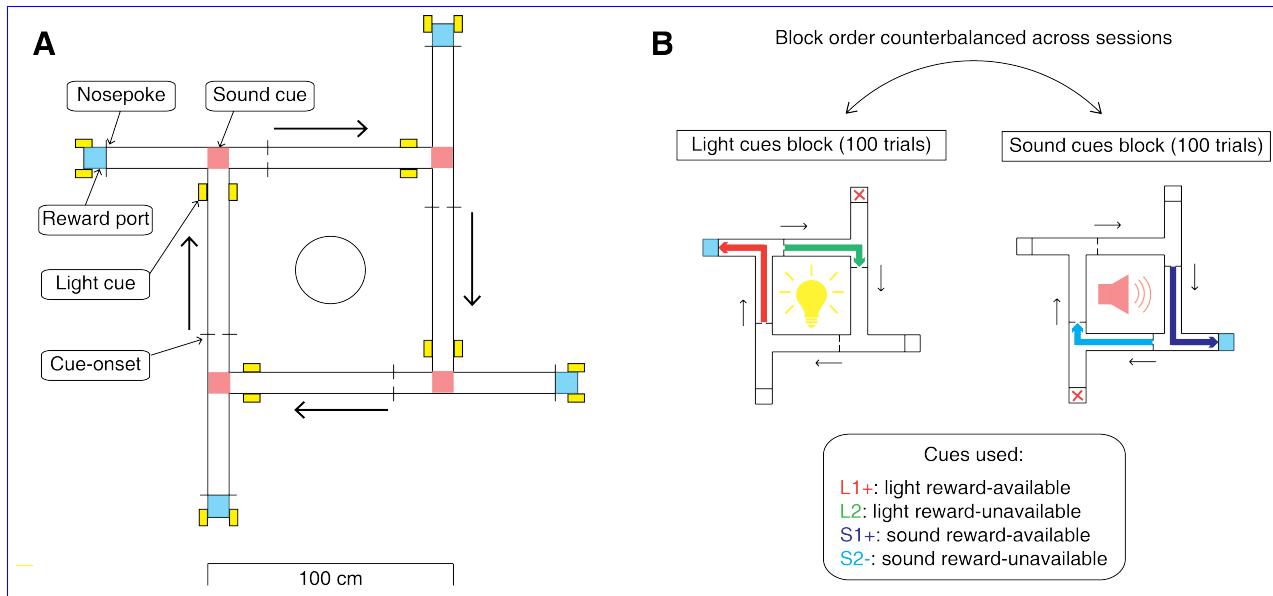


Figure 2: Schematic of behavioral task. **A:** To scale depiction of square track consisting of multiple identical T-choice points. At each choice point, the availability of 12% sucrose reward at the nearest reward receptacle (light blue fill) was signaled by one of four possible cues, presented when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

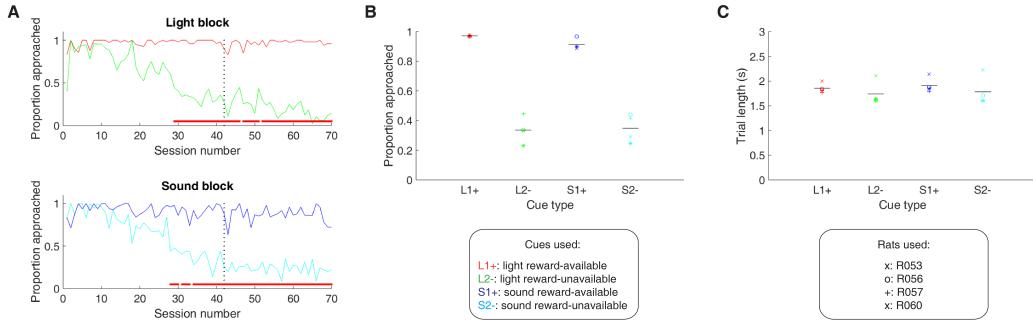


Figure 3: Performance on the behavioral task. **A.** Example learning curves across sessions from a single subject (R060) showing the proportion approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light block, light blue line for sound block) for light (top) and sound (bottom) blocks. Fully correct performance corresponds to an approach proportion of 1 for reward-available trials and 0 for reward-unavailable trials. Rats initially approach on both reward-available and reward-unavailable trials, and learn with experience to skip non-rewarded trials. Red bars indicate days in which a rat statistically discriminated between reward-available and reward-unavailable cues, determined by a chi square test. Dashed line indicates time of electrode implant surgery. **B-C:** Summary of performance during recording sessions for each rat. **B:** Proportion approached for all rats, averaged across all recording sessions. Different columns indicate the different cues (reward-available (red) and reward-unavailable (green) light cues, reward-available (navy blue) and reward-unavailable (light blue) sound cues). Different symbols correspond to individual subjects; horizontal black line shows the mean. All rats learned to discriminate between reward-available and reward-unavailable cues, as indicated by the clear difference of proportion approached between reward-available (~90% approached) and reward-unavailable cues (~30% approached), for both blocks (see Results for statistics). **C:** Average trial length for each cue. Note that the time to complete a trial was comparable for the different cues.

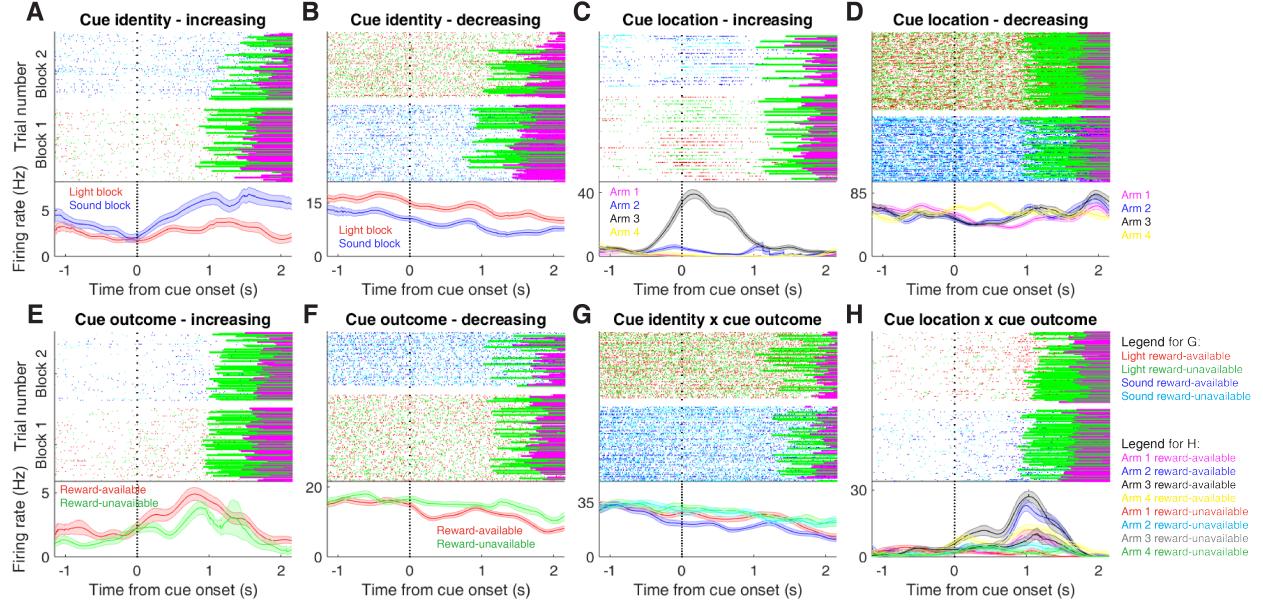


Figure 4: Examples of different cue-modulated NAc units influenced by various task parameters. **A:** Example of a cue-modulated NAc unit that showed an increase in firing following the cue, and encoded cue identity. Top: rasterplot showing the spiking activity across all trials aligned to cue-onset. 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.

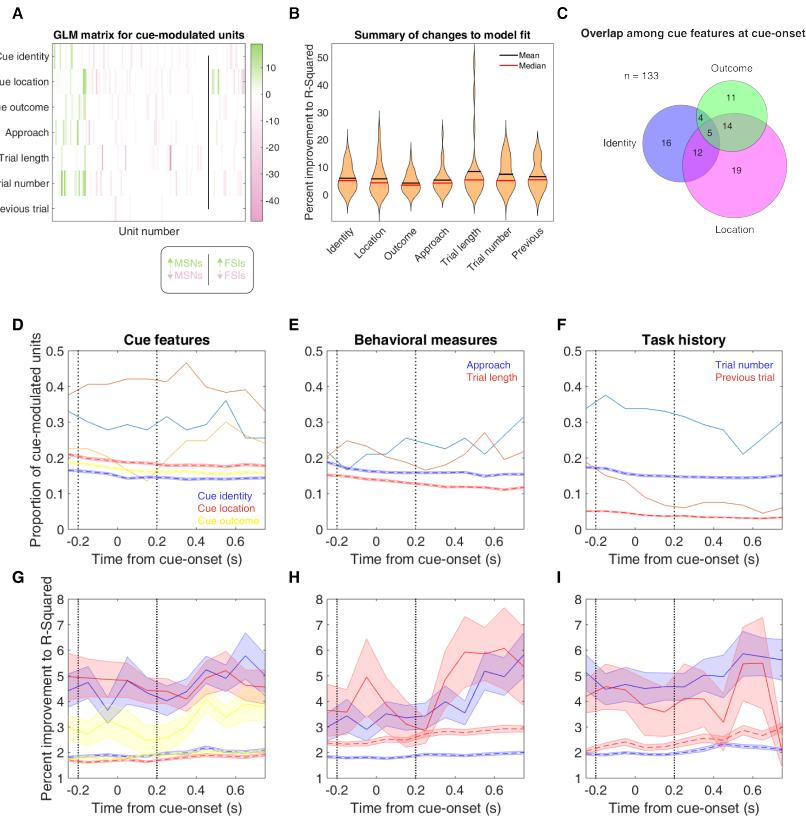


Figure 5: Summary of influence of various task parameters on cue-modulated NAc units after cue-onset. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank 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 G, but for trial number and previous trial.

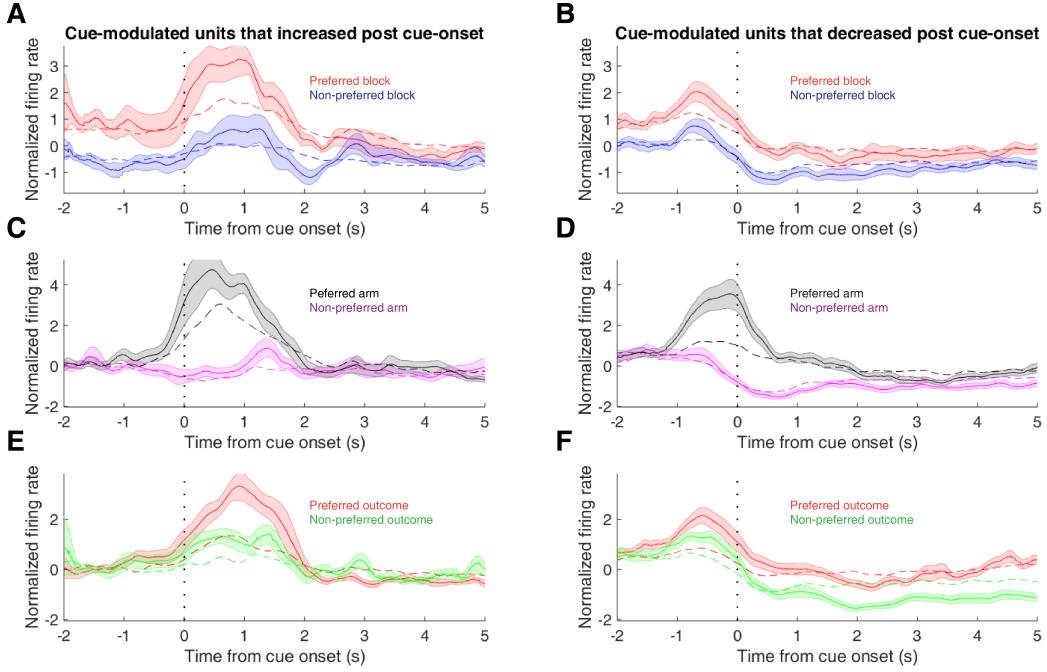


Figure 6: Population-level averages of cue feature sensitive NAc units. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to cue-onset. Activity is plotted for preferred stimulus block (red) and non-preferred stimulus block (blue). Dashed vertical line indicates onset of cue. Dashed color lines indicate the result of shuffling the identity of the units used for this average 1000 times. Lightly shaded area indicates standard error of the mean. Note larger increase to preferred stimulus block over nonpreferred stimulus block. Black lines indicate the average of 1000 rounds of random sampling of units from the non-drifting population for the preferred and non-preferred blocks. **B:** Same as A but for units that decreased in response to cue. Note population level activity reveals units classified as decreasing in response to cue show a biphasic response at the population level, with a transient increase around the time the rat starts on the arm, followed by a minimum after cue onset. Also, note the sustained difference in firing between the two blocks. **C-D:** Same as A-B for cue location. Activity is plotted for most preferred arm (black) and least preferred arm (magenta). **E-F:** Same as A-B for cue outcome. Activity is plotted for preferred expected outcome (red), and nonpreferred outcome (green). Note the larger increase to the cue representing the units preferred outcome (E), and the sustained decrease to the nonpreferred outcome (F).

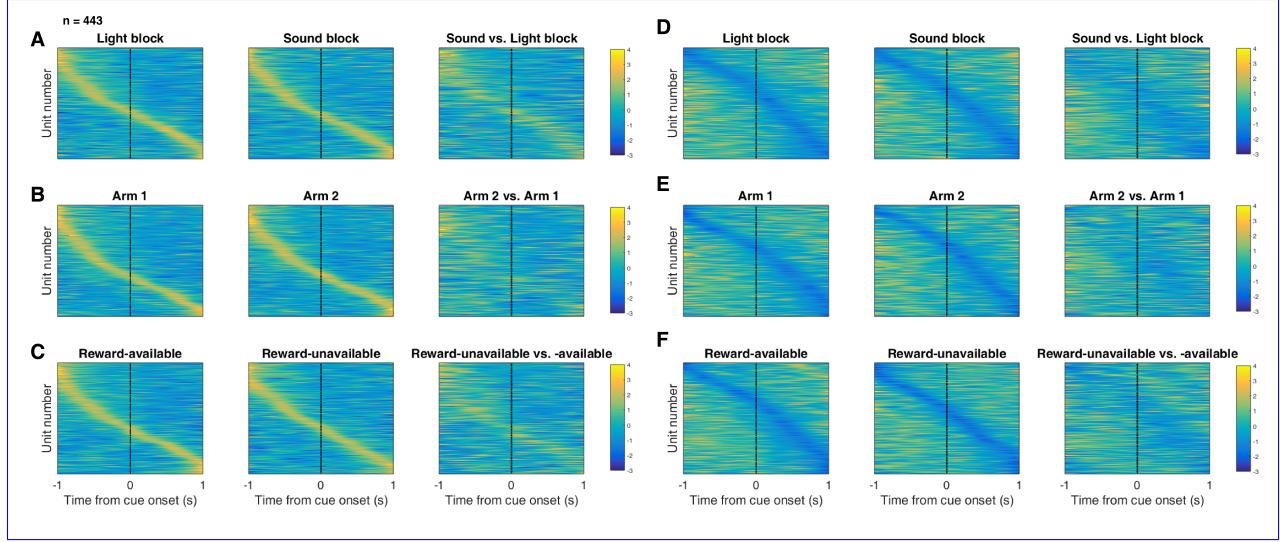


Figure 7: Distribution of NAc firing rates across time surrounding cue onset. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue onset), averaged across all trials for a specific cue type, indicated by text labels. **A-C:** Heat plots aligned to normalized peak firing rates. **A**, left: Heat plot showing smoothed normalized firing activity of all recorded NAc units ordered according to the time of their peak firing rate during the light block. Each row is a units average activity across time to the light block. Dashed line indicates cue onset. Notice the yellow band across time, indicating all aspects of visualized task space were captured by the peak firing rates of various units. **A**, middle: Same units ordered according to the time of the peak firing rate during the sound block. Note that for both blocks, units tile time approximately uniformly with a clear diagonal of elevated firing rates. **A**, right: Unit firing rates taken from the sound block, ordered according to peak firing rate taken from the light block. Note that a weaker but still discernible diagonal persists, indicating partial similarity between firing rates in the two blocks. A similar pattern exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text). **B:** Same layout as in **A**, except that the panels now compare two different locations on the track instead of two cue modalities. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C:** Same layout as in **A**, except that panels now compare reward-available and reward-unavailable trials. **D-F:** Heat plots aligned to normalized minimum firing rates. **D:** Responses during different stimulus blocks as in **A**, but with units ordered according to the time of their minimum firing rate. **E:** Responses during trials on different arms as in **B**, but with units ordered by their minimum firing rate. **F:** Responses during cues signalling different outcomes as in **C**, but with units ordered by their minimum firing rate. Overall, NAc units "tiled" experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

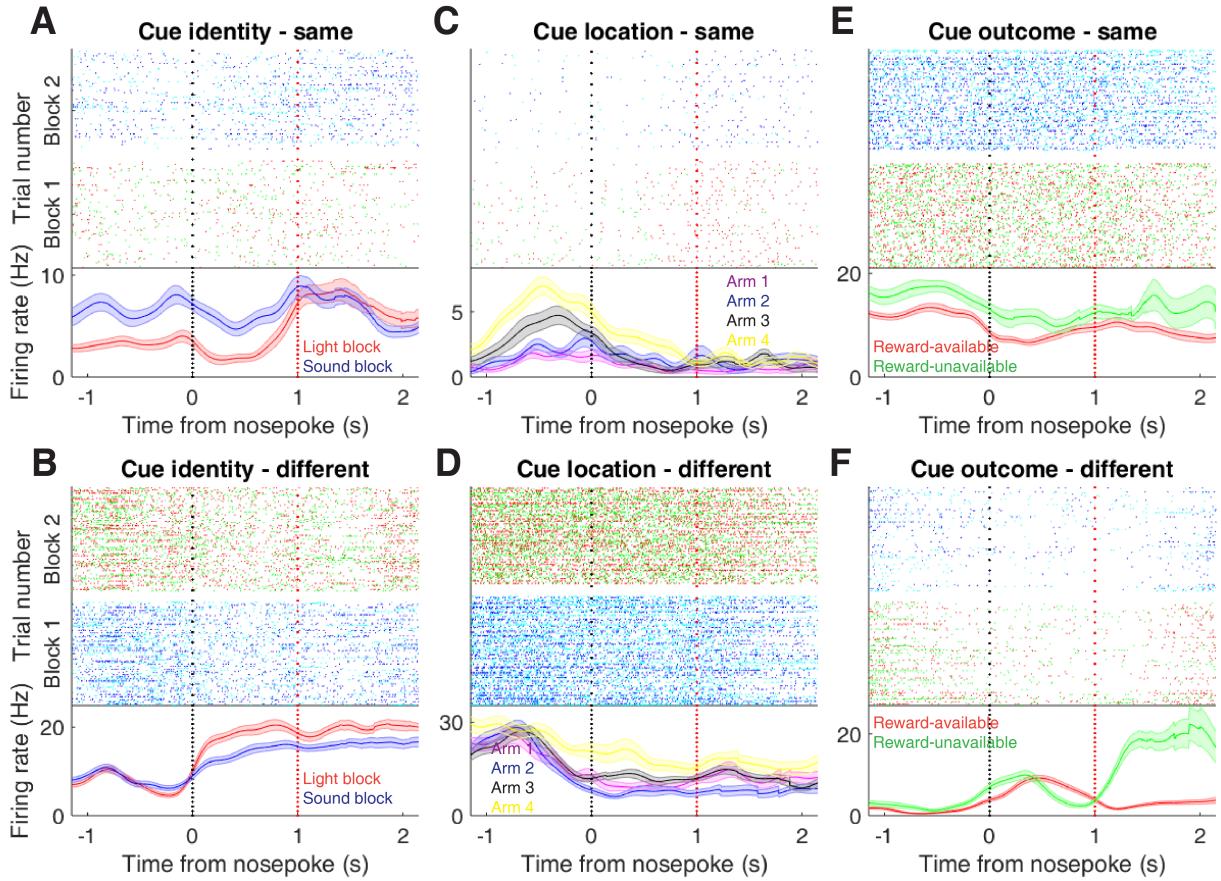


Figure 8: Examples of cue-modulated NAc units influenced by various task parameters at time of nosepoke. **A:** Example of a cue-modulated NAc unit that encoded cue identity at both cue-onset and during nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to nosepoke. 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.

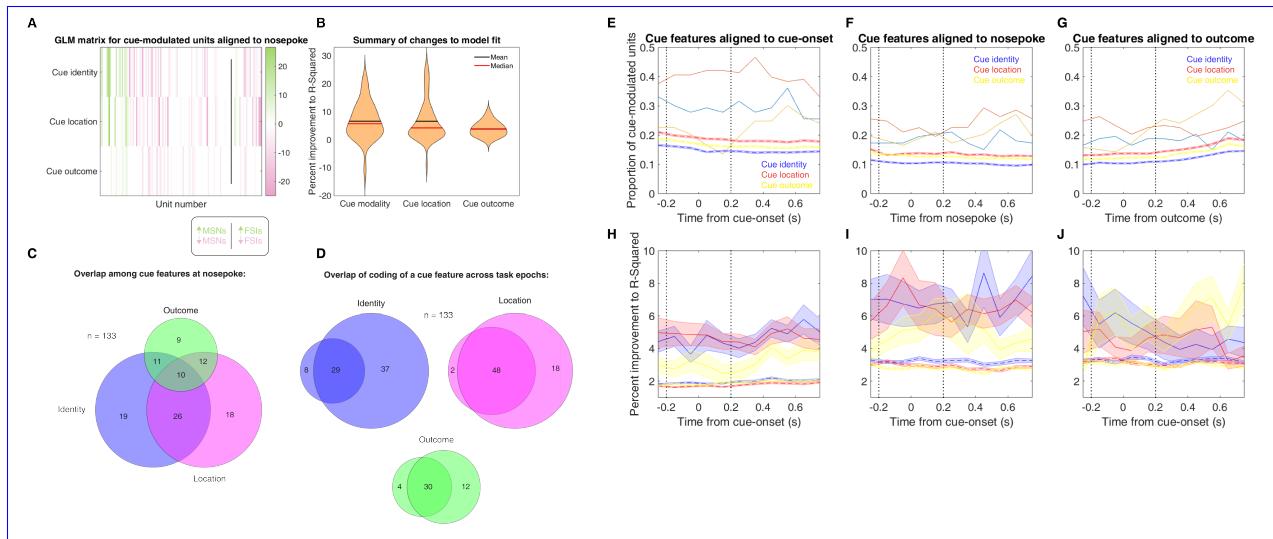


Figure 9: Summary of influence of various task parameters on cue-modulated NAc units during nosepoke. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank test. Each row represents a given task parameter, and each column corresponds to a single unit. Colors indicate how much of the firing rate variance an individual predictor contributed to the model, as measured by differences in R-squared between the final model and the model minus the predictor of Interest. Ordering from left to right: MSNs that increased firing in response to the cue (green, left of line), MSNs with a decreasing response (red, left of line), FSIs with an increasing response (green, right of line), FSIs with a decreasing response (red, right of line). Darker shades indicate more firing rate variance explained by a given predictor. Black line indicates separation of MSNs and FSIs. **B:** Violin plots demonstrating changes in R-squared values with the addition of each of the individual predictors. The mean, median, and distribution of changes in R-squared values is plotted for each of the three task parameters that were significant predictors in the GLM. **C:** Venn diagram illustrating the number of cue-modulated units encoding cue identity (blue circle), cue location (green circle), cue outcome (pink circle), as well as the overlap among units that encoded multiple cue features. **D:** Venn diagram illustrating the number of cue-modulated units encoding cue identity, cue location, and cue outcome during cue-onset (left circle), during nosepoke hold (right circle), and during both epochs (overlap). **E-G:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of cue-onset (E), nosepoke (F), and outcome (G). **E:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), cue location (red solid line), and cue outcome (yellow solid line) significantly contributed to the model at various time epochs relative to cue-onset. Dashed colored lines indicate the average of shuffling the firing rate order that went into the GLM 100 times. Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the GLM. **F:** Same as E, but for time epochs relative to nosepoke where the rat waited for the outcome. **G:** Same as E, but for time epochs relative to receipt of outcome after the rat got feedback about his approach. **H-J:** Average improvement to model fit. **H:** Average percent improvement to R-squared for units where cue identity, cue location, or cue outcome were significant contributors to the final model for time epochs relative to cue-onset. Shaded area around mean represents the standard error of the mean. **I:** Same as H, but for time epochs relative to nosepoke. **J:** Same H, but for time epochs relative to receipt of outcome.

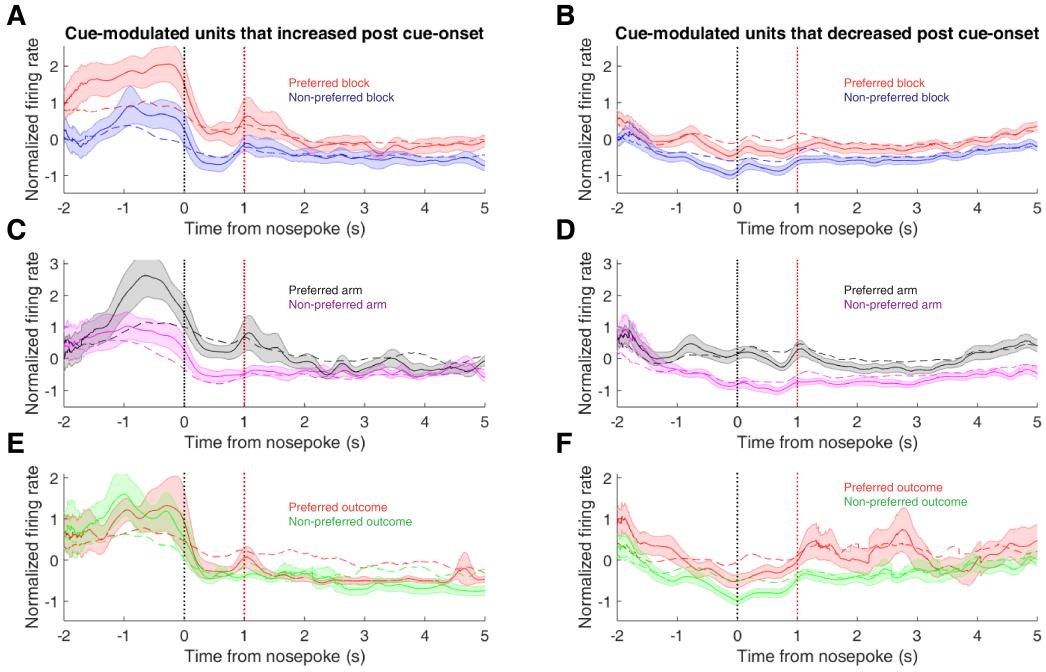


Figure 10: Population-level averages of cue feature sensitive NAc units during a nosepoke. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to nosepoke with reward delivery occurring 1 s after nosepoke. Activity is plotted for preferred stimulus block (red) and nonpreferred stimulus block (blue). Black vertical dashed line indicates nosepoke. Red vertical dashed line indicates reward delivery occurring 1 s after nosepoke for reward-available trials. Dashed color lines indicate the result of shuffling the identity of the units used for this average 1000 times. Lightly shaded area indicates standard error of the mean. Note larger increase leading up to nosepoke to preferred stimulus block over nonpreferred stimulus block. **B:** Same as A but for units that decreased in firing. Note the sustained difference in firing between the two blocks. **C-D:** Same as A-B for cue location. Activity is plotted for most preferred arm (black) and least preferred arm (magenta). **E-F:** Same as A-B for cue outcome. Activity is plotted for preferred expected outcome (red), and nonpreferred outcome (green). Note the peak after outcome receipt for preferred outcome in decreasing units (F).

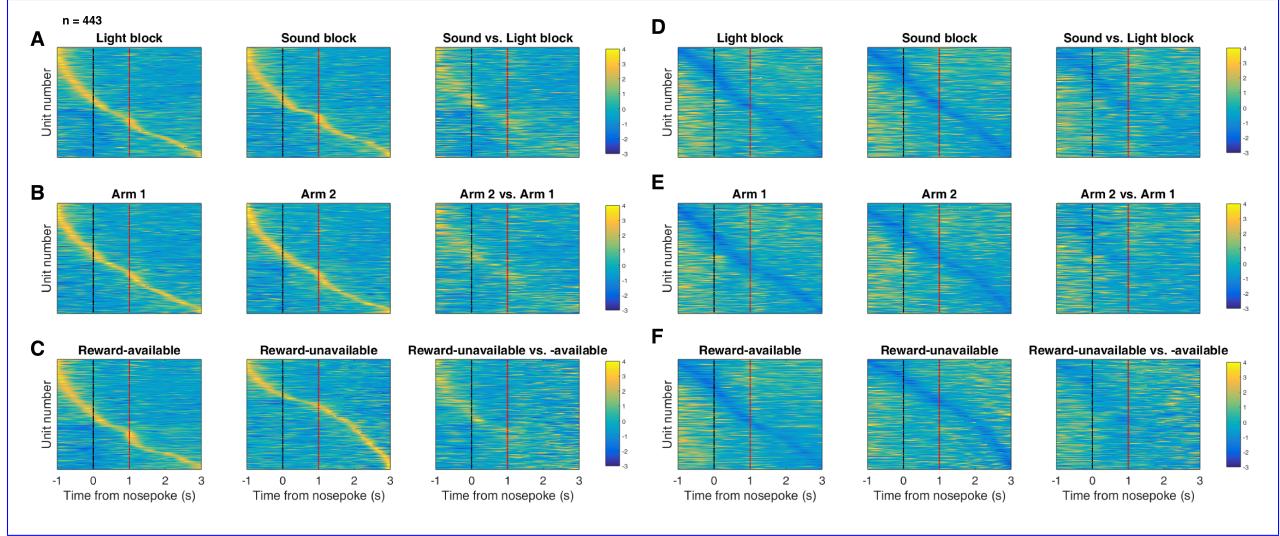


Figure 11: Distribution of NAc firing rates across time surrounding nosepoke for approach trials. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates nosepoke), averaged across all approach trials for a specific cue type, indicated by text labels. **A-C:** Heat plots aligned to normalized peak firing rates. **A, far left:** Heat plot showing smoothed normalized firing activity of all recorded NAc units ordered according to the time of their peak firing rate during the light block. Each row is a units average activity across time to the light block. Black dashed line indicates nosepoke. Red dashed line indicates reward delivery occurring 1 s after nosepoke for reward-available trials. Notice the yellow band across time, indicating all aspects of visualized task space were captured by the peak firing rates of various units. **A, middle:** Same units ordered according to the time of the peak firing rate during the sound block. Note that for both blocks, units tile time approximately uniformly with a clear diagonal of elevated firing rates, and a clustering around outcome receipt. **A, right:** Unit firing rates taken from the sound block, ordered according to peak firing rate taken from the light block. Note that a weaker but still discernible diagonal persists, indicating partial similarity between firing rates in the two blocks. A similar pattern exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text). **B:** Same layout as in A, except that the panels now compare two different locations on the track instead of two cue modalities. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C:** Same layout as in A, except that panels now compare correct reward-available and incorrect reward-unavailable trials. The disproportionate tiling around outcome receipt for reward-available, but not reward-unavailable trials suggests encoding of reward receipt by NAc units. **D-F:** Heat plots aligned to normalized minimum firing rates. **D:** Responses during different stimulus blocks as in A, but with units ordered according to the time of their minimum firing rate. **E:** Responses during trials on different arms as in B, but with units ordered by their minimum firing rate. **F:** Responses during cues signalling different outcomes as in C, but with units ordered by their minimum firing rate. Overall, NAc units "tiled" experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

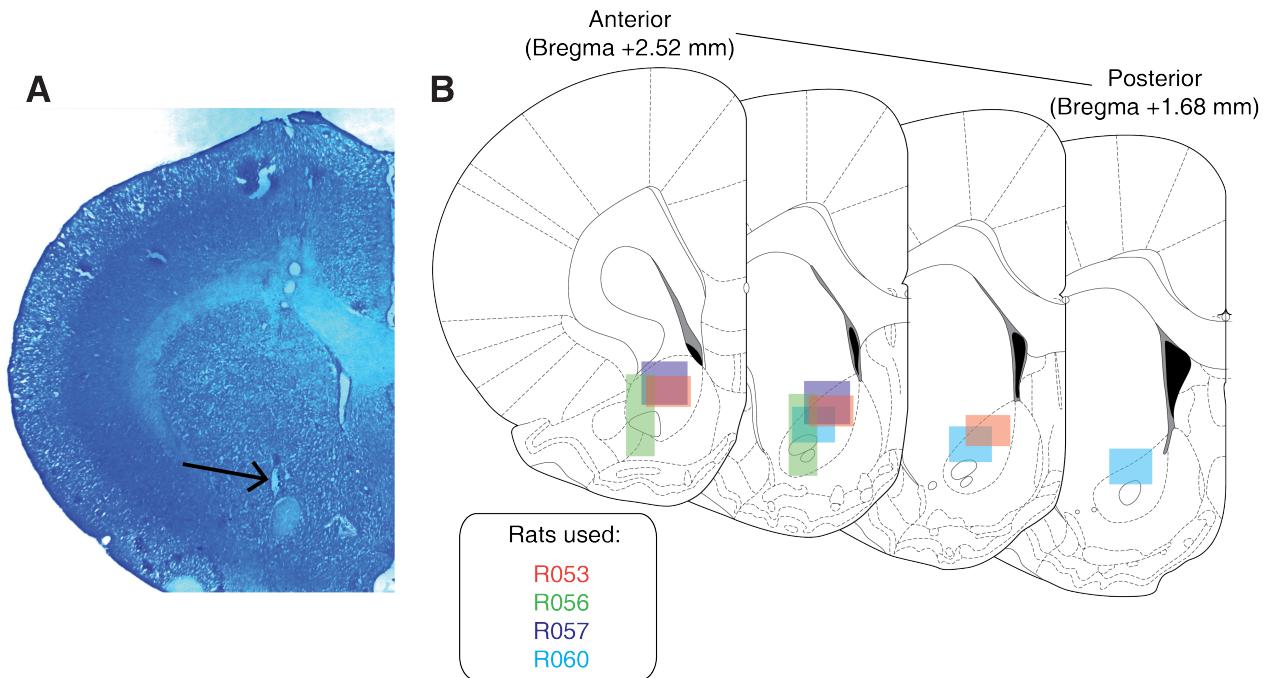


Figure 12: Histological verification of recording sites. Upon completion of experiments, brains were sectioned and tetrode placement was confirmed. **A:** Example section from R060 showing a recording site in the NAc core just dorsal to the anterior commissure (arrow). **B:** Schematic showing recording areas for all subjects.

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

Table 1: Units overview