

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

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

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

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

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

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

## **1 Abstract**

2 The nucleus accumbens (NAc) is important for learning from feedback, and for biasing and invigorating  
3 behavior in response to cues that predict motivationally relevant outcomes. NAc encodes outcome-related  
4 cue features such as the magnitude and identity of reward. However, little is known about how features  
5 of cues themselves are encoded. We designed a decision making task where rats learned multiple sets  
6 of outcome-predictive cues, and recorded single-unit activity in the NAc during performance. We found  
7 that coding of cue identity and location occurred alongside coding of expected outcome. Furthermore, this  
8 coding persisted both during a delay period, after the rat made a decision and was waiting for an outcome,  
9 and after the outcome was revealed. Encoding of cue features in the NAc may enable contextual modulation  
10 of ongoing behavior, and provide an eligibility trace of outcome-predictive stimuli for updating stimulus-  
11 outcome associations to 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 reward prediction errors (RPEs) (Averbeck  
15 & Costa 2017; Joel et al. 2002; Khamassi & Humphries 2012; Lee et al. 2012; Maia 2009; Schultz 2016; see  
16 also the addiction literature on effects of drug rewards; Carelli 2010; Hyman et al. 2006; Kalivas & Volkow  
17 2005), and others a role in the modulation of ongoing behavior through stimuli associated with motivation-  
18 ally relevant outcomes (invigorating, directing; Floresco, 2015; Nicola, 2010; Salamone & Correa, 2012).

19 These proposals echo similar ideas on the functions of the neuromodulator dopamine (Berridge, 2012; Maia,  
20 2009; Salamone & Correa, 2012; Schultz, 2016), with which the NAc is tightly linked functionally as well  
21 as anatomically (Cheer et al., 2007; du Hoffmann & Nicola, 2014; Ikemoto, 2007; Takahashi et al., 2016).

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

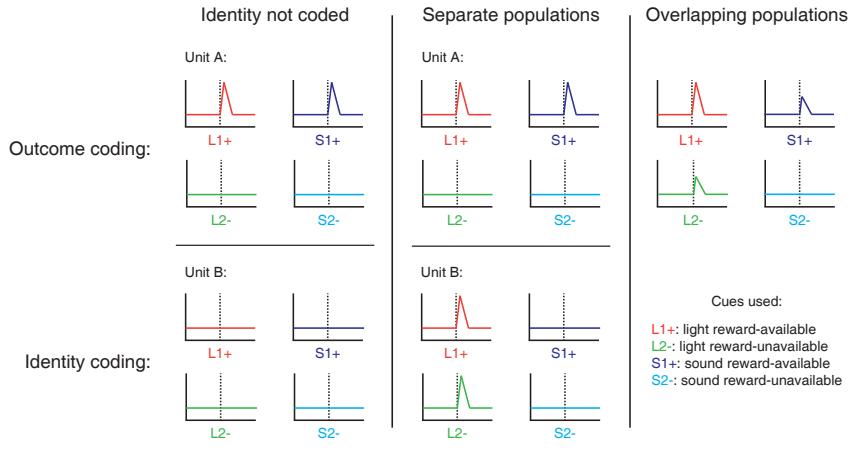
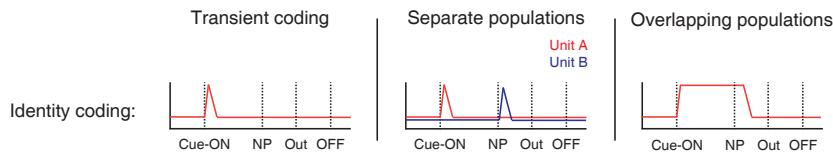
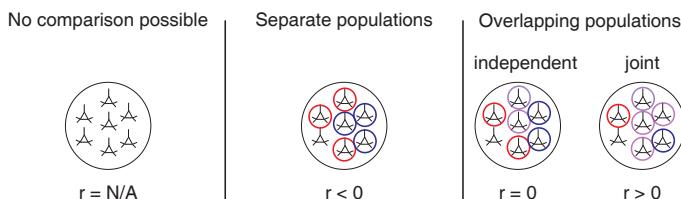
32 Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues  
33 (Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Cromwell & Schultz, 2003; Day et al., 2006;  
34 Goldstein et al., 2012; Hassani et al., 2001; Hollerman et al., 1998; Lansink et al., 2012; McGinty et al., 2013;  
35 Nicola, 2004; Roesch et al., 2009; Roitman et al., 2005; Saddoris et al., 2011; Schultz et al., 1992; Setlow et

36 al., 2003; Sugam et al., 2014; West & Carelli, 2016), much less is known about how outcome-predictive cues  
37 themselves are encoded in the NAc (but see; Sleezer et al., 2016). This is an important issue for at least two  
38 reasons. First, in reinforcement learning, motivationally relevant outcomes are typically temporally delayed  
39 relative to the cues that predict them. In order to solve the problem of assigning credit (or blame) across such  
40 temporal gaps, some trace of preceding activity needs to be maintained (Lee et al., 2012; Sutton & Barto,  
41 1998). For example, if you become ill after eating food X in restaurant A, depending on if you remember the  
42 identity of the restaurant or the food at the time of illness, you may learn to avoid all restaurants, restaurant  
43 A only, food X only, or the specific pairing of X-in-A. Therefore, a complete understanding of what is  
44 learned following feedback requires understanding what trace is maintained. Since NAc is a primary target  
45 of dopamine signals interpretable as RPEs, and NAc lesions impair RPEs related to timing, its activity trace  
46 will help determine what can be learned when RPEs arrive (Hamid et al., 2015; Hart et al., 2014; Ikemoto,  
47 2007; McDannald et al., 2011; Takahashi et al., 2016). Similarly, in a neuroeconomic framework, NAc is  
48 thought to represent a domain-general subjective value signal for different offers (Bartra et al., 2013; Levy  
49 & Glimcher, 2012; Peters & Büchel, 2009; Sescousse et al., 2015); having a representation of the offer itself  
50 alongside this value signal would provide a potential neural substrate for updating offer value.

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 et al., 2014), and unsignaled rule changes, as occurs in set shifting and other cogni-  
54 tive control tasks (Cohen & Servan-Schreiber, 1992; Floresco et al., 2006; Grant & Berg, 1948; Sleezer et  
55 al., 2016). In such situations, the question arises how selective, context-dependent processing of outcome-  
56 predictive cues is implemented. For instance, is there a gate prior to NAc such that only currently relevant  
57 cues are encoded in NAc, or are all cues represented in NAc but their current values dynamically updated  
58 (FitzGerald et al., 2014; Goto & Grace, 2008; Sleezer et al., 2016)? Representation of cue identity would  
59 allow for context-dependent mapping of outcomes predicted by specific cues.

60 Thus, both from a learning and a flexible performance perspective, it is of interest to determine how cue iden-

61 tity is represented in the brain, with NAc of particular interest given its anatomical and functional position at  
62 the center of motivational systems. We sought to determine whether cue identity is represented in the NAc,  
63 if cue identity is represented alongside other motivationally relevant variables, such as cue outcome, and if  
64 these representations are maintained after a behavioral decision has been made (see Figure 1 for a schematic  
65 representation of the specific hypotheses tested). To address these questions, we recorded the activity of NAc  
66 units as rats performed a task in which multiple, distinct sets of cues predicted the same outcome.

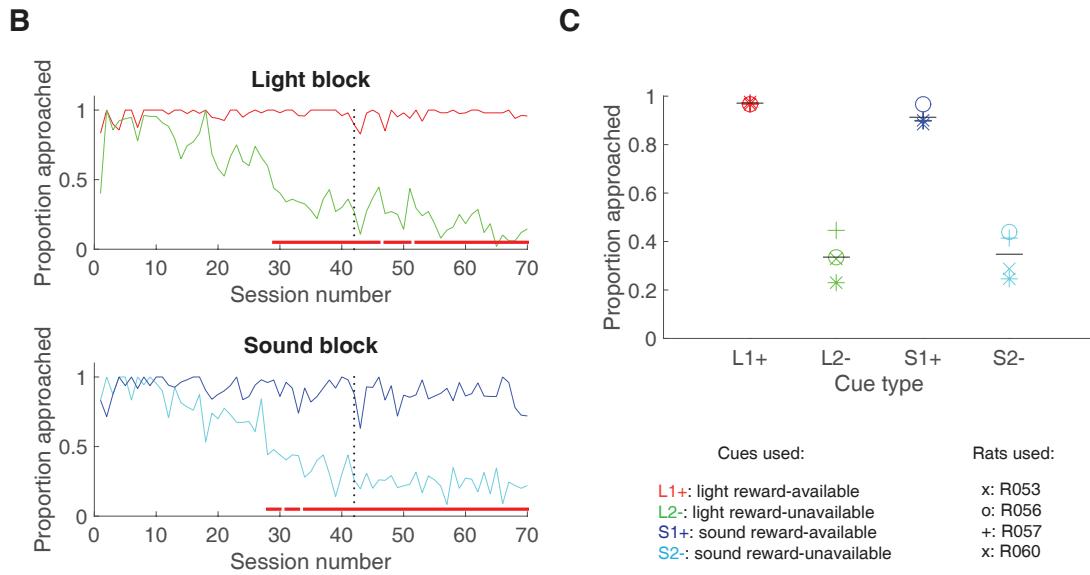
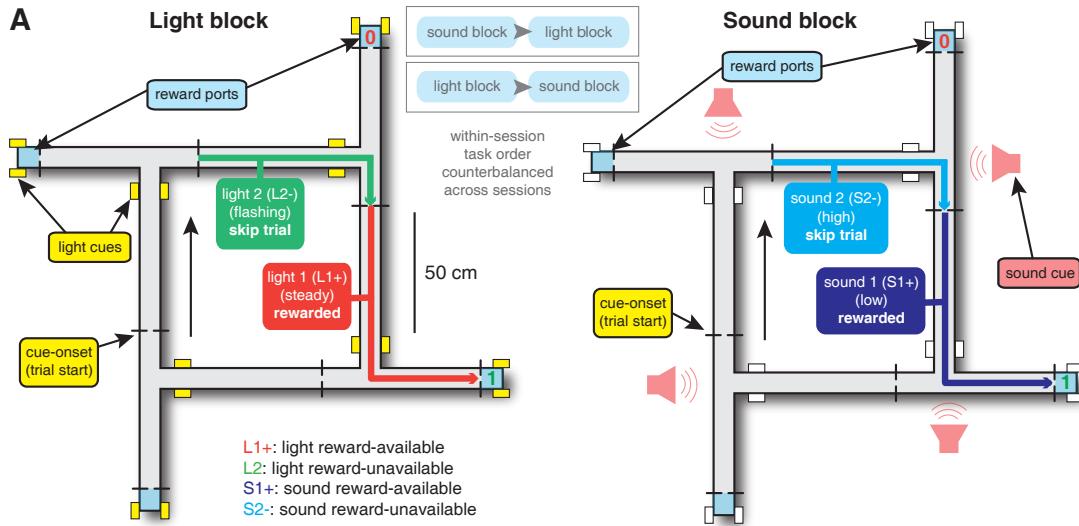
**A****Presence of cue identity coding****B****Persistence of cue identity coding****C****Quantification of coding across units and time epochs**

**Figure 1:** Schematic of hypothetical coding scenarios for cue feature coding employed by single units in the NAc across different cue features (A) and phases of a trial (B). **A:** Displayed are schematic peri-event time histograms (PETHs) illustrating putative responses to different cues under different hypotheses of how cue identity (light, sound; L, S) and outcome (reward-available, reward-unavailable; +, -) are coded. Left panel: Coding of 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 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. Middle panel: Coding of identity occurs in a separate population of units from coding of other cue features such as expected outcome or physical location. Top: Same as left panel, 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. Note lack of coding overlap in both units. Right panel: Coding of identity occurs in an overlapping population of cells with coding of other motivationally relevant variables. Hypothetical example demonstrating a unit that responds to reward-available cues, but firing rate is also modulated by the stimulus modality of the cue, firing most for the reward-available light cue. **B:** Displayed are schematic PETHs illustrating potential ways in which identity coding may persist over time. 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 cue-onset, 'NP' signifies nosepoke at a reward receptacle, 'Out' signifies when the outcome is revealed, 'OFF' signifies cue-offset. Middle and right panel: Identity coding persists at other time points, shown here during a nosepoke hold period until outcome is revealed. Coding can either be maintained by a sequence of units (middle panel) or by the same unit as during cue-onset (right panel). **C:** Schematic pool of NAc units, illustrating different analysis outcomes that discriminate between hypotheses.  $R$  values represent the correlation between sets of recoded regression coefficients (see text for analysis details). Left panel: Cue identity is not coded (A: left panel), or is only transiently represented in response to the cue (B: left panel). Middle panel: Negative correlation ( $r < 0$ ) suggests that identity and outcome coding are represented by separate populations of units (A: middle panel), or identity coding is represented by distinct units across different points in a trial (B: middle panel). Red circles represent coding for one cue feature or point in time, blue circles for the other cue feature or point in time. Right panel: Identity and outcome coding (A: right panel), or identity coding at cue-onset and nosepoke (B: right panel) are represented by overlapping populations of units, shown here by the purple circles. The absence of a correlation ( $r = 0$ ) suggests that the overlap of identity and outcome coding, or identity coding at cue-onset and nosepoke, is expected by chance and that the two cue features, or points in time, are coded by overlapping but independent populations from one another. A positive correlation ( $r > 0$ ) implies a higher overlap than expected by chance, suggesting coding by a joint population. Note: The same logic applies to other aspects of the environment when the cue is presented, such as the physical location of the cue, as well as other time epochs within the task, such as when the animal receives feedback about an approach.

67 **Results**

68 **Behavior**

69 Rats were trained to discriminate between cues signaling the availability and absence of reward on a square  
70 track with four identical arms for two distinct set of cues (Figure 2A). During each session, rats were pre-  
71 sented sequentially with two behavioral blocks containing cues from different sensory modalities, a light and  
72 a sound block, with each block containing a cue that signaled the availability of reward (reward-available),  
73 and a cue that signaled the absence of reward (reward-unavailable). To maximize reward receipt, rats should  
74 approach reward sites on reward-available trials, and skip reward sites on reward-unavailable trials (see Fig-  
75 ure 2B for an example learning curve). All four rats learned to discriminate between the reward-available and  
76 reward-unavailable cues for both the light and sound blocks as determined by reaching significance ( $p < .05$ )  
77 on a daily chi-square test comparing approach behavior for reward-available and reward-unavailable cues for  
78 each block, for at least three consecutive days (range for time to criterion: 22 - 57 days). Maintenance of  
79 behavioral performance during recording sessions was assessed using linear mixed effects models for pro-  
80 portion of trials where the rat approached the receptacle. Analyses revealed that the likelihood of a rat to  
81 make an approach was influenced by whether a reward-available or reward-unavailable cue was presented,  
82 but was not significantly modulated by whether the rat was presented with a light or sound cue (Percent-  
83 age approached: light reward-available = 97%; light reward-unavailable = 34%; sound reward-available =  
84 91%; sound reward-unavailable 35%; cue identity  $p = .115$ ; cue outcome  $p < .001$ ; Figure 2C). Additional  
85 analyses separated each block into two halves to assess possible within session learning. Adding block half  
86 into the model did not improve prediction of behavioral performance ( $p = .86$ ), arguing against within ses-  
87 sion learning. Thus, rats successfully discriminated the cues according to whether or not they signaled the  
88 availability of reward at the reward receptacle.



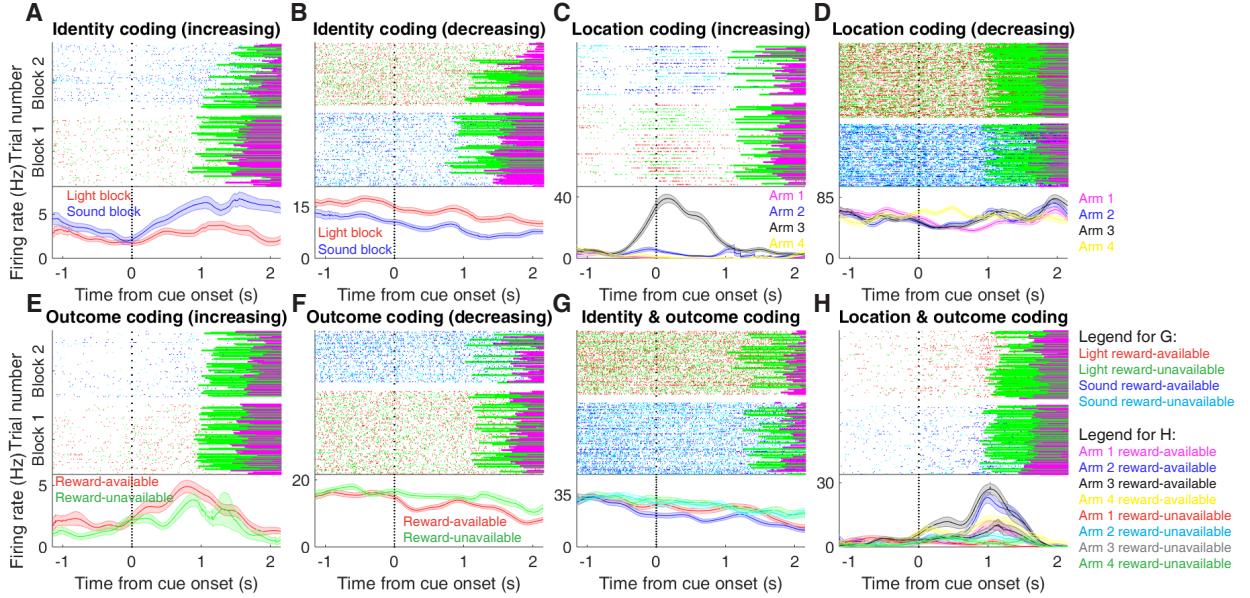
**Figure 2:** Schematic and performance of the behavioral task. **A:** Apparatus was a square track consisting of multiple identical T-choice points. At each choice point, the availability of 12% sucrose reward at the nearest reward receptacle (light blue fill) was signaled by one of four possible cues, 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. Arrows outside of track indicate correct running direction. Left: light block showing an example trajectory for a correct reward-available (approach trial; red) and reward-unavailable (skip trial; green) trial. Rectangular boxes with yellow fill indicate location of LEDs used for light cues. Right: sound block with a correct reward-available (approach trial; navy blue) and reward-unavailable (skip trial; light blue) trial. Speakers for sound cues were placed underneath the choice points, indicated by magenta speaker icons. 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. **B-C:** Performance on the behavioral task. **B:** Example learning curves across sessions from a single subject (R060) showing the proportion of trials 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 reward-unavailable trials. Red bars indicate days in which a rat statistically discriminated between reward-available and reward-unavailable cues, determined by a chi square test. Dashed line indicates time of electrode implant surgery. **C:** Proportion of trials approached for each cue, averaged across all recording sessions and shown for each rat. 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 (~90% approached) and reward-unavailable cues (~30% approached), for both blocks (see Results for statistics).

89 **NAc encodes behaviorally relevant and irrelevant cue features**

90 We sought to address which parameters of our task were encoded by NAc activity, specifically whether the  
91 NAc encodes aspects of motivationally relevant cues not directly tied to reward, such as the identity and  
92 location of the cue, and whether this coding is accomplished by separate or overlapping populations (Figure  
93 1A). To do this we recorded a total of 443 units with > 200 spikes in the NAc from 4 rats over 57 sessions  
94 (range: 12 - 18 sessions per rat) while they performed a cue discrimination task (Table 1). Units that exhibited  
95 a drift in firing rate over the course of either block, as measured by a Mann-Whitney U test comparing firing  
96 rates for the first and second half of trials within a block, were excluded from further analysis, leaving  
97 344 units for further analysis. The activity of 133 (39%) of these 344 units were modulated by the cue, as  
98 determined by comparing 1 s pre- and post-cue activity with a Wilcoxon signed-rank test, with more showing  
99 a decrease in firing ( $n = 103$ ) than an increase ( $n = 30$ ) around the time of cue-onset (Table 1). Within this  
100 group, 24 were classified as putative fast spiking interneurons (FSIs), while 109 were classified as putative  
101 medium spiny neurons (MSNs). Upon visual inspection, we observed several patterns of firing activity,  
102 including units that discriminated firing upon cue-onset across various cue conditions, showed sustained  
103 differences in firing across cue conditions, had transient responses to the cue, showed a ramping of activity  
104 starting at cue-onset, and showed elevated activity immediately preceding cue-onset (Figure 3, supplement  
105 1, supplement 2).

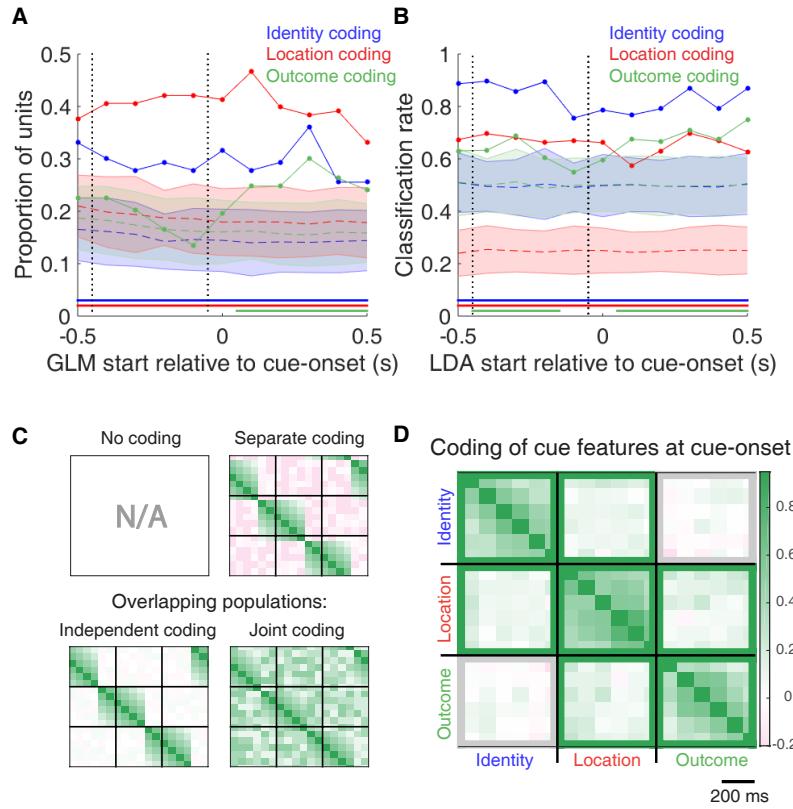
<b>Task parameter</b>	<b>Total</b>	$\uparrow$ <b>MSN</b>	$\downarrow$ <b>MSN</b>	$\uparrow$ <b>FSI</b>	$\downarrow$ <b>FSI</b>
All units	443	155	216	27	45
<i>Rat ID</i>					
R053	145	51	79	4	11
R056	70	12	13	17	28
R057	136	55	75	3	3
R060	92	37	49	3	3
Analyzed units	344	117	175	18	34
Cue modulated units	133	24	85	6	18
<i>GLM aligned to cue-onset</i>					
Cue identity	42 (32%)	9 (38%)	25 (29%)	0 (-)	8 (44%)
Cue location	55 (41%)	11 (46%)	33 (39%)	3 (50%)	8 (44%)
Cue outcome	26 (20%)	5 (21%)	15 (18%)	1 (17%)	5 (28%)
Approach behavior	32 (24%)	8 (33%)	19 (22%)	2 (33%)	3 (17%)
Trial length	22 (17%)	5 (21%)	14 (16%)	0 (-)	3 (17%)
Trial number	42 (32%)	11 (46%)	20 (24%)	1 (17%)	10 (56%)
Trial history	8 (6%)	1 (4%)	5 (6%)	0 (-)	1 (6%)
<i>GLM aligned to nosepoke</i>					
Cue identity	28 (21%)	3 (13%)	17 (20%)	2 (33%)	6 (33%)
Cue location	30 (23%)	2 (8%)	21 (25%)	2 (33%)	5 (28%)
Cue outcome	23 (17%)	2 (8%)	14 (16%)	1 (17%)	6 (33%)
<i>GLM aligned to outcome</i>					
Cue identity	25 (19%)	4 (17%)	15 (18%)	2 (33%)	4 (22%)
Cue location	31 (23%)	5 (21%)	23 (27%)	0 (-)	3 (17%)
Cue outcome	34 (26%)	6 (25%)	15 (18%)	4 (67%)	9 (50%)

**Table 1:** Overview of recorded NAc units and their relationship to task variables at various time epochs. Percentage is relative to the number of cue-modulated units (n = 133).



**Figure 3:** Examples of cue-modulated NAc units influenced by different task parameters. **A:** Example of a cue-modulated NAc unit that showed an increase in firing following the cue, and exhibited identity coding. 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: PCTHs 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 exhibited location coding. Each color in the PCTHs 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 rate decreased for this unit on all arms but arm 4. **E-F:** Cue-modulated units that exhibited outcome coding, with the PCTHs 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 showed both identity and outcome coding. **H:** An example of a unit that coded for both identity and location.

106 To characterize more formally whether these cue-modulated responses were influenced by various aspects of  
107 the task, we fit a sliding window generalized linear model (GLM) to the firing rate of each cue-modulated unit  
108 surrounding cue-onset, using a forward selection stepwise procedure for variable selection, a bin size of 500  
109 ms for firing rate and a step size of 100 ms for the sliding window. Fitting GLMs to all trials within a session  
110 revealed that a variety of task parameters accounted for a significant portion of firing rate variance in NAc  
111 cue-modulated units (Figure 4A, supplement 1, supplement 2, Table 1). Notably, a significant proportion  
112 of units discriminated between the light and sound block (*identity coding*: ~32% of cue-modulated units,  
113 accounting for ~5% of firing rate variance) or the arms of the apparatus (*location coding*: ~41% of cue-  
114 modulated units, accounting for ~4% of firing rate variance) throughout the entire window surrounding  
115 cue-onset. Additionally, a substantial proportion of units discriminating between the common portion of  
116 reward-available and reward-unavailable trials (*outcome coding*: ~20% of cue-modulated units, accounting  
117 for ~4% of firing rate variance) was not observed until after the onset of the cue ( $z$ -score  $> 1.96$  when  
118 comparing observed proportion of units to a shuffled distribution obtained when shuffling the firing rates of  
119 each unit across trials before running the GLM). Furthermore, our variable selection method ensured that the  
120 observed coding was not due to potential confounds from other task variables, such as; behavioral response  
121 at the choice point (*approach behavior*; left vs. right), variability in response vigor (*trial length*; see McGinty  
122 et al. 2013), drift due to the passage of time (*trial number*), and the pseudorandom nature of cue presentation  
123 (*trial history*). In addition to accounting for firing rate variance explained due to whether the rat turned left  
124 or right, we reran our cue-onset GLM using only approach trials, and found a similar proportion of outcome  
125 coding units (34 units; ~26% of cue-modulated units), providing further support that these units were coding  
126 the expected outcome of the cue. Taken together, these results from the GLMs suggest that the NAc encodes  
127 features of outcome-predictive cues in addition to expected outcome.



**Figure 4:** Summary of influence of cue features on cue-modulated NAc units at time points surrounding cue-onset. **A:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), location (red solid line), and outcome (green 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. Error bars indicate 1.96 standard deviations from the shuffled mean. Solid lines at the bottom indicate when the proportion of units observed was greater than the shuffled distribution ( $z$ -score  $> 1.96$ ). Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the GLM. **B:** Sliding window LDA (bin size: 500 ms; step size: 100 ms) demonstrating the classification rate for cue identity (blue solid line), location (red solid line), and outcome (green solid line) using a pseudoensemble consisting of the 133 cue-modulated units. Dashed colored lines indicate the average of shuffling the firing rate order that went into the cross-validated LDA 100 times. Solid lines at the bottom indicate when the classifier performance greater than the shuffled distribution ( $z$ -score  $> 1.96$ ). Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the classifier. **C-D:** Correlation matrices testing the presence and overlap of cue feature coding at cue-onset. **C:** Schematic outlining the possible outcomes for coding across cue features at cue-onset, generated by correlating the recoded beta coefficients from the GLMs and comparing to a shuffled distribution (see text for analysis details). Top left: coding is not present, therefore no comparison is possible. Top right: cue features are coded by separate populations of units. Displayed is a correlation matrix with each of the 9 blocks representing correlations for two cue features across the post-cue-onset time bins from the sliding window GLM, with green representing positive correlations ( $r > 0$ ), pink representing negative correlations ( $r < 0$ ), and white representing no correlation ( $r = 0$ ). X- and y-axis have the same axis labels, therefore the diagonal represents the correlation of a cue feature against itself at that particular time point ( $r = 1$ ). Here the large amount of pink in the off-diagonal elements suggests that coding of cue features occur separately from one another. Bottom left: Coding of cue features occurs in overlapping but independent populations of units, shown here by the abundance of white and relative lack of green and pink in the off-diagonal elements. Bottom right: Coding of cue features occurs in a joint overlapping population, shown here by the large amount of green in the off-diagonal elements. **D:** Correlation matrix showing the correlation among cue identity, location, and outcome coding surrounding cue-onset. The window of GLMs used in each block is from cue-onset to the 500 ms window post-cue-onset, in 100 ms steps. Each individual value is for a sliding window GLM within that range, with the scale bar contextualizing step size. Color bar displays relationship between correlation value and color. Colored square borders around each block indicate the result of a comparison of the mean correlation of a block to a shuffled distribution, with pink indicating separate populations ( $z$ -score  $< -1.96$ ), grey indicating overlapping but independent populations, and green indicating joint overlapping populations ( $z$ -score  $> 1.96$ ).

128 To assess what information may be encoded at the population level, we trained a classifier on a pseudoensemble  
129 of the 133 cue-modulated units (Figure 4B). Specifically, we used the firing rate of each unit for each  
130 trial as an observation, and different cue conditions as trial labels (e.g. light block, sound block). A lin-  
131 ear discriminant analysis (LDA) classifier with 10-fold cross-validation could correctly predict a trial above  
132 chance levels for the identity and location of a cue across all time points surrounding cue-onset ( $z$ -score  $>$   
133 1.96 when comparing classification accuracy of data versus a shuffled distribution), whereas the ability to  
134 predict whether a trial was reward-available or reward-unavailable (outcome coding) was not significantly  
135 higher than the shuffled distribution for the time point containing 500 ms of pre-cue firing rate, and increased  
136 gradually as a trial progressed, providing evidence that cue information is also present in the pseudoensemble  
137 level.

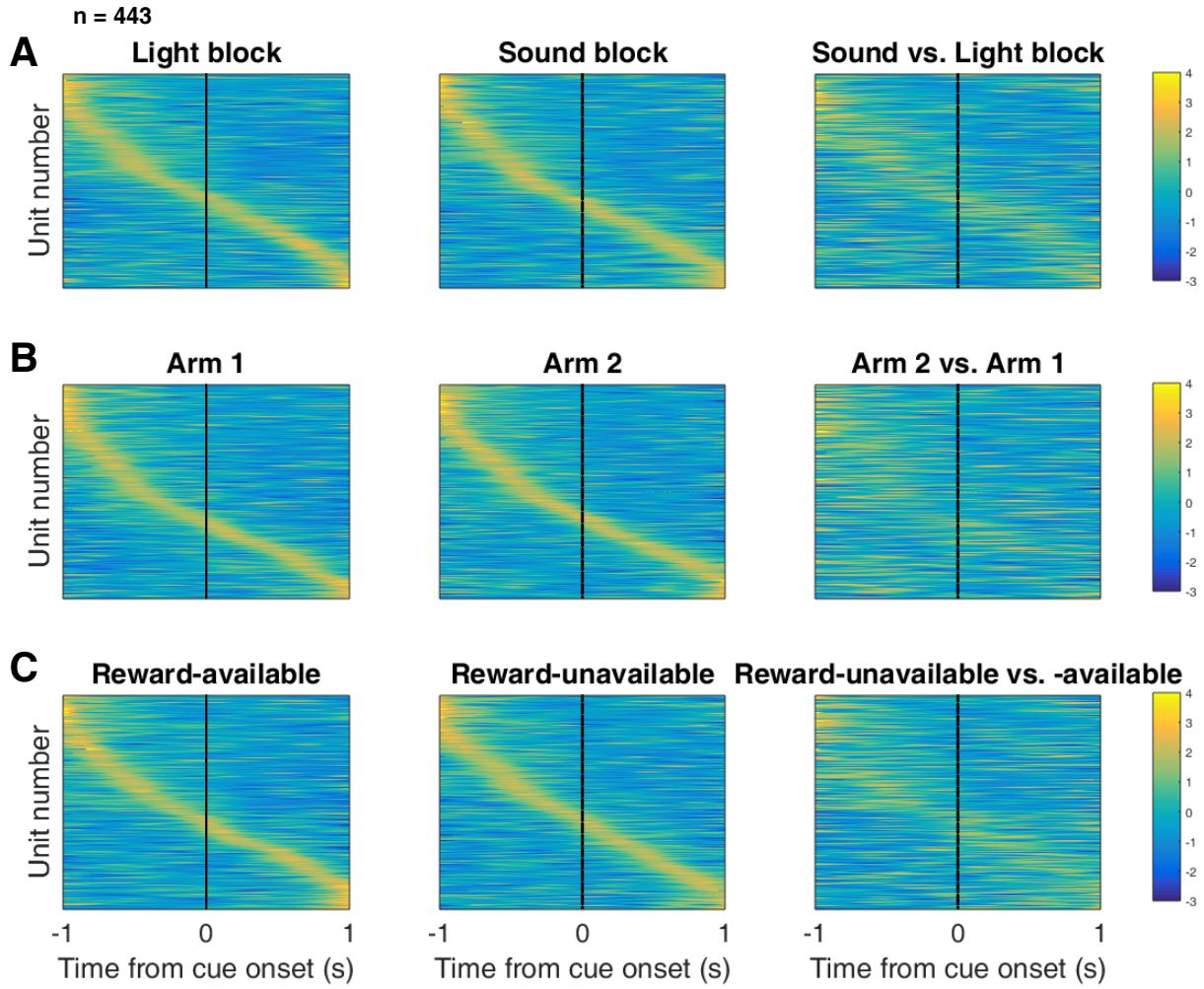
138 To quantify the overlap of cue feature coding we correlated recoded beta coefficients from the GLMs, as-  
139 signing a value of ‘1’ if a cue feature was a significant predictor for that unit and ‘0’ if not, and calculated  
140 a z-score comparing the mean of the obtained correlations to the mean and standard deviation of a shuffled  
141 distribution, generated by shuffling the unit ordering within an array (Figure 1A,C, 4C,D). This revealed that  
142 identity was coded independently from outcome (mean  $r = .009$ ,  $z$ -score = 0.81), and by a joint population  
143 with location (mean  $r = .097$ ,  $z$ -score = 6.61), while location and outcome were coded by a joint popula-  
144 tion of units (mean  $r = .119$ ,  $z$ -score = 8.07). Together, these findings show that various cue features are  
145 represented in the NAc at both the single-unit and pseudoensemble level, with location being coded by joint  
146 populations with identity and outcome, but that identity is coded independently from outcome.

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

148 Next, we sought to determine how coding of cue features evolved over time. Two main possibilities can  
149 be distinguished (Figure 1B); a unit coding for a feature such as cue identity could remain persistently  
150 active, or a progression of distinct units could activate in sequence. To visualize the distribution of responses  
151 throughout our task space and test if this distribution is modulated by cue features, we z-scored the firing rate

152 of each unit, plotted the normalized firing rates of all units aligned to cue-onset, and sorted them according  
153 to the time of peak firing rate (Figure 5). We did this separately for both the light and sound blocks, and  
154 found a nearly uniform distribution of firing fields in task space that was not limited to alignment to the  
155 cue (Figure 5A). Furthermore, to determine if this population level activity was similar across blocks, we  
156 also organized firing during the sound blocks according to the ordering derived from the light blocks. This  
157 revealed that while there was some preservation of order, the overall firing was qualitatively different across  
158 the two blocks, implying that population activity distinguishes between light and sound blocks.

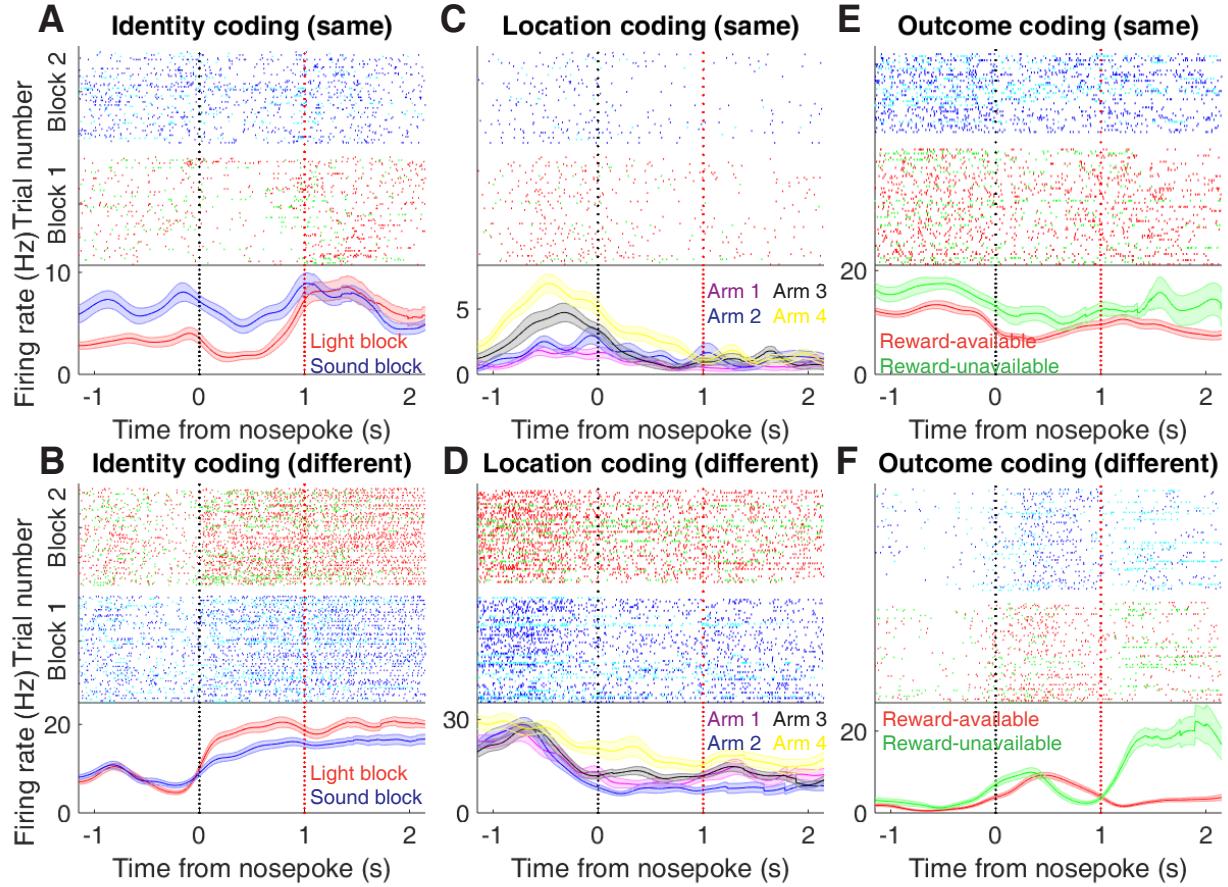
159 To control for the possibility that any comparison of trials would produce this effect, we divided each block  
160 into two halves and looked at the correlation of the average smoothed firing rates across various combina-  
161 tions of these halves across our cue-onset centered epoch to see if the across block comparisons were less  
162 correlated than the within block correlations. A linear mixed effects model revealed that within block corre-  
163 lations (e.g. one half of light trials vs other half of light trials) were higher and more similar than across block  
164 correlations (e.g. half of light trials vs half of sound trials) suggesting that activity in the NAc discriminates  
165 across light and sound blocks (mean within block correlation = .381; mean across block correlation = .342;  
166  $p < .001$ ). This process was repeated for cue location (Figure 5B; mean within block correlation = .360;  
167 mean across block correlation = .288;  $p < .001$ ) and cue outcome (Figure 5C; mean within block correlation  
168 = .345; mean across block correlation = .254;  $p < .001$ ). Additionally, given that the majority of our units  
169 showed an inhibitory response to the cue, we also plotted the firing rates according to the lowest time in fir-  
170 ing, and again found some maintenance of order, but largely different ordering across the two blocks (Figure  
171 5 supplement 1). Together, these results illustrate that NAc coding of task space was not confined to salient  
172 events such as cue-onset, but was approximately uniformly distributed throughout the task.



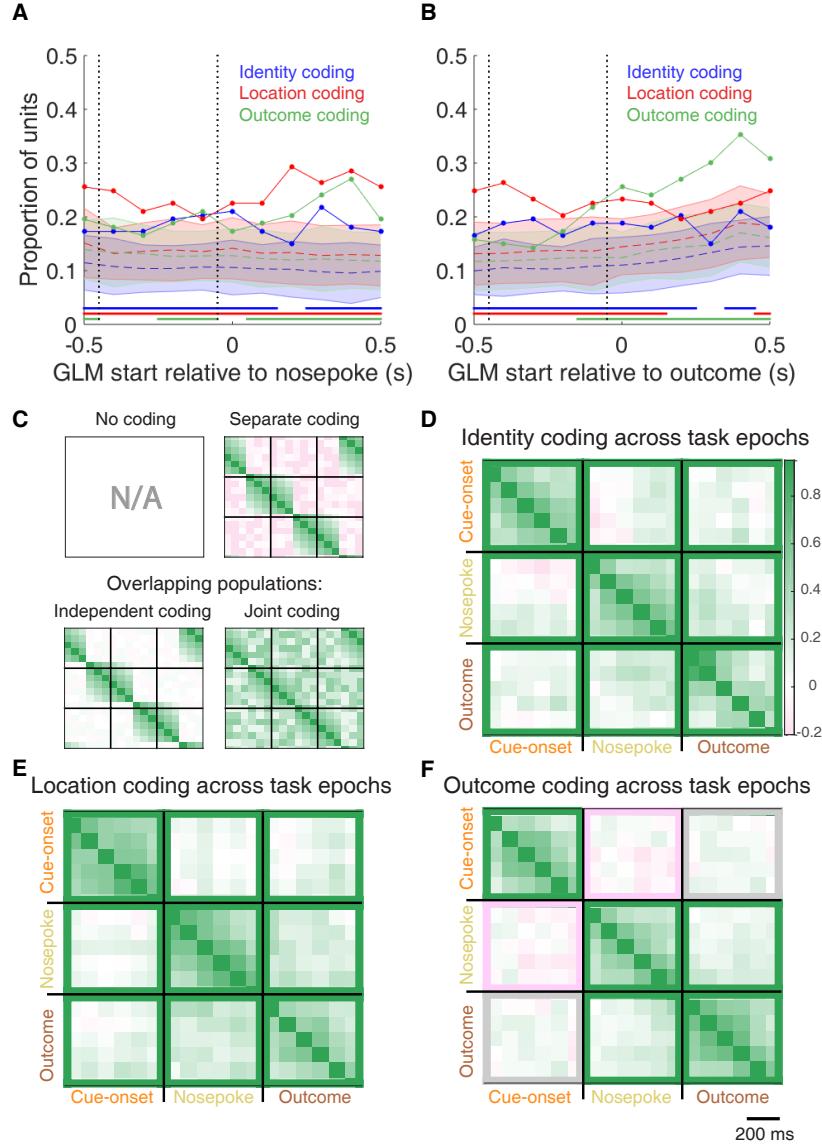
**Figure 5:** Distribution of NAc firing rates across time surrounding cue-onset. Each panel shows normalized (z-score) peak firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue-onset), averaged across all trials for a specific cue type, indicated by text labels. **A**, 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. Color bar displays relationship between z-score and color. **B**: Same layout as in **A**, except that the panels now compare two different locations on the track instead of two 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. Overall, NAc units coded 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.

<sup>173</sup> **NAc encoding of cue features persists until outcome:**

<sup>174</sup> In order to be useful for credit assignment in reinforcement learning, a trace of the cue must be maintained  
<sup>175</sup> until the outcome, so that information about the outcome can be associated with the outcome-predictive cue  
<sup>176</sup> (Figure 1B). Investigation into the post-approach period during nosepoke revealed units that discriminated  
<sup>177</sup> various cue features, with some units showing discriminative activity at both cue-onset and nosepoke (Figure  
<sup>178</sup> 6, supplement 1, supplement 2). To quantitatively test whether representations of cue features persisted  
<sup>179</sup> post-approach until the outcome was revealed, we fit sliding window GLMs to the post-approach firing  
<sup>180</sup> rates of cue-modulated units aligned to both the time of nosepoke into the reward receptacle, and after the  
<sup>181</sup> outcome was revealed (Figure 7A,B, supplement 1 A-D, Table 1). This analysis showed that a variety of  
<sup>182</sup> units discriminated firing according to cue identity (~20% of cue-modulated units), location (~25% of cue-  
<sup>183</sup> modulated units), and outcome (~25% of cue-modulated units), but not other task parameters, showing that  
<sup>184</sup> NAc activity discriminates various cue conditions well into a trial.



**Figure 6:** Examples of cue-modulated NAc units influenced by cue features at time of nosepoke. **A:** Example of a cue-modulated NAc unit that exhibited identity coding at both cue-onset and during subsequent 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 exhibited location coding, 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 exhibited outcome coding, at both cue-onset and nosepoke (E), and only nosepoke (F), with the PETHs comparing reward-available (red) and reward-unavailable (green) trials.



**Figure 7:** Summary of influence of cue features on cue-modulated NAc units at time points surrounding nosepoke and subsequent receipt of outcome. **A-B:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of nosepoke (A), and outcome (B). **A:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), location (red solid line), and outcome (green solid line) significantly contributed to the model at various time epochs relative to when the rat made a nosepoke. Dashed colored lines indicate the average of shuffling the firing rate order that went into the GLM 100 times. Error bars indicate 1.96 standard deviations from the shuffled mean. Solid lines at the bottom indicate when the proportion of units observed was greater than the shuffled distribution ( $z\text{-score} > 1.96$ ). Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the GLM. **B:** Same as A, but for time epochs relative to receipt of outcome after the rat got feedback about his approach. **C-F:** Correlation matrices testing the persistence of cue feature coding across points in time.

**C:** Schematic outlining the possible outcomes for coding of a cue feature across various points in a trial, generated by correlating the recoded beta coefficients from the GLMs and comparing to a shuffled distribution (see text for analysis details). Top left: coding is not present, therefore no comparison is possible. Top right: a cue feature is coded by separate populations of units across time. Displayed is a correlation matrix with each of the 9 blocks representing correlations for a cue feature across time bins for two task events from the sliding window GLM, with green representing positive correlations ( $r > 0$ ), pink negative correlations ( $r < 0$ ), and white representing significant correlation ( $r = 0$ ). X- and y-axis have the same axis labels, therefore the diagonal represents the correlation of cue feature against itself at that particular time point ( $r = 1$ ). Here the large amount of pink in the off-diagonal elements suggests that coding of a cue feature is accomplished by separate populations of units across time. Bottom left: Coding of a cue feature across time occurs in overlapping but independent populations of units, shown here by the abundance of white and relative lack of green and pink in the off-diagonal elements. Bottom right: Coding of a cue feature across time occurs in a joint overlapping population, shown here by the large amount of green in the off-diagonal elements. **D:** Correlation matrix showing the correlation of units that exhibited identity coding across time points after cue-onset, nosepoke, and outcome receipt. The window of GLMs used in each block is from the onset of the task phase to the 1900 ms window post-onset, in 100 ms steps. Each individual value is for a sliding window GLM within that range, with the scale bar contextualizing step size. Color bar displays relationship between correlation value and color. Colored square borders around each block indicate the result of a comparison of the mean correlation of a block to a shuffled distribution, with pink indicating separate populations ( $z\text{-score} < -1.96$ ), grey indicating overlapping but independent populations, and green indicating joint overlapping populations ( $z\text{-score} > 1.96$ ). **E-F:** Same as D, but for location and outcome coding, respectively.

185 To determine whether NAc representations of cue features at nosepoke and outcome were encoded by a  
186 similar pool of units as during cue-onset, we correlated recoded beta coefficients from the GLMs for a cue  
187 feature across time points in the task, and compared the obtained correlations to correlations generated by  
188 shuffling unit ordering within a recoded array (Figure 1B,C, 7C-F). This revealed that identity coding was  
189 accomplished by a joint population across all three task events (cue-onset and nosepoke: mean  $r = .048$ , z-  
190 score = 3.47; cue-onset and outcome: mean  $r = .081$ , z-score = 5.55; nosepoke and outcome: 10.91), while  
191 joint coding was observed between nosepoke and outcome (mean  $r = .147$ ; 3.94 standard deviations from the  
192 shuffled mean). Applying this same analysis for cue location revealed a similar pattern for location coding  
193 (cue-onset and nosepoke: mean  $r = .058$ , z-score = 4.15; cue-onset and outcome: mean  $r = .093$ , z-score  
194 = 6.40; nosepoke and outcome: mean  $r = .204$ , z-score = 14.29). However, outcome coding at cue-onset  
195 was separate from coding at nosepoke (mean  $r = -.040$ , z-score = -3.10), and independent from coding at  
196 outcome (mean  $r = .025$ , z-score = 1.65), while joint coding was observed between nosepoke and outcome  
197 (mean  $r = .148$ , z-score = 9.74). Together, these findings show that the NAc maintains representations of  
198 cue identity and location by a joint overlapping population throughout a trial, while separate populations of  
199 units encode cue outcome before and after a behavioral decision has been made.

200 To assess overlap among cue features at nosepoke and outcome receipt, we applied the same recoded coeffi-  
201 cient analysis (Figure 7 supplement 1 E,F). This revealed joint coding of cue features at the time of nosepoke  
202 (cue identity and location: mean  $r = .124$ , z-score = 8.26; cue identity and outcome: mean  $r = .052$ , z-score  
203 = 3.65; mean  $r = .097$ , z-score = 6.60); while at outcome, identity was coded by a joint population with  
204 both location (mean  $r = .085$ , z-score = 5.58), and outcome (mean  $r = .039$ , z-score = 2.93), and location  
205 and outcome were coded by a separate population of units (mean  $r = .004$ , z-score = 0.28).

206 To assess the distributed coding of units for task space around outcome receipt, we aligned normalized  
207 peak firing rates to nosepoke onset (Figure 7 supplement 2). This revealed a clustering of responses around  
208 outcome receipt for all cue conditions where the rat would have received reward, in addition to the same  
209 pattern of higher within- vs across-block correlations for cue identity (Figure 7 supplement 2 A,C; mean

210 within block correlation = .551; mean across block correlation = .484; p < .001), cue location (Figure 7  
211 supplement 2 B,E; mean within block correlation = .468; mean across block correlation = .412; p < .001),  
212 and cue outcome (Figure 7 supplement 2 C,F; mean within block correlation = .511; mean across block  
213 correlation = .408; p < .001), further reinforcing that the NAc segments the task and represents all aspects  
214 of task space.

## 215 Discussion

216 The main result of the present study is that NAc units encode not only the expected outcome of outcome-  
217 predictive cues, but also the identity of such cues (Figure 1A). The population of units that coded for cue  
218 identity was statistically independent from the population coding for expected outcome at cue-onset (i.e.  
219 overlap as expected from chance), while a joint overlapping population coded for identity and outcome  
220 at both nosepoke and outcome receipt (i.e. overlap greater than that expected from chance, Figure 1C).  
221 Importantly, this identity coding was maintained on approach trials by a similar population of units both  
222 during a delay period where the rat held a nosepoke until the outcome was received, and immediately after  
223 outcome receipt (Figure 1B,C). Cue identity information was also present at the population level, as indicated  
224 by high classification performance based on pseudoensembles. More generally, NAc unit activity profiles  
225 were not limited to salient task events such as the cue, nosepoke and outcome, but were distributed more  
226 uniformly throughout the task. This temporally distributed activity differed systematically between cue  
227 identities, expected outcomes and locations. We discuss these observations and their implications below.

### 228 Identity coding:

229 Our finding that NAc units can discriminate between different outcome-predictive stimuli with similar mo-  
230 tivational significance (i.e. encode cue identity) expands upon an extensive rodent literature examining NAc

correlates of conditioned stimuli (Ambroggi et al., 2008; Atallah et al., 2014; Bissonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Dejean et al., 2017; Goldstein et al., 2012; Ishikawa et al., 2008; Lansink et al., 2012; McGinty et al., 2013; Nicola, 2004; Roesch et al., 2009; Roitman et al., 2005; Saddoris et al., 2011; Setlow et al., 2003; Sugam et al., 2014; West & Carelli, 2016; Yun et al., 2004). Perhaps the most comparable work in rodents comes from a study that found a subset of NAc units that modulated their firing for an odor when it predicted distinct but equally valued rewards (Cooch et al., 2015). The present study is complementary to such *outcome identity* coding, in showing that NAc units encode *cue identity* in addition to the reward it predicts (Figure 1A). Setlow et al. paired two distinct odor cues with appetitive and aversive odor cues respectively in a Go/NoGo task, such that cue identity and cue outcome were linked. Although reversal sessions were run that uncoupled identity and outcome, the resulting changes in reinforcement history and behavioral performance precluded a clear test of cue identity coding. Thus, our study was different in asking how distinct cues encoding the same anticipated outcome are encoded. Our results suggest that the NAc dissociates cue identity representations at multiple levels of analysis (e.g. single-unit and population) even when the motivational significance of these stimuli is identical. Viewed within the neuroeconomic framework of decision making, functional magnetic resonance imaging (fMRI) studies have found support for NAc representations of *offer value*, a domain-general common currency signal that enables comparison of different attributes such as reward identity, effort, and temporal proximity (Bartra et al., 2013; Levy & Glimcher, 2012; Peters & Büchel, 2009; Sescousse et al., 2015). Our study adds to a growing body of electrophysiological research that suggests the view of the NAc as a value center, while informative and capturing major aspects of NAc processing, neglects additional contributions of NAc to learning and decision making such as the offer (cue) identity signal reported here.

Our analyses were designed to eliminate several potential alternative interpretations to cue identity coding. Because the different cues were separated into different blocks, units that discriminated between cue identities could instead be encoding time or other slowly-changing quantities. We excluded this possible confound by excluding units that showed a drift in firing between the first and second half within a block. Additionally, we included time as a nuisance variable in our GLMs, to exclude firing rate variance in the remaining units

257 that could be attributed to this confound. Furthermore, we found that the temporally evolving firing rate  
258 throughout a trial was more strongly correlated within a block than across blocks. However, the possibility  
259 remains that instead of, or in addition to, stimulus identity, these units encode a preferred context, or even  
260 a macroscale representation of progress through the session. Indeed, encoding of the current strategy could  
261 be an explanation for the presence of pre-cue identity coding (Figure 4A), as well as for the differential  
262 distributed coding of task structure across blocks observed in the current study (Figure 5).

263 An overall limitation of the current study is that rats were never presented with both sets of cues simul-  
264 taneously, and were not required to switch strategies between multiple sets of cues (this was attempted in  
265 behavioral pilots, however animals took several days of training to successfully switch strategies). Addition-  
266 ally, our recordings were done during performance on the well-learned behavior, and not during the initial  
267 acquisition of the cue-outcome relationships when an eligibility trace would be most useful. Thus, it is  
268 unknown to what extent the cue identity encoding we observed is behaviorally relevant, although extrapo-  
269 lating data from other work (Sleeker et al., 2016) suggests that cue identity coding would be modulated by  
270 relevance. Furthermore, NAc core lesions have been shown to impair shifting between different behavioral  
271 strategies (Floresco et al., 2006), and it is possible that selectively silencing the units that prefer responding  
272 for a given modality or rule would impair performance when the animal is required to use that information,  
273 or artificial enhancement of those units would cause them to use the rule when it is the inappropriate strategy.

274 **NAc activity provides a rich task representation beyond reward alone:**

275 Beyond coding of cue identity, we found several other notable features of NAc activity. First, a substantial  
276 number of cue-modulated units was differentially active depending on location, consistent with previous re-  
277 ports (Lavoie & Mizumori, 1994; Mulder et al., 2005; Strait et al., 2016; Wiener et al., 2003). However, it is  
278 notable that in our task, location is explicitly uninformative about reward, yet coding of this uninformative  
279 variable persists. This is unlike previous work of location coding in the dorsolateral striatum, which was  
280 present when location was predictive of reward, and absent when it was uninformative (Schmitzer-Torbert

281 and Redish, 2008). Persistent coding of location in the NAc is likely attributable to inputs from the hip-  
282 pocampus (Lansink et al., 2016; Sjulson et al., 2017; Tabuchi et al., 2000; van der Meer & Redish, 2011);  
283 speculatively, this coding may map onto a bias in credit assignment, such that motivationally relevant events  
284 are likely to be associated with the locations where they occur.

285 A second striking feature of NAc activity evident from this task is that NAc units were not only active at  
286 salient events such as cue presentation, nosepoking, and feedback about the outcome, but distributed their  
287 activity throughout a trial (Figure 5). This observation is consistent with previous work reporting that NAc  
288 units can signal progress through a sequence of cues and/or actions (Atallah et al., 2014; Berke et al., 2009;  
289 Khamassi et al., 2008; Lansink et al., 2012; Mulder et al., 2004; Shidara et al., 1998) and reminiscent of  
290 similar observations in the ventral pallidum (Tingley & Buzsáki, 2018) to which the NAc projects. Extending  
291 this previous work, we show that the specific pattern of NAc units throughout a trial can be modified by task  
292 variables such as cue identity. This richer view of NAc activity recalls a dynamical systems perspective, in  
293 which different task conditions correspond to different trajectories in a neural state space (e.g. Buonomano  
294 and Maass, 2009; Shenoy et al. 2013). In any case, this view of NAc activity provides a substantially richer  
295 picture than that expected from encoding of reward-related variables alone.

296 **Functional relevance of cue identity coding:**

297 One possible function of cue identity coding is to support contextual modulation of the motivational rel-  
298 evance of specific cues. A context can be understood as a particular mapping between specific cues and  
299 their outcomes: for instance, in context 1 cue A but not cue B is rewarded, whereas in context 2 cue B  
300 but not cue A is rewarded. Successfully implementing such contextual mappings requires representation of  
301 the cue identities. Indeed, Sleezer et al. (2016) recorded NAc responses during the Wisconsin Card Sort-  
302 ing Task, a common set-shifting task used in both the laboratory and clinic, and found units that preferred  
303 firing to stimuli when a certain rule, or rule category was currently active. Further support for a modula-  
304 tion of NAc responses by strategy comes from an fMRI study that examined blood-oxygen-level dependent

305 (BOLD) levels during a set-shifting task (FitzGerald et al., 2014). In this task, participants learned two sets  
306 of stimulus-outcome contingencies, a visual set and an auditory set. During testing they were presented with  
307 both simultaneously, and the stimulus dimension that was relevant was periodically shifted between the two.  
308 It was found that bilateral NAc activity reflected value representations for the currently relevant stimulus  
309 dimension, and not the irrelevant stimulus. Given that BOLD activity is thought to reflect the processing  
310 of incoming and local information, and not spiking output (Logothetis et al., 2001), it is possible that the  
311 relevance-gated value representations observed by FitzGerald et al. (2014) are integrated with the relevant  
312 identity coding in the output of the NAc, as observed in the current study.

313 A different potential role for cue identity coding is in learning to associate rewards with reward-predictive  
314 features of the environment, a process referred to as *credit assignment* in the reinforcement learning literature  
315 (Sutton & Barto, 1998). Maladaptive decision making, as occurs in schizophrenia, addiction, Parkinson's  
316 Disease and others can result from dysfunctional reward prediction errors (RPEs) and value signals (Frank et  
317 al., 2004; Gradin et al., 2011; Maia & Frank, 2011). This view has been successful in explaining both positive  
318 and negative symptoms in schizophrenia, and deficits in learning from feedback in Parkinson's (Frank et al.,  
319 2004; Gradin et al., 2011). However, the effects of RPE and value updating are contingent upon encoding  
320 of preceding action and cue features, the eligibility trace (Lee et al., 2012; Sutton & Barto, 1998). Value  
321 updates can only be performed on these aspects of preceding experience that are encoded when the update  
322 occurs. Therefore, maladaptive learning and decision making can result from not only aberrant RPEs but also  
323 from altered cue feature encoding. For instance, on this task the environmental stimulus that signaled the  
324 availability of reward was conveyed by two distinct cues that were presented in four locations. Although in  
325 our current study, the location and identity of the cue did not require any adjustments in the animals behavior,  
326 we found coding of these features alongside the expected outcome of the cue that could be the outcome of  
327 credit assignment computations computed upstream (Akaishi et al., 2016; Asaad et al., 2017; Chau et al.,  
328 2015; Noonan et al., 2017). Identifying neural coding related to an aspect of credit assignment is important  
329 as inappropriate credit assignment could be a contributor to conditioned fear overgeneralization seen in  
330 disorders with pathological anxiety such as generalized anxiety disorder, post-traumatic stress disorder, and

331 obsessive-compulsive disorder (Kaczkurkin et al., 2017; Kaczkurkin & Lissek, 2013; Lissek et al., 2014),  
332 and delusions observed in disorders such as schizophrenia, Alzheimer's and Parkinson's (Corlett et al., 2010;  
333 Kapur, 2003). Thus, our results provide a starting point for studies of the neural basis of credit assignment,  
334 and the extent and specific manner in which this process fails in syndromes such as schizophrenia, obsessive-  
335 compulsive disorder, and others.

## 336 Methods

### 337 Subjects:

338 A sample size of 4 adult male Long-Evans rats (Charles River, Saint Constant, QC) from an a priori de-  
339 termined sample of 5 were used as subjects (1 rat was excluded from the data set due to poor cell yield).  
340 Rats were individually housed with a 12/12-h light-dark cycle, and tested during the light cycle. Rats were  
341 food deprived to 85-90% of their free feeding weight (weight at time of implantation was 440 - 470 g), and  
342 water restricted 4-6 hours before testing. All experimental procedures were approved by the the University  
343 of Waterloo Animal Care Committee (protocol# 11-06) and carried out in accordance with Canadian Council  
344 for Animal Care (CCAC) guidelines.

### 345 Overall timeline:

346 Each rat was first handled for seven days during which they were exposed to the experiment room, the  
347 sucrose solution used as a reinforcer, and the click of the sucrose dispenser valves. Rats were then trained  
348 on the behavioral task (described in the next section) until they reached performance criterion. At this point  
349 they underwent hyperdrive implantation targeted at the NAc. Rats were allowed to recover for a minimum  
350 of five days before being retrained on the task, and recording began once performance returned to pre-

351 surgery levels. Upon completion of recording, animals were gloised, euthanized and recording sites were  
352 histologically confirmed.

353 **Behavioral task and training:**

354 The behavioral apparatus was an elevated, square-shaped track (100 x 100 cm, track width 10 cm) containing  
355 four possible reward locations at the end of track “arms” (Figure 2A). Rats initiated a *trial* by triggering a  
356 photobeam located 24 cm from the start of each arm. Upon trial initiation, one of two possible light cues (L1,  
357 L2), or one of two possible sound cues (S1, S2), was presented that signaled the presence (*reward-available*  
358 *trial*, L1+, S1+) or absence (*reward-unavailable trial*, L2-, S2-) of a 12% sucrose water reward (0.1 mL) at  
359 the upcoming reward site. A trial was classified as an *approach trial* if the rat turned left at the decision point  
360 and made a nosepoke at the reward receptacle (40 cm from the decision point), while trials were classified as  
361 a *skip trial* if the rat instead turned right at the decision point and triggered the photobeam to initiate the next  
362 trial. A trial was labeled *correct* if the rat approached (i.e. nosepoked) on reward-available trials, and skipped  
363 (i.e. did not nosepoke) on reward-unavailable trials. On reward-available trials there was a 1 second delay  
364 between a nosepoke and subsequent reward delivery. *Trial length* was determined by measuring the length  
365 of time from cue-onset until nosepoke (for approach trials), or from cue-onset until the start of the following  
366 trial (for skip trials). Trials could only be initiated through clockwise progression through the series of arms,  
367 and each entry into the subsequent arm on the track counted as a trial. Cues were present until 1 second after  
368 outcome receipt on approach trials, and until initiating the following trial on skip trials.

369 Each session consisted of both a *light block* and a *sound block* with 100 trials each. Within a block, one cue  
370 signaled reward was available on that trial (L1+ or S1+), while the other signaled reward was not available  
371 (L2- or S2-). Light block cues were a flashing white light, and a steady yellow light. Sound block cues  
372 were a 2 kHz sine wave (low) and a 8 kHz sine wave (high) whose amplitude was modulated from 0 to  
373 maximum by a 2 Hz sine wave. Outcome-cue associations were counterbalanced across rats, e.g. for some  
374 rats L1+ was the flashing white light, and for others L1+ was the steady yellow light. The order of cue

375 presentation was pseudorandomized so that the same cue could not be presented more than twice in a row.  
376 Block order within each day was also pseudorandomized, such that the rat could not begin a session with  
377 the same block for more than two days in a row. Each session consisted of a 5 minute pre-session period  
378 on a pedestal (a terracotta planter filled with towels), followed by the first block, then the second block,  
379 then a 5 minute post-session period on the pedestal. For approximately the first week of training, rats were  
380 restricted to running in the clockwise direction by presenting a physical barrier to running counterclockwise.  
381 Cues signaling the availability and unavailability of reward, as described above, were present from the start  
382 of training. Rats were trained for 200 trials per day (100 trials per block) until they discriminated between  
383 the reward-available and reward-unavailable cues for both light and sound blocks for three consecutive days,  
384 according to a chi-square test rejecting the null hypothesis of equal approaches for reward-available and  
385 reward-unavailable trials, at which point they underwent electrode implant surgery.

386 **Surgery:**

387 Surgical procedures were as described previously (Malhotra et al., 2015). Briefly, animals were administered  
388 analgesics and antibiotics, anesthetized with isoflurane, induced with 5% in medical grade oxygen and main-  
389 tained at 2% throughout the surgery (~0.8 L/min). Rats were then chronically implanted with a “hyperdrive”  
390 consisting of 20 independently drivable tetrodes, with 4 designated as references tetrodes, and the remaining  
391 16 either all targeted for the right NAc (AP +1.4 mm and ML +1.6 mm relative to bregma; Paxinos & Watson  
392 1998), or 12 in the right NAc and 4 targeted at the mPFC (AP +3.0 mm and ML +0.6 mm, relative to bregma;  
393 only data from NAc tetrodes was analyzed). Following surgery, all animals were given at least five days to  
394 recover while receiving post-operative care, and tetrodes were lowered to the target (DV -6.0 mm) before  
395 being reintroduced to the behavioral task.

396 **Data acquisition and preprocessing:**

397 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with a

398 RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex). Sig-  
399 nals were referenced against a tetrode placed in the corpus callosum above the NAc. Candidate spikes for  
400 sorting into putative single units were obtained by band-pass filtering the data between 600-9000 Hz, thresh-  
401 olding and aligning the peaks (UltraMegaSort2k, Hill et al., 2011). Spike waveforms were then clustered  
402 with KlustaKwik using energy and the first derivative of energy as features, and manually sorted into units  
403 (MClust 3.5, A.D. Redish et al., <http://redishlab.neuroscience.umn.edu/MClust/MClust.html>). Isolated units  
404 containing a minimum of 200 spikes within a session were included for subsequent analysis. Units were  
405 classified as FSIs by an absence of interspike intervals (ISIs) > 2 s, while MSNs had a combination of ISIs  
406 > 2 s and phasic activity with shorter ISIs (Atallah et al., 2014; Barnes et al., 2005).

407 **Data analysis:**

408 *Behavior.* To determine if rats distinguished behaviorally between the reward-available and reward-unavailable  
409 cues (*cue outcome*), we generated linear mixed effects models to investigate the relationships between cue  
410 type and the proportion of trials approached, with *cue outcome* (reward available or not) and *cue identity*  
411 (light or sound) as fixed effects, and the addition of an intercept for rat identity as a random effect. For each  
412 cue, the average proportion of trials approached for a session was used as the response variable. Contribu-  
413 tion of cue outcome to behavior was determined by comparing the full model to a model with cue outcome  
414 removed. To assess within session learning we divided each block into two halves, and compared a model  
415 including a block half variable to a null model excluding this variable, to see if adding block half improved  
416 prediction of overall behavioral performance.

417 *Neural data.* Given that some of our analyses compare firing rates across time, particularly comparisons  
418 across blocks, we sought to exclude units with unstable firing rates that would generate spurious results  
419 reflecting a drift in firing rate over time unrelated to our task. We used a multipronged strategy to address this  
420 potential confound. As a first step, we ran a Mann-Whitney U test comparing the cue-modulated firing rates  
421 for the first and second half of trials within a block, and excluded 99 of 443 units from analysis that showed

422 a significant change for either block, leaving 344 units for further analyses by our GLM. Furthermore, we  
423 included time (trial number) as a nuisance variable in our GLMs to control for firing rate variance account  
424 for by this confound (see below). To investigate the contribution of different cue features (*cue identity*, *cue*  
425 *location* and *cue outcome*) on the firing rates of NAc single units, we first determined whether firing rates for  
426 a unit were modulated by the onset of a cue by collapsing across all cues and comparing the firing rates for the  
427 1 s preceding cue-onset with the 1 s following cue-onset. Single units were considered to be *cue-modulated*  
428 if a Wilcoxon signed-rank test comparing pre- and post-cue firing was significant at  $p < .01$ . Cue-modulated  
429 units were then classified as either increasing or decreasing if the post-cue activity was higher or lower than  
430 the pre-cue activity, respectively.

431 To determine the relative contribution of different task parameters to firing rate variance (as in Figure 4A,  
432 supplement 1), a forward selection stepwise GLM using a Poisson distribution for the response variable was  
433 fit to each cue-modulated unit, using data from every trial in a session. Cue identity (light block, sound  
434 block), cue location (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available, reward-unavailable),  
435 behavior (approach, skip), trial length, trial number, and trial history (reward availability on the previous 2  
436 trials) were used as predictors, with firing rate as the response variable. The GLMs were fit using a 500 ms  
437 sliding window moving in 100 ms steps centered at 250 ms pre-cue (so no post-cue activity was included)  
438 to centered at 750 ms post-cue, such that 11 different GLMs were fit for each unit, tracking the temporal  
439 dynamics of the influence of task parameters on firing rate around the onset of the cue. Units were classified  
440 as being modulated by a given task parameter if addition of the parameter significantly improved model fit  
441 using deviance as the criterion ( $p < .01$ ), and the total proportion of cue-modulated units influenced by a  
442 task parameter was counted for each time bin. A comparison of the R-squared value between the final model  
443 and the final model minus the predictor of interest was used to determine the amount of firing rate variance  
444 explained by the addition of that predictor for a given unit. To control for the amount of units that would be  
445 affected by a predictor by chance, we shuffled the trial order of firing rates for a particular unit within a time  
446 bin, ran the GLM with the shuffled firing rates, counted the proportion of units encoding a predictor, and took  
447 the average of this value over 100 shuffles. We then calculated how many z-scores the observed proportion

448 was from the mean of the shuffled distribution. For this and all subsequent shuffle analyses, we used a z-score  
449 of greater than 1.96 or less than -1.96 as a marker of significance. To further control for whether outcome  
450 coding could be attributed to subsequent behavioral variability at the choice point, we reran our cue-onset  
451 GLM for approach trials only.

452 To get a sense of the predictive power of these cue feature representations we trained a classifier using firing  
453 rates from a pseudoensemble comprised of our 133 cue-modulated units (Figure 4B). We created a matrix  
454 of firing rates for each time epoch surrounding cue-onset where each row was an observation representing  
455 the firing rate for a trial, and each column was a variable representing the firing rate for a given unit. Trial  
456 labels, or classes, were each condition for a cue feature (e.g. light and sound for cue identity), making sure  
457 to align trial labels across units. We then ran LDA on these matrices, using 10-fold cross validation to train  
458 the classifier on 90% of the trials and testing its predictions on the held out 10% of trials, and repeated this  
459 approach to get the classification accuracy for 100 iterations. To test if the classification accuracy was greater  
460 by chance, we shuffled the order of firing rates for each unit before we trained the classifier. We repeated  
461 this for 100 shuffled matrices for each time point, and calculated how many z-scores the mean classification  
462 rate of the observed data was from the mean of the shuffled distribution.

463 To determine the degree to which coding of cue identity, cue location, and cue outcome overlapped within  
464 units we correlated the recoded beta coefficients from the GLMs for the cue features (Figure 4C,D). Specifi-  
465 cally, we generated an array for each cue feature at each point in time where for all cue-modulated units we  
466 coded a ‘1’ if the cue feature was a significant predictor in the final model, and ‘0’ if it was not. We then  
467 correlated an array of the coded 0s and 1s for one cue feature with a similar array for another cue feature, re-  
468 peating this process for all post cue-onset sliding window combinations. The NAc was determined as coding  
469 a pair of cue features in a) separate populations of units if there was a significant negative correlation ( $r <$   
470 0), b) an independently coded overlapping population of units if there was no significant correlation ( $r = 0$ ),  
471 or c) a jointly coded overlapping population of units if there was a significant positive correlation ( $r > 0$ ).  
472 To summarize the correlation matrices generated from this analysis, we shuffled the unit ordering for each

473 array 100 times, took the mean of the 36 correlations for a block comparison for each of the 100 shuffles  
474 for an analysis window, and used the mean and standard deviation of these shuffled correlation averages to  
475 compare to the mean of the comparison block for the actual data.

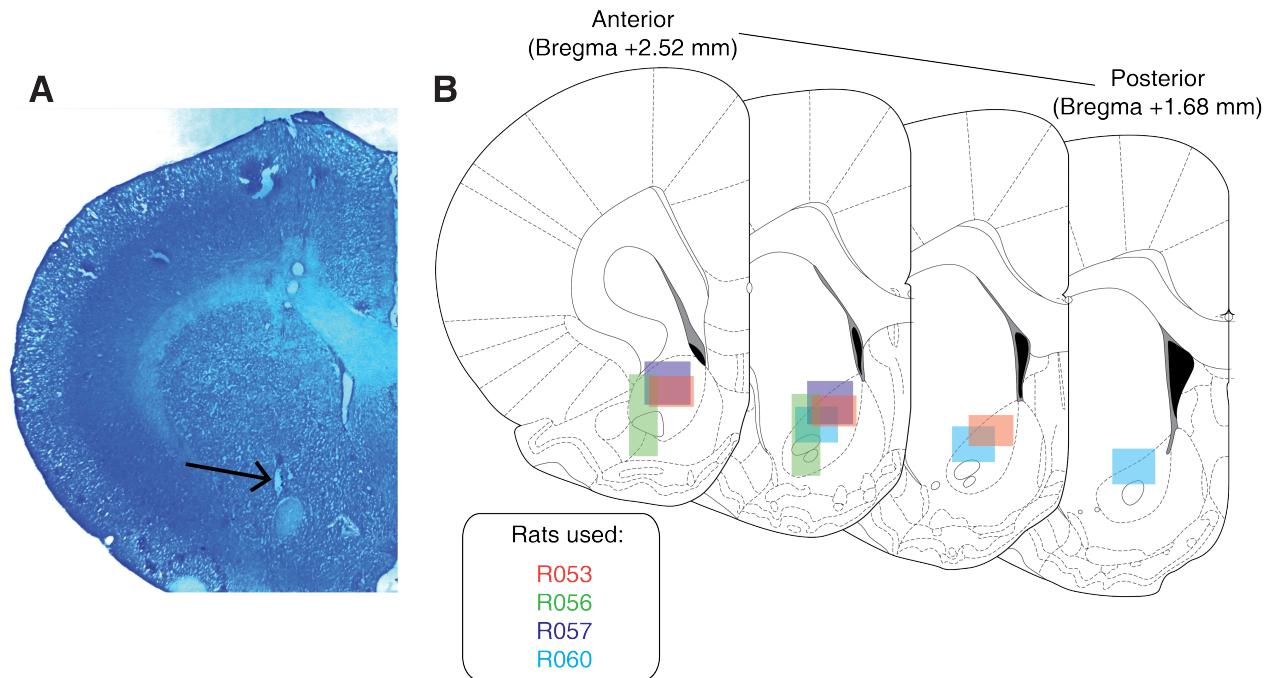
476 To better visualize responses to cues and enable subsequent population level analyses (as in Figures 3, 5),  
477 spike trains were convolved with a Gaussian kernel ( $\sigma = 100$  ms), and peri-event time histograms (PETHs)  
478 were generated by taking the average of the convolved spike trains across all trials for a given task condition.  
479 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units  
480 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure  
481 5). To compare representations of task space across cue conditions for a cue feature, the ordering of units  
482 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other  
483 condition (e.g. sound block). To assess whether the task distributions were different across cue conditions,  
484 we split each cue condition into two halves, controlling for the effects of time by shuffling trial ordering  
485 before the split, and calculated the correlation of the temporally evolving smoothed firing rate across each of  
486 these halves, giving us 6 correlation values for each unit. We then concatenated these 6 values across all 443  
487 units to give us an array of 2658 correlation coefficients. We then fit a linear mixed effects model, trying to  
488 predict these block comparison correlations with comparison type (e.g. 1st half of light block vs. 1st half of  
489 sound sound) as a fixed-effect term, and unit number as a random-effect term. Comparison type is nominal,  
490 so dummy variables were created for the various levels of comparison type, and coefficients were generated  
491 for each condition, referenced against one of the within-within comparison types (e.g. 1st half of light block  
492 vs. 2nd half of light block). The NAc was considered to discriminate across cue conditions if across-block  
493 correlations were lower than within-block correlations. Additionally, we ran a model comparison between  
494 the above model and a null model with just unit number, to see if adding comparison type improved model  
495 fit.

496 To identify the responsivity of units to different cue features at the time of nosepoke into a reward receptacle,  
497 and subsequent reward delivery, the same cue-modulated units from the cue-onset analyses were analyzed at

498 the time of nosepoke and outcome receipt using identical analysis techniques for all approach trials (Figures  
499 6, 7). To compare whether coding of a given cue feature was accomplished by the same or distinct population  
500 of units across time epochs, we ran the recoded coefficient correlation that was used to assess the degree of  
501 overlap among cue features within a time epoch. All analyses were completed in MATLAB R2015a, the  
502 code is available on our public GitHub repository (<http://github.com/vandermeerlab/papers>), and the data  
503 can be accessed through DataLad.

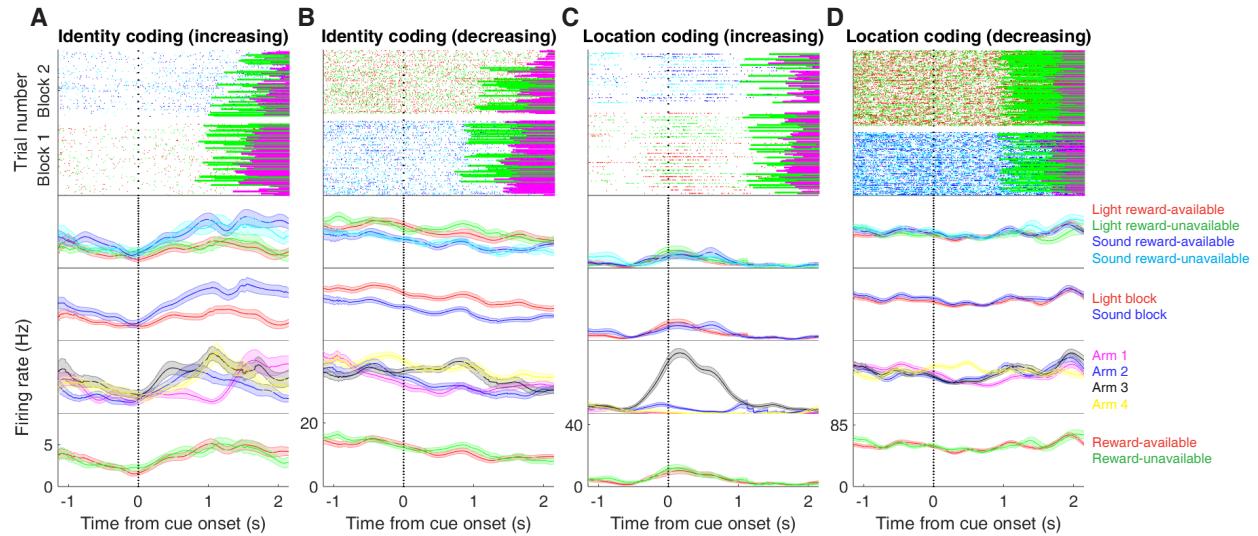
504 **Histology:**

505 Upon completion of the experiment, recording channels were glosed by passing  $10 \mu A$  current for 10 sec-  
506 onds and waiting 5 days before euthanasia, except for rat R057 whose implant detached prematurely. Rats  
507 were anesthetized with 5% isoflurane, then asphyxiated with carbon dioxide. Transcardial perfusions were  
508 performed, and brains were fixed and removed. Brains were sliced in  $50 \mu m$  coronal sections and stained  
509 with thionin. Slices were visualized under light microscopy, tetrode placement was determined, and elec-  
510 trodes with recording locations in the NAc were analyzed (Figure 8).

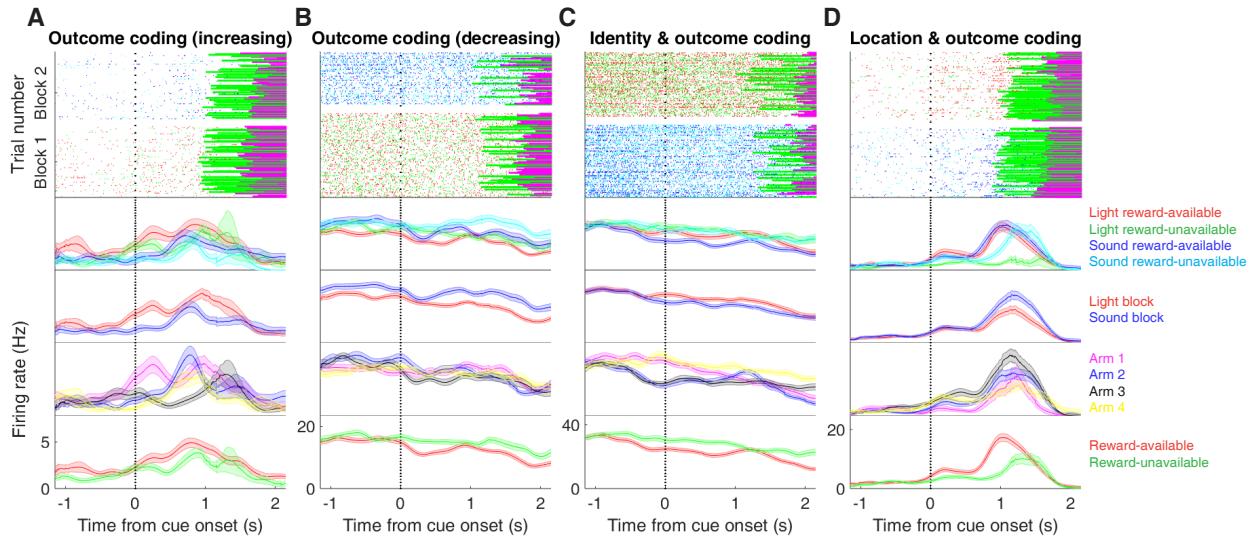


**Figure 8:** Histological verification of recording sites. Upon completion of experiments, brains were sectioned and tetrode placement was confirmed. **A:** Example section from R060 showing a recording site in the NAc core just dorsal to the anterior commissure (arrow). **B:** Schematic showing recording areas for all subjects.

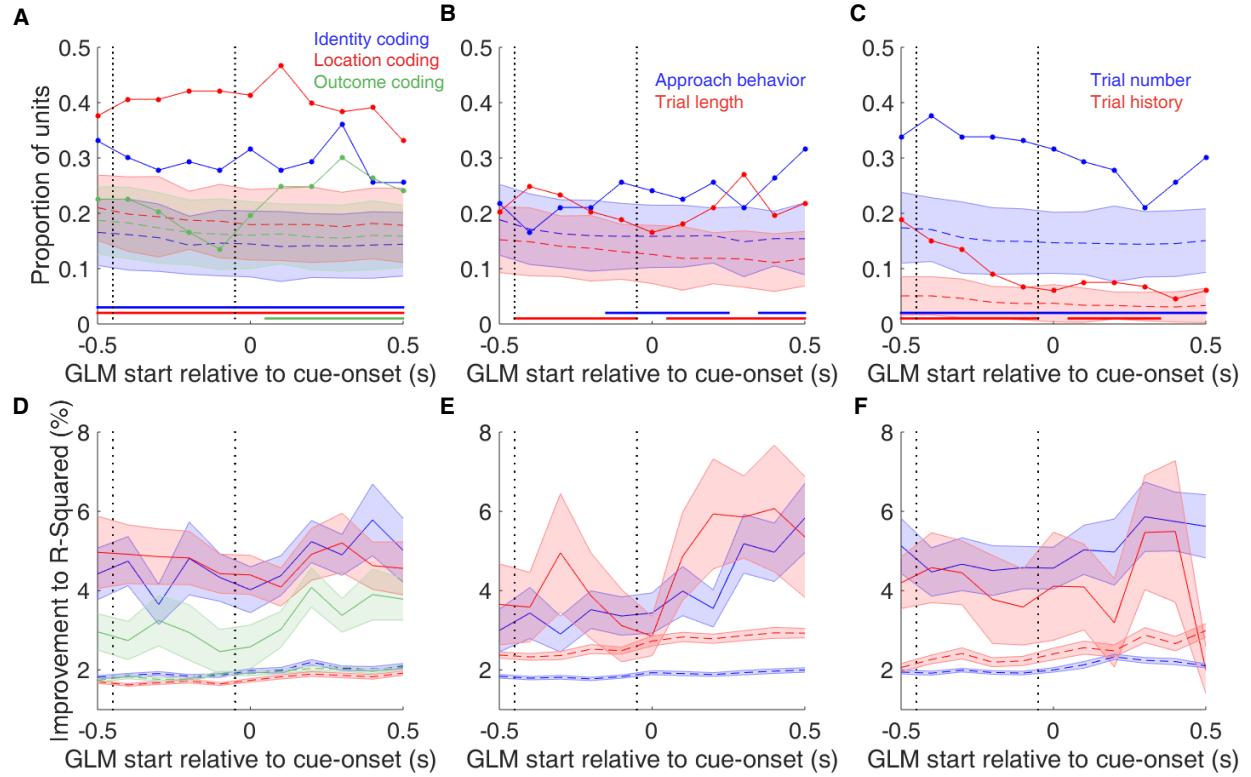
511 **Figure supplements**



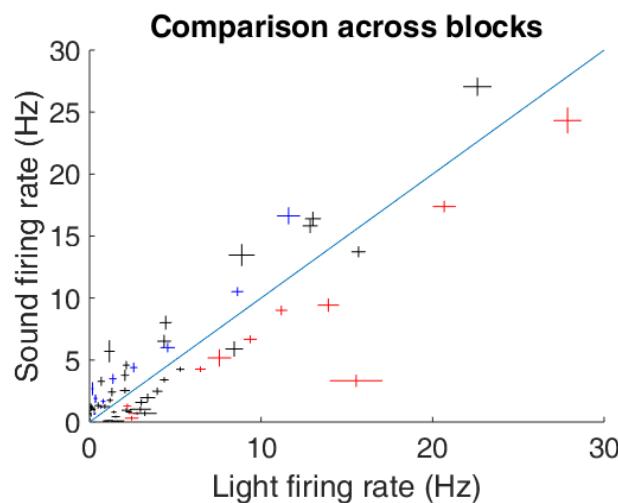
**Figure 3 supplement 1:** Expanded examples of cue-modulated NAc units influenced by different task parameters for Figure 3A-D, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).



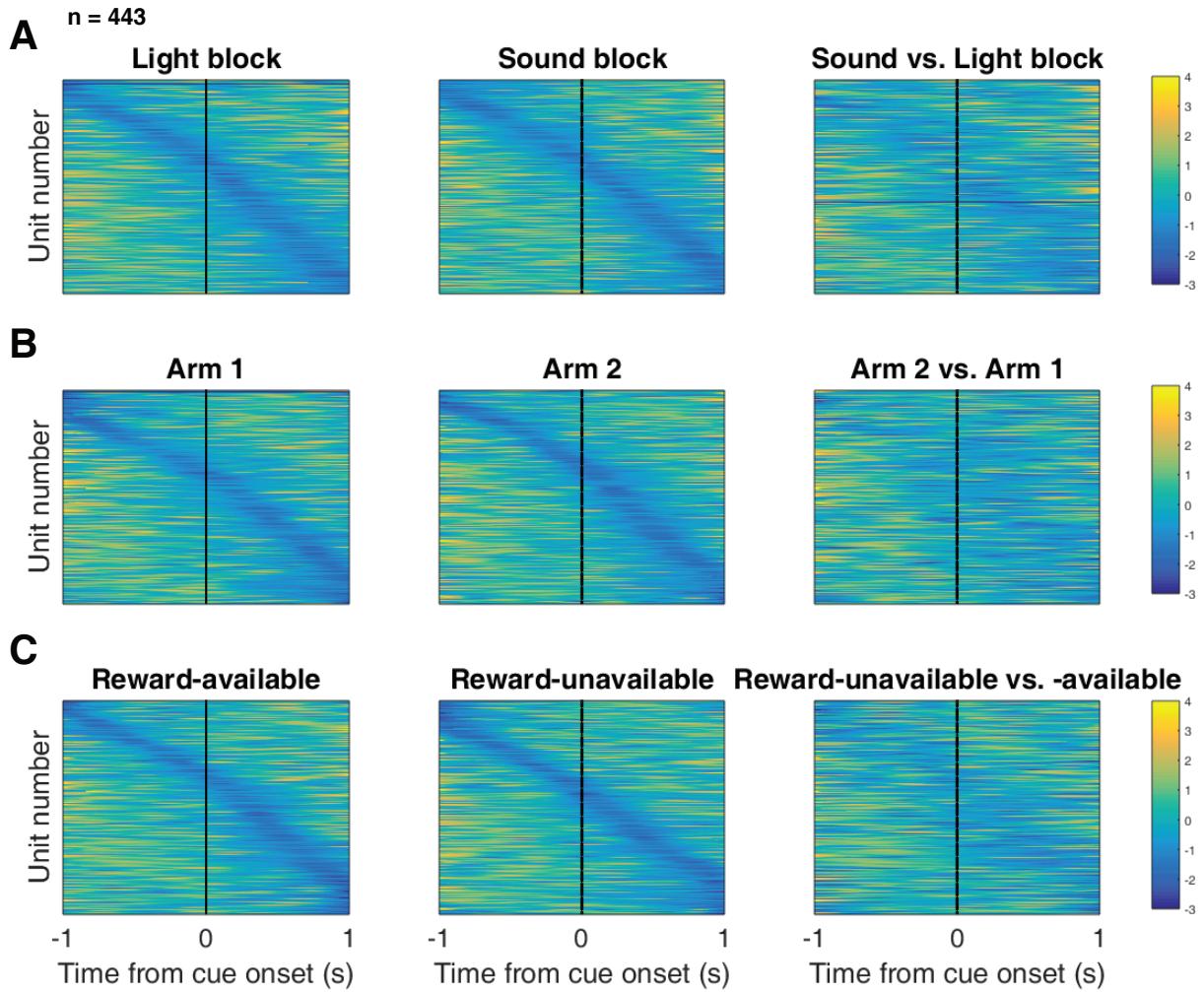
**Figure 3 supplement 2:** Expanded examples of cue-modulated NAc units influenced by different task parameters for Figure 3E-H, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).



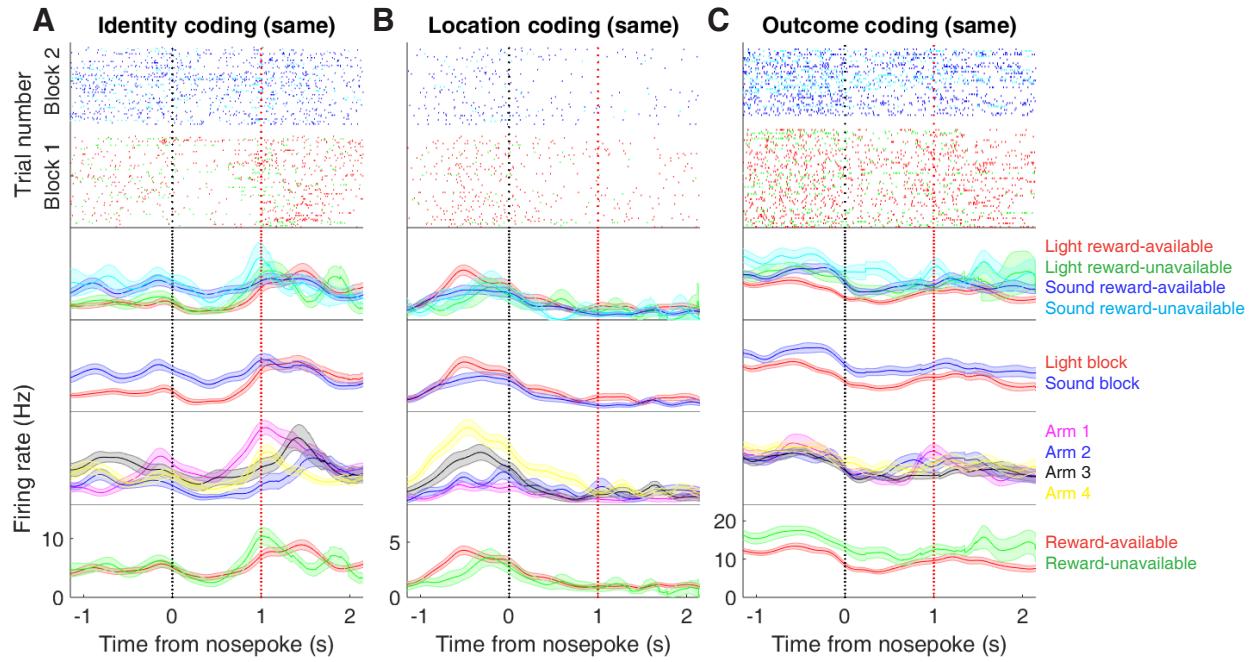
**Figure 4 supplement 1:** Summary of influence of various task parameters on cue-modulated NAc units at time points surrounding cue-onset. **A-C:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of cue-onset. **A:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), location (red solid line), and outcome (green 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. Error bars indicate 1.96 standard deviations from the shuffled mean. Solid lines at the bottom indicate when the proportion of units observed was greater than the shuffled distribution ( $z$ -score  $> 1.96$ ). Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the GLM. **B:** Same as A, but for approach behavior and trial length. **C:** Same as A, but for trial number and trial history. **D-F:** Average improvement to model fit. **D:** Average percent improvement to  $R^2$  for units where cue identity, location, or 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. **E:** Same as D, but for approach behavior and trial length. **F:** Same D, but for trial number and trial history.



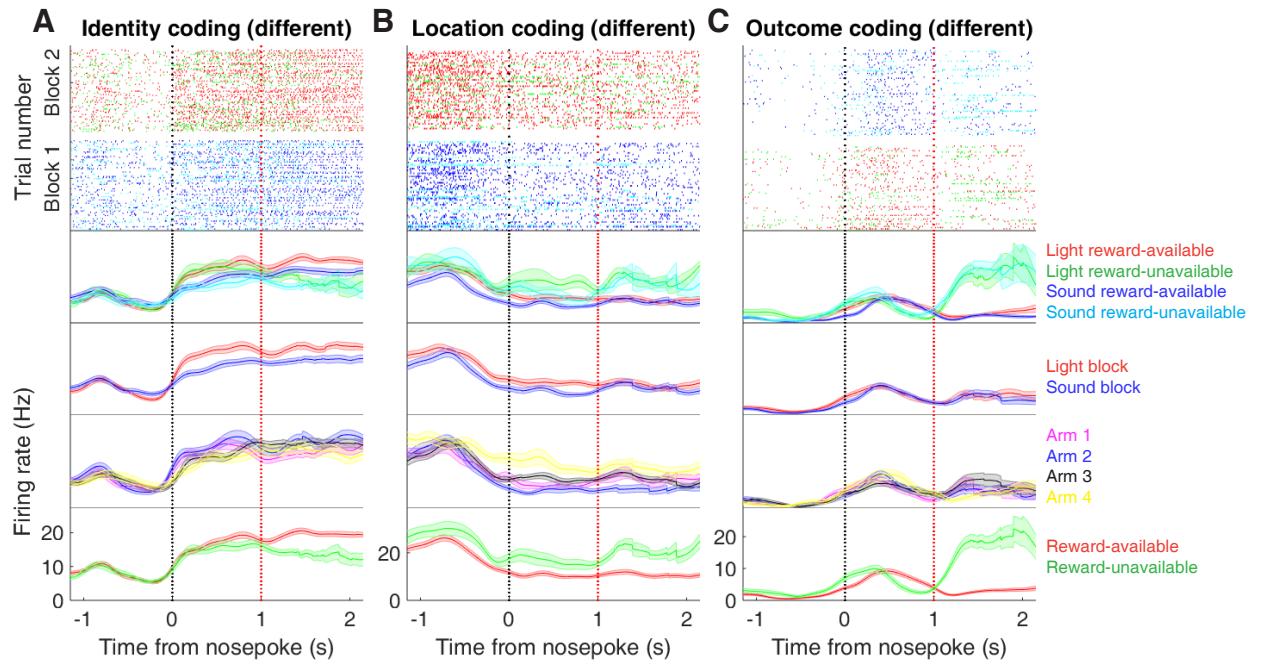
**Figure 4 supplement 2:** Scatter plot depicting comparison of firing rates for cue-modulated units across light and sound blocks. Crosses are centered on the mean firing rate, range represents the standard error of the mean. Colored crosses represents units that had cue identity as a significant predictor of firing rate variance in the GLM centered at cue-onset (blue are sound block preferring, red are light block preferring), whereas black crosses represent units where cue identity was not a significant predictor of firing rate variance. Diagonal dashed line indicates point of equal firing across blocks.



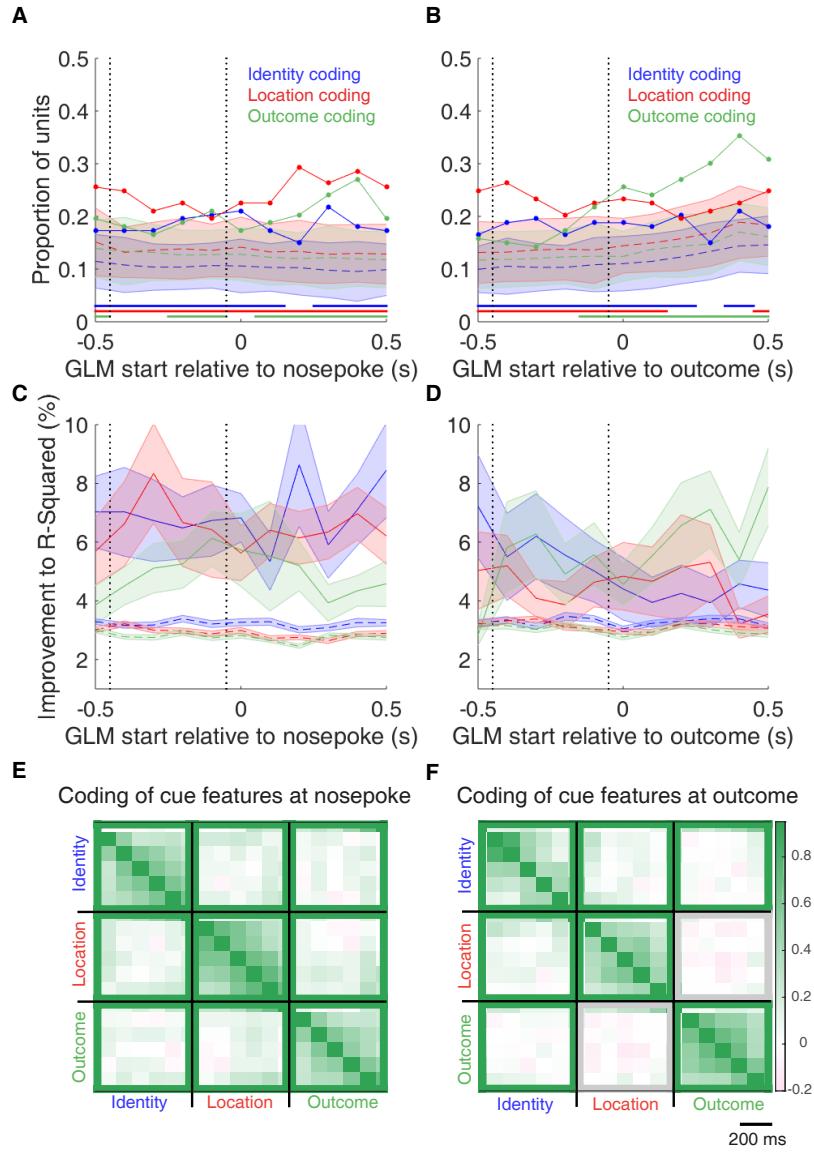
**Figure 5 supplement 1:** Distribution of NAc firing rates across time surrounding cue-onset. Each panel shows normalized (z-score) minimum firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue-onset), averaged across all trials for a specific cue type, indicated by text labels. **A:** Responses during different stimulus blocks as in Figure 5A, but with units ordered according to the time of their minimum firing rate. **B:** Responses during trials on different arms as in Figure 5B, but with units ordered by their minimum firing rate. **C:** Responses during cues signaling different outcomes as in Figure 5C, but with units ordered by their minimum firing rate. Overall, NAc units coded 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.



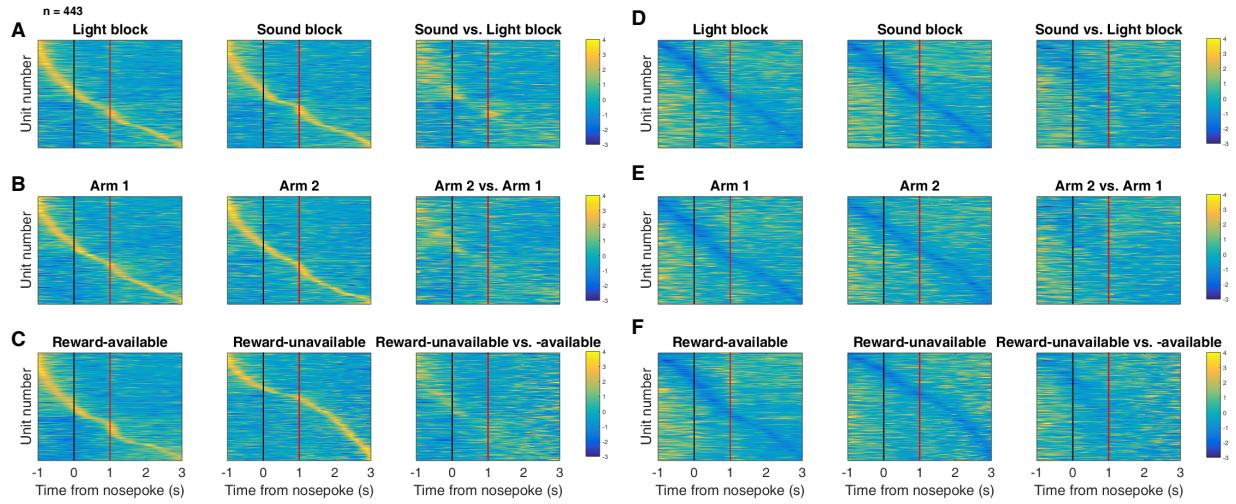
**Figure 6 supplement 1:** Expanded examples of cue-modulated NAc units influenced by different cue features at both cue-onset and during subsequent nosepoke hold for Figure 6A,C,E, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).



**Figure 6 supplement 2:** Expanded examples of cue-modulated NAc units influenced by different cue features at time of nosepoke for Figure 6B,D,F, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).



**Figure 7 supplement 1:** Summary of influence of cue features on cue-modulated NAc units at time points surrounding nosepoke and subsequent receipt of outcome. **A-B:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of nosepoke (A), and outcome (B). **A:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), location (red solid line), and outcome (green solid line) significantly contributed to the model at various time epochs relative to when the rat made a nosepoke. Dashed colored lines indicate the average of shuffling the firing rate order that went into the GLM 100 times. Error bars indicate 1.96 standard deviations from the shuffled mean. Solid lines at the bottom indicate when the proportion of units observed was greater than the shuffled distribution ( $z\text{-score} > 1.96$ ). Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the GLM. **B:** Same as A, but for time epochs relative to receipt of outcome after the rat got feedback about his approach. **C-D:** Average improvement to model fit. **C:** Average percent improvement to  $R^2$  for units where cue identity (blue solid line), location (red solid line), or outcome (green solid line) were significant contributors to the final model for time epochs relative to nosepoke. Dashed colored lines indicate the average of shuffling the firing rate order that went into the GLM 100 times. Shaded area around mean represents the standard error of the mean. **D:** Same C, but for time epochs relative to receipt of outcome. **E-F:** Correlation matrices testing the presence and overlap of cue feature coding at nosepoke (E) and outcome (F). **E:** Correlation matrix showing the correlation among identity, location, and outcome coding at nosepoke. Each of the 9 blocks represents correlations for two cue features across various nosepoke-centered time bins from the sliding window GLM, with green representing positive correlations ( $r > 0$ ), pink negative correlations ( $r < 0$ ), and grey representing no significant correlation ( $r = 0$ ). X- and y-axis have the same axis labels, therefore the diagonal represents the correlation of a cue feature against itself at that particular time point ( $r = 1$ ). The window of GLMs used in each block is from the onset of the task phase to the 500 ms window post-onset, in 100 ms steps. Each individual value is for a sliding window GLM within that range, with the scale bar contextualizing step size. Colored square borders around each block indicate the result of a comparison of the mean correlation to a shuffled distribution, with pink indicating separate populations ( $z\text{-score} < -1.96$ ), grey indicating overlapping but independent populations, and green indicating joint overlapping populations ( $z\text{-score} > 1.96$ ). **F:** Same as E, but for time bins following outcome receipt. Color bar displays relationship between correlation value and color.



**Figure 7 supplement 2:** 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. Color bar displays relationship between z-score and color. **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 coding 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 signaling different outcomes as in C, but with units ordered by their minimum firing rate. Overall, NAc units coded 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.

512 **References**

- 513 Akaishi, R., Kolling, N., Brown, J. W., & Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue  
514 Environment. *Journal of Neuroscience*, 36(4), 1096–1112. doi: 10.1523/JNEUROSCI.3159-15.2016
- 515 Ambroggi, F., Ishikawa, A., Fields, H. L., & Nicola, S. M. (2008). Basolateral Amygdala Neurons Facilitate Reward-Seeking  
516 Behavior by Exciting Nucleus Accumbens Neurons. *Neuron*, 59(4), 648–661. doi: 10.1016/j.neuron.2008.07.004
- 517 Asaad, W. F., Lauro, P. M., Perge, J. A., & Eskandar, E. N. (2017). Prefrontal Neurons Encode a Solution to the Credit-Assignment  
518 Problem. *The Journal of Neuroscience*, 37(29), 6995–7007. doi: 10.1523/JNEUROSCI.3311-16.2017
- 519 Atallah, H. E., McCool, A. D., Howe, M. W., & Graybiel, A. M. (2014). Neurons in the ventral striatum exhibit cell-type-specific  
520 representations of outcome during learning. *Neuron*, 82(5), 1145–1156. doi: 10.1016/j.neuron.2014.04.021
- 521 Averbeck, B. B., & Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*,  
522 20(4), 505–512. doi: 10.1038/nn.4506
- 523 Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and  
524 recoding of procedural memories. *Nature*, 437(7062), 1158–1161. doi: 10.1038/nature04053
- 525 Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis  
526 of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, 76, 412–427. doi:  
527 10.1016/J.NEUROIMAGE.2013.02.063
- 528 Berke, J. D., Breck, J. T., & Eichenbaum, H. (2009). Striatal Versus Hippocampal Representations During Win-Stay Maze Perform-  
529 ance. *Journal of Neurophysiology*, 101(3), 1575–1587. doi: 10.1152/jn.91106.2008
- 530 Berridge, K. C. (2012). From prediction error to incentive salience: Mesolimbic computation of reward motivation. *European  
531 Journal of Neuroscience*, 35(7), 1124–1143. doi: 10.1111/j.1460-9568.2012.07990.x
- 532 Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., ... Roesch, M. R. (2013). Separate  
533 Populations of Neurons in Ventral Striatum Encode Value and Motivation. *PLoS ONE*, 8(5), e64673. doi: 10.1371/jour-  
534 nal.pone.0064673
- 535 Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological  
536 Bulletin*, 114(1), 80–99. doi: 10.1371/journal.pone.0064673
- 537 Carelli, R. M. (2010). Drug Addiction: Behavioral Neurophysiology. In *Encyclopedia of neuroscience* (pp. 677–682). Elsevier.  
538 doi: 10.1016/B978-008045046-9.01546-1
- 539 Chang, S. E., & Holland, P. C. (2013). Effects of nucleus accumbens core and shell lesions on autosshaped lever-pressing. *Be-  
540 havioural Brain Research*, 256, 36–42. doi: 10.1016/j.bbr.2013.07.046
- 541 Chang, S. E., Wheeler, D. S., & Holland, P. C. (2012). Roles of nucleus accumbens and basolateral amygdala in autosshaped lever  
542 pressing. *Neurobiology of Learning and Memory*, 97(4), 441–451. doi: 10.1016/j.nlm.2012.03.008

- 543 Chau, B. K. H., Sallet, J., Papageorgiou, G. K., Noonan, M. A. P., Bell, A. H., Walton, M. E., & Rushworth, M. F. S. (2015).  
544 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5),  
545 1106–1118. doi: 10.1016/j.neuron.2015.08.018
- 546 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., & Wightman, R. M. (2007). Coordinated  
547 Accumbal Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2), 237–244. doi:  
548 10.1016/j.neuron.2007.03.021
- 549 Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in  
550 schizophrenia. *Psychological Review*, 99(1), 45.doi: 10.1016/j.neuron.2007.03.021
- 551 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., & Schoenbaum, G. (2015).  
552 Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature Com-  
553 munications*, 6, 7195. doi: 10.1038/ncomms8195
- 554 Corbit, L. H., & Balleine, B. W. (2011). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are  
555 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33), 11786–11794. doi:  
556 10.1523/JNEUROSCI.2711-11.2011
- 557 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in  
558 Neurobiology*, 92, 345–369. doi: 10.1016/j.pneurobio.2010.06.007
- 559 Cromwell, H. C., & Schultz, W. (2003). Effects of Expectations for Different Reward Magnitudes on Neuronal Activity in Primate  
560 Striatum. *Journal of Neurophysiology*, 89(5), 2823–2838. doi: 10.1152/jn.01014.2002
- 561 Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine  
562 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8), 1020–1028. doi: 10.1038/nn1923
- 563 Day, J. J., Wheeler, R. A., Roitman, M. F., & Carelli, R. M. (2006). Nucleus accumbens neurons encode Pavlovian approach behav-  
564 iors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5), 1341–1351. doi: 10.1111/j.1460-  
565 9568.2006.04654.x
- 566 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., ... Le Moine, C. (2017). Memories of Opiate Withdrawal  
567 Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychopharmacology*,  
568 42(5), 1157–1168. doi: 10.1038/npp.2016.272
- 569 du Hoffmann, J., & Nicola, S. M. (2014). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the  
570 Nucleus Accumbens. *Journal of Neuroscience*, 34(43), 14349–14364. doi: 10.1523/JNEUROSCI.3492-14.2014
- 571 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental  
572 Psychology*, 32(2), 150–155. doi: 10.1037/h0058316
- 573 Fitzgerald, T. H. B., Schwartenbeck, P., & Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent  
574 and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4), 1271–1279. doi: 10.1523/JNEUROSCI.4389-  
575 13.2014

- 576 Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., ... Akil, H. (2011). A selective role for dopamine in  
577 stimulus-reward learning. *Nature*, 469(7328), 53–59. doi: 10.1038/nature09588
- 578 Floresco, S. B. (2015). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of  
579 Psychology*, 66(1), 25–52. doi: 10.1146/annurev-psych-010213-115159
- 580 Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., & Magyar, O. (2006). Dissociable roles for the nucleus accumbens core and shell  
581 in regulating set shifting. *Journal of Neuroscience*, 26(9), 2449–2457. doi: 10.1523/JNEUROSCI.4431-05.2006
- 582 Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism.  
583 *Science*, 306(5703), 1940–1943. doi: 10.1126/science.1102941
- 584 Giertler, C., Bohn, I., & Hauber, W. (2004). Transient inactivation of the rat nucleus accumbens does not impair guidance of  
585 instrumental behaviour by stimuli predicting reward magnitude. *Behavioural Pharmacology*, 15(1), 55–63. doi: 10.1126/sci-  
586 ence.1102941
- 587 Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., ... Roesch, M. R. (2012). Ventral Striatum  
588 Encodes Past and Predicted Value Independent of Motor Contingencies. *Journal of Neuroscience*, 32(6), 2027–2036. doi:  
589 10.1523/JNEUROSCI.5349-11.2012
- 590 Goto, Y., & Grace, A. A. (2008). Limbic and cortical information processing in the nucleus accumbens. *Trends in Neurosciences*,  
591 31(11), 552–558. doi: 10.1016/j.tins.2008.08.002
- 592 Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., ... Steele, J. D. (2011). Expected value and prediction  
593 error abnormalities in depression and schizophrenia. *Brain*, 134(6), 1751–1764. doi: 10.1093/brain/awr059
- 594 Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a  
595 Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. doi: 10.1037/h0059831
- 596 Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., ... Berke, J. D. (2015).  
597 Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1), 117–126. doi: 10.1038/nn.4173
- 598 Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014). Phasic Dopamine Release in the Rat Nucleus Ac-  
599 cumbens Symmetrically Encodes a Reward Prediction Error Term. *The Journal of Neuroscience*, 34(3), 698–704. doi:  
600 10.1523/JNEUROSCI.2489-13.2014
- 601 Hassani, O. K., Cromwell, H. C., & Schultz, W. (2001). Influence of Expectation of Different Rewards on Behavior-Related  
602 Neuronal Activity in the Striatum. *Journal of Neurophysiology*, 85(6), 2477–2489. doi: 10.1152/jn.2001.85.6.2477
- 603 Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: the stimulus-reinforcer relation and directed action*. Psychonomic Society. doi:  
604 10.1152/jn.2001.85.6.2477
- 605 Hill, D. N., Mehta, S. B., & Kleinfeld, D. (2011). Quality Metrics to Accompany Spike Sorting of Extracellular Signals. *Journal of  
606 Neuroscience*, 31(24), 8699–8705. doi: 10.1523/JNEUROSCI.0971-11.2011
- 607 Holland, P. C. (1992). Occasion setting in pavlovian conditioning. *Psychology of Learning and Motivation*, 28(C), 69–125. doi:  
608 10.1016/S0079-7421(08)60488-0

- 609 Hollerman, J. R., Tremblay, L., & Schultz, W. (1998). Influence of Reward Expectation on Behavior-Related Neuronal Activity in  
610 Primate Striatum. *Journal of Neurophysiology*, 80(2), 947–963. doi: 10.1152/jn.1998.80.2.947
- 611 Honey, R. C., Iordanova, M. D., & Good, M. (2014). Associative structures in animal learning: Dissociating elemental and  
612 configural processes. *Neurobiology of Learning and Memory*, 108, 96–103. doi: 10.1016/j.nlm.2013.06.002
- 613 Hyman, S. E., Malenka, R. C., & Nestler, E. J. (2006). NEURAL MECHANISMS OF ADDICTION: The Role of Reward-Related  
614 Learning and Memory. *Annual Review of Neuroscience*, 29(1), 565–598. doi: 10.1146/annurev.neuro.29.051605.113009
- 615 Ikemoto, S. (2007). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-  
616 olfactory tubercle complex. *Brain Research Reviews*, 56(1), 27–78. doi: 10.1016/j.brainresrev.2007.05.004
- 617 Ishikawa, A., Ambroggi, F., Nicola, S. M., & Fields, H. L. (2008). Dorsomedial Prefrontal Cortex Contribution to Behav-  
618 ioral and Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, 28(19), 5088–5098. doi:  
619 10.1523/JNEUROSCI.0253-08.2008
- 620 Joel, D., Niv, Y., & Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives.  
621 *Neural Networks*, 15(4-6), 535–547. doi: 10.1016/S0893-6080(02)00047-3
- 622 Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., ... Lissek, S. (2017).  
623 Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174(2), 125–134. doi:  
624 10.1176/appi.ajp.2016.15121549
- 625 Kaczkurkin, A. N., & Lissek, S. (2013). Generalization of Conditioned Fear and Obsessive-Compulsive Traits. *Journal of Psychol-  
626 ogy & Psychotherapy*, 7, 3. doi: 10.4172/2161-0487.S7-003
- 627 Kalivas, P. W., & Volkow, N. D. (2005). The neural basis of addiction: A pathology of motivation and choice. *American Journal of  
628 Psychiatry*, 162(8), 1403–1413. doi: 10.1176/appi.ajp.162.8.1403
- 629 Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in  
630 schizophrenia. *American Journal of Psychiatry*, 160(1), 13–23. doi: 10.1176/appi.ajp.160.1.13
- 631 Khamassi, M., & Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and  
632 model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, 6, 79. doi: 10.3389/fnbeh.2012.00079
- 633 Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., & Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal  
634 neurons of behaving rats. *European Journal of Neuroscience*, 28(9), 1849–1866. doi: 10.1111/j.1460-9568.2008.06480.x
- 635 Lansink, C. S., Jackson, J. C., Lankelma, J. V., Ito, R., Robbins, T. W., Everitt, B. J., & Pennartz, C. M. A. (2012). Reward Cues  
636 in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles. *Journal of  
637 Neuroscience*, 32(36), 12444–12459. doi: 10.1523/JNEUROSCI.0593-12.2012
- 638 Lansink, C. S., Meijer, G. T., Lankelma, J. V., Vinck, M. A., Jackson, J. C., & Pennartz, C. M. A. (2016). Reward Expectancy  
639 Strengthens CA1 Theta and Beta Band Synchronization and Hippocampal-Ventral Striatal Coupling. *Journal of Neuro-  
640 science*, 36(41), 10598–10610. doi: 10.1523/JNEUROSCI.0682-16.2016
- 641 Lavoie, A. M., & Mizumori, S. J. (1994). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of

- 642 rats. *Brain Research*, 638(1-2), 157–168. doi: 10.1016/0006-8993(94)90645-9
- 643 Lee, D., Seo, H., & Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of*  
644 *Neuroscience*, 35(1), 287–308. doi: 10.1146/annurev-neuro-062111-150512
- 645 Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*,  
646 22, 1027–1038. doi: 10.1016/j.conb.2012.06.001
- 647 Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized anxiety disorder is associated with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75(11), 909–915. doi:  
648 10.1016/j.biopsych.2013.07.025
- 650 Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the  
651 fMRI signal. *Nature*, 412(6843), 150–157. doi: 10.1038/35084005
- 652 Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and*  
653 *Behavioral Neuroscience*, 9(4), 343–364. doi: 10.3758/CABN.9.4.343
- 654 Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature*  
655 *Neuroscience*, 14(2), 154–162. doi: 10.1038/nn.2723
- 656 Malhotra, S., Cross, R. W., Zhang, A., & Van Der Meer, M. A. A. (2015). Ventral striatal gamma oscillations are highly variable  
657 from trial to trial, and are dominated by behavioural state, and only weakly influenced by outcome value. *European Journal*  
658 *of Neuroscience*, 42(10), 2818–2832. doi: 10.1038/nn.2723
- 659 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011). Ventral Striatum and Orbitofrontal Cortex  
660 Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, 31(7), 2700–  
661 2705. doi: 10.1523/JNEUROSCI.5499-10.2011
- 662 McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., & Nicola, S. M. (2013). Invigoration of reward seeking by cue and proximity  
663 encoding in the nucleus accumbens. *Neuron*, 78(5), 910–922. doi: 10.1016/j.neuron.2013.04.010
- 664 Mulder, A. B., Shibata, R., Trullier, O., & Wiener, S. I. (2005). Spatially selective reward site responses in tonically active neurons  
665 of the nucleus accumbens in behaving rats. *Experimental Brain Research*, 163(1), 32–43. doi: 10.1007/s00221-004-2135-3
- 666 Mulder, A. B., Tabuchi, E., & Wiener, S. I. (2004). Neurons in hippocampal afferent zones of rat striatum parse routes into  
667 multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7), 1923–1932. doi: 10.1111/j.1460-  
668 9568.2004.03301.x
- 669 Nicola, S. M. (2004). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a Discrimi-  
670 native Stimulus Task. *Journal of Neurophysiology*, 91(4), 1840–1865. doi: 10.1152/jn.00657.2003
- 671 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of  
672 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49), 16585–  
673 16600. doi: 10.1523/JNEUROSCI.3958-10.2010
- 674 Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psy-*

- chopharmacology, 191(3), 507–520. doi: 10.1007/s00213-006-0502-4

676 Noonan, M. P., Chau, B. K., Rushworth, M. F., & Fellows, L. K. (2017). Contrasting Effects of Medial and Lateral Orbitofrontal  
677 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29), 7023–7035.  
678 doi: 10.1523/JNEUROSCI.0692-17.2017

679 Paxinos, G., & Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates* (4th ed.). San Diego: Academic Press.doi:  
680 10.1523/JNEUROSCI.0692-17.2017

681 Peters, J., & Büchel, C. (2009). Overlapping and distinct neural systems code for subjective value during intertemporal and risky  
682 decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(50), 15727–34. doi:  
683 10.1523/JNEUROSCI.3489-09.2009

684 Rescorla, R. A., & Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and  
685 Instrumental Learning. *Psychological Review*, 74(3), 151–182. doi: 10.1037/h0024475

686 Robinson, T. E., & Flagel, S. B. (2009). Dissociating the Predictive and Incentive Motivational Properties of Reward-Related Cues  
687 Through the Study of Individual Differences. *Biological Psychiatry*, 65(10), 869–873. doi: 10.1016/j.biopsych.2008.09.006

688 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value  
689 of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42),  
690 13365–13376. doi: 10.1523/JNEUROSCI.2572-09.2009

691 Roitman, M. F., Wheeler, R. A., & Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for reward-  
692 ing and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4), 587–597. doi:  
693 10.1016/j.neuron.2004.12.055

694 Saddoris, M. P., Stamatakis, A., & Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nucleus ac-  
695 cumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*, 33(12),  
696 2274–2287. doi: 10.1111/j.1460-9568.2011.07683.x

697 Salamone, J. D., & Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3), 470–485.  
698 doi: 10.1016/j.neuron.2012.10.021

699 Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32. doi:  
700 10.1038/nrn.2015.26

701 Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the  
702 expectation of reward. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 12(12), 4595–  
703 610. doi: 10.1523/JNEUROSCI.12-12-04595.1992

704 Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.  
705 doi: 10.1126/science.275.5306.1593

706 Sescousse, G., Li, Y., & Dreher, J.-C. (2015). A common currency for the computation of motivational values in the human striatum.  
707 *Social Cognitive and Affective Neuroscience*, 10(4), 467–473. doi: 10.1093/scan/nsu074

- 708 Setlow, B., Schoenbaum, G., & Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning.  
709 *Neuron*, 38(4), 625–636. doi: 10.1016/S0896-6273(03)00264-2
- 710 Shidara, M., Aigner, T. G., & Richmond, B. J. (1998). Neuronal signals in the monkey ventral striatum related to progress through  
711 a predictable series of trials. *Journal of Neuroscience*, 18(7), 2613–25. doi: 10.1016/S0896-6273(03)00264-2
- 712 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., & Buszaki, G. (2017). Cocaine place conditioning strengthens location-  
713 specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, 1–10. doi: 10.1101/105890
- 714 Sleezer, B. J., Castagno, M. D., & Hayden, B. Y. (2016). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.  
715 *Journal of Neuroscience*, 36(44), 11223–11237. doi: 10.1523/JNEUROSCI.1766-16.2016
- 716 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., & Hayden, B. Y. (2016). Neuronal selectivity  
717 for spatial positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3), 1098–1111. doi:  
718 10.1152/jn.00325.2015
- 719 Sugam, J. A., Saddoris, M. P., & Carelli, R. M. (2014). Nucleus accumbens neurons track behavioral preferences and reward  
720 outcomes during risky decision making. *Biological Psychiatry*, 75(10), 807–816. doi: 10.1016/j.biopsych.2013.09.010
- 721 Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction* (Vol. 9) (No. 5). MIT Press, Cambridge, MA. doi:  
722 10.1109/TNN.1998.712192
- 723 Tabuchi, E. T., Mulder, A. B., & Wiener, S. I. (2000). Position and behavioral modulation of synchronization of hippocam-  
724 pal and accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6), 717–728. doi: 10.1002/1098-  
725 1063(2000)10:6;717::AID-HIPO1009\_3.0.CO;2-3
- 726 Takahashi, Y. K., Langdon, A. J., Niv, Y., & Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors  
727 Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1), 182–193. doi:  
728 10.1016/j.neuron.2016.05.015
- 729 Tingley, D., & Buszaki, G. (2018). Transformation of a Spatial Map across the Hippocampal-Lateral Septal Circuit. *Neuron*, 98(6),  
730 1229–1242.e5. doi: 10.1016/J.NEURON.2018.04.028
- 731 van der Meer, M. A. A., & Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward  
732 Information. *Journal of Neuroscience*, 31(8), 2843–2854. doi: 10.1523/JNEUROSCI.4869-10.2011
- 733 West, E. A., & Carelli, R. M. (2016). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after  
734 Reinforcer Devaluation. *Journal of Neuroscience*, 36(4), 1128–1139. doi: 10.1523/JNEUROSCI.2976-15.2016
- 735 Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., & Mulder, A. B. (2003). Spatial and behavioral correlates in  
736 nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,  
737 1250(C), 275–292. doi: 10.1016/S0531-5131(03)00978-6
- 738 Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral  
739 and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, 24(12), 2923–2933. doi:  
740 10.1523/JNEUROSCI.5282-03.2004