

# **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 of  
6 outcome-predictive cues, and recorded single-unit activity in the NAc during performance. We found that  
7 coding of various cue features occurred alongside coding of expected outcome. Furthermore, this coding  
8 persisted both during a delay period, after the rat made a decision and was waiting for an outcome, and  
9 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.

## <sup>12</sup> Introduction

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

<sup>23</sup> Much of our understanding of NAc function comes from studies of how cues that predict motivationally  
<sup>24</sup> relevant outcomes (e.g. reward) influence behavior and neural activity in the NAc. Task designs that asso-  
<sup>25</sup> ciate such cues with rewarding outcomes provide a convenient access point, eliciting conditioned responses  
<sup>26</sup> such as sign-tracking and goal-tracking (Hearst & Jenkins, 1974; Robinson & Flagel, 2009), pavlovian-  
<sup>27</sup> instrumental transfer (Estes, 1943; Rescorla & Solomon, 1967) and enhanced response vigor (Nicola, 2010;  
<sup>28</sup> Niv, Daw, Joel, & Dayan, 2007), which tend to be affected by NAc manipulations (Chang, Wheeler, &  
<sup>29</sup> Holland 2012; Corbit & Balleine 2011; Flagel et al. 2011; although not always straightforwardly; Chang &  
<sup>30</sup> Holland 2013; Giertler, Bohn, & Hauber 2004). Similarly, analysis of RPEs typically proceeds by estab-  
<sup>31</sup> lishing an association between a cue and subsequent reward, with NAc responses transferring from outcome  
<sup>32</sup> to the cue with learning (Day, Roitman, Wightman, & Carelli, 2007; Roitman, Wheeler, & Carelli, 2005;  
<sup>33</sup> Schultz, Dayan, & Montague, 1997; Setlow, Schoenbaum, & Gallagher, 2003).

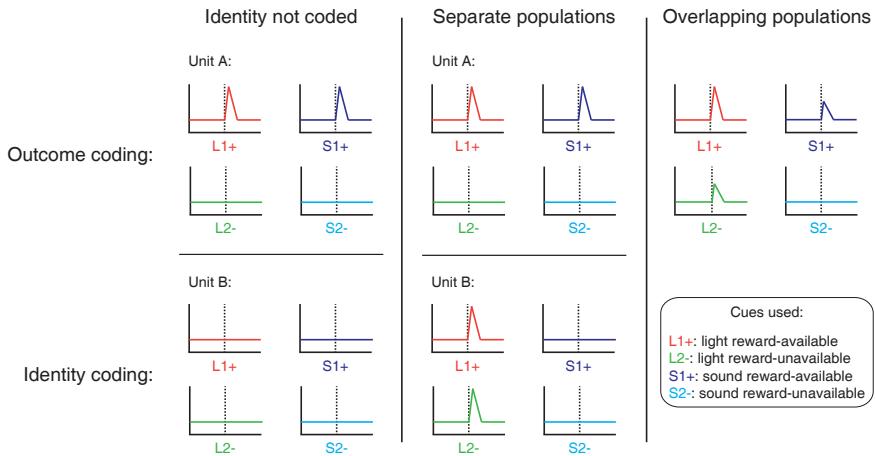
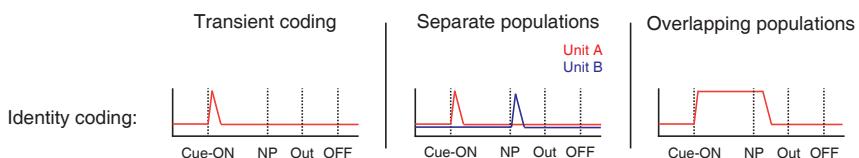
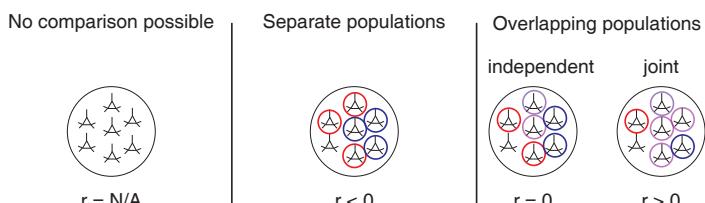
<sup>34</sup> Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues  
<sup>35</sup> (Atallah, McCool, Howe, & Graybiel, 2014; Bissonette et al., 2013; Cooch et al., 2015; Cromwell & Schultz,

36 2003; Day, Wheeler, Roitman, & Carelli, 2006; Goldstein et al., 2012; Hassani, Cromwell, & Schultz, 2001;  
37 Hollerman, Tremblay, & Schultz, 1998; Lansink et al., 2012; McGinty, Lardeux, Taha, Kim, & Nicola,  
38 2013; Nicola, 2004; Roesch, Singh, Brown, Mullins, & Schoenbaum, 2009; Roitman et al., 2005; Saddoris,  
39 Stamatakis, & Carelli, 2011; Schultz, Apicella, Scarnati, & Ljungberg, 1992; Setlow et al., 2003; Sugam,  
40 Saddoris, & Carelli, 2014; West & Carelli, 2016), much less is known about how outcome-predictive cues  
41 themselves are encoded in the NAc (but see; Sleezer, Castagno, & Hayden, 2016). This is an important  
42 issue for at least two reasons. First, in reinforcement learning, motivationally relevant outcomes are typi-  
43 cally temporally delayed relative to the cues that predict them. In order to solve the problem of assigning  
44 credit (or blame) across such temporal gaps, some trace of preceding activity needs to be maintained (Lee  
45 et al., 2012; Sutton & Barto, 1998). For example, if you become ill after eating food X in restaurant A,  
46 depending on if you remember the identity of the restaurant or the food at the time of illness, you may learn  
47 to avoid all restaurants, restaurant A only, food X only, or the specific pairing of X-in-A. Therefore, a com-  
48 plete understanding of what is learned following feedback requires understanding what trace is maintained.  
49 Since NAc is a primary target of dopamine signals interpretable as RPEs, and NAc lesions impair RPEs  
50 related to timing, its activity trace will help determine what can be learned when RPEs arrive (Hamid et  
51 al., 2015; Hart, Rutledge, Glimcher, & Phillips, 2014; Ikemoto, 2007; McDannald, Lucantonio, Burke, Niv,  
52 & Schoenbaum, 2011; Takahashi et al., 2016). Similarly, in a neuroeconomic framework, NAc is thought  
53 to represent a domain-general subjective value singal for different offers (Bartra, McGuire, & Kable, 2013;  
54 Levy & Glimcher, 2012; Peters & Büchel, 2009; Sescousse, Li, & Dreher, 2015); having a representation of  
55 the offer itself alongside this value signal would provide a potential neural substrate for updating offer value.

56 Second, for ongoing behavior, the relevance of cues typically depends on context. In experimental set-  
57 tings, context may include the identity of a preceding cue, spatial or configural arrangements (Bouton, 1993;  
58 Holland, 1992; Honey, Iordanova, & Good, 2014), and unsignaled rules as occurs in set shifting and other  
59 cognitive control tasks (Cohen & Servan-Schreiber, 1992; Floresco, Ghods-Sharifi, Vexelman, & Magyar,  
60 2006; Grant & Berg, 1948; Sleezer et al., 2016). In such situations, the question arises how selective, context-  
61 dependent processing of outcome-predictive cues is implemented. For instance, is there a gate prior to NAc

62 such that only currently relevant cues are encoded in NAc, or are all cues represented in NAc but their current  
63 values dynamically updated (FitzGerald, Schwartenbeck, & Dolan, 2014; Goto & Grace, 2008; Sleezer et  
64 al., 2016). Representation of cue identity would allow for context-dependent mapping of outcomes predicted  
65 by specific cues.

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

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

**Figure 1:** Schematic of potential coding strategies 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 PETHs illustrating putative responses to different cues under different hypotheses of how cue identity (light, sound) 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. 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 onset of cue, 'NP' signifies when the rat holds a nosepoke at a reward receptacle, 'Out' signifies when the outcome is revealed, 'OFF' signifies when the cue turns off. 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).

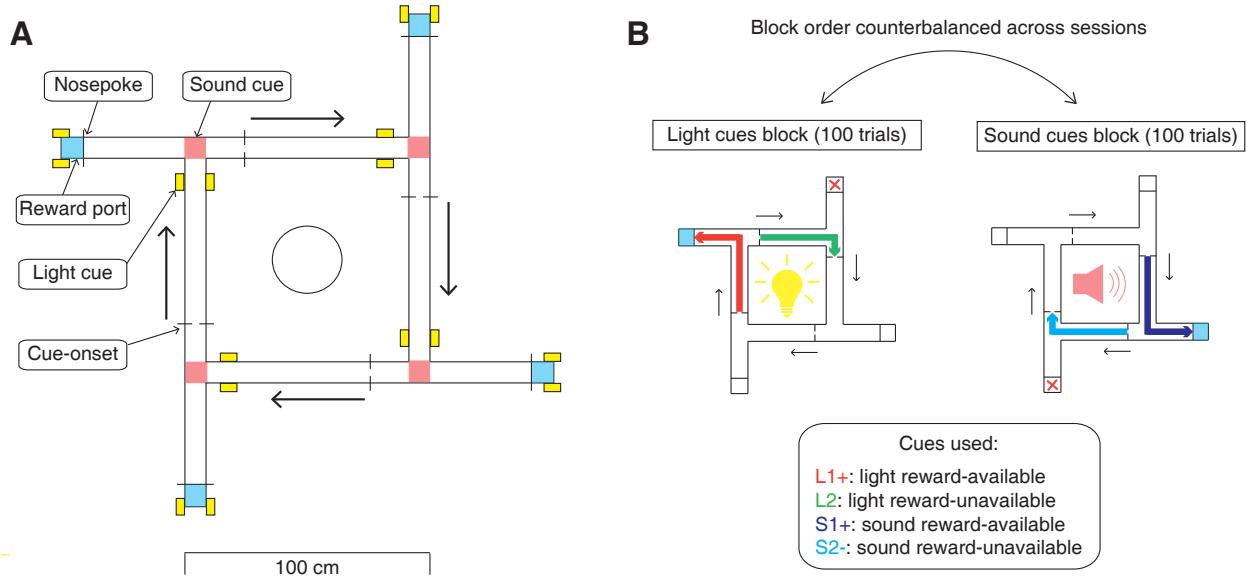
**Figure 1:** (Previous page.) **C:** Schematic pool of NAc units, illustrating the different possible outcomes for each of the analyses.  $R$  values represent the correlation between sets of recoded regression coefficients (see text for analysis details). Left panel: Cue identity is not coded (A: left 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 represents 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 overlapping population. Note: The same hypotheses apply 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.

## 73 Results

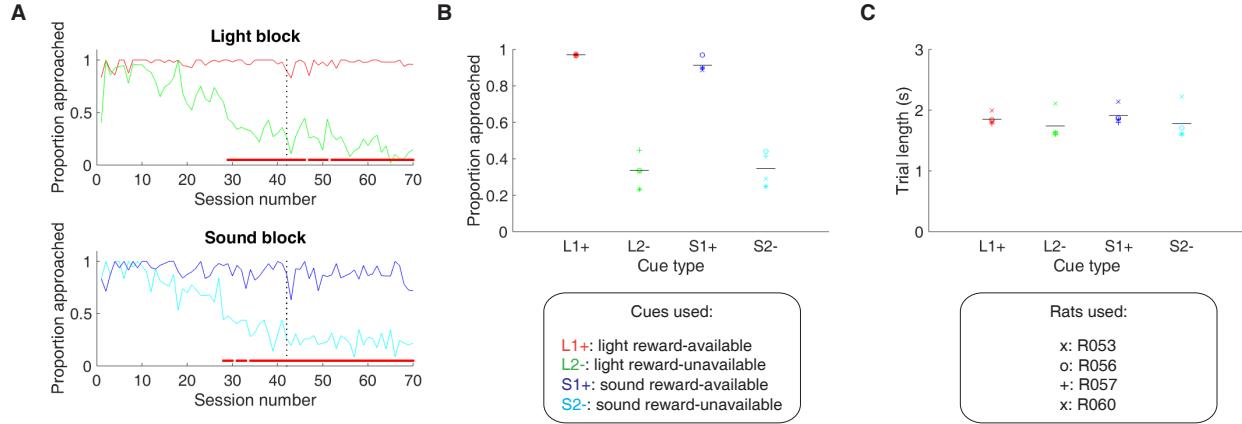
### 74 Behavior

75 Rats were trained to discriminate between cues signaling the availability and absence of reward on a square  
76 track with four identical arms for two distinct set of cues (Figure 2). During each session, rats were pre-  
77 sented sequentially with two behavioral blocks containing cues from different sensory modalities, a light and  
78 a sound block, with each block containing a cue that signalled the availability of reward (reward-available),  
79 and a cue that signalled the absence of reward (reward-unavailable). To maximize reward receipt, rats should  
80 approach reward sites on reward-available trials, and skip reward sites on reward-unavailable trials (see Fig-  
81 ure 3A for an example learning curve). All four rats learned to discriminate between the reward-available  
82 and reward-unavailable cues for both the light and sound blocks as determined by reaching significance ( $p <$   
83 .05) on a daily chi-square test comparing approach behavior for reward-available and reward-unavailable  
84 cues for each block, for at least three consecutive days (range for time to criterion: 22 - 57 days). Mainte-  
85 nance of behavioral performance during recording sessions was assessed using linear mixed effects models  
86 for both proportion of trials where the rat approached the receptacle, and trial length. Analyses revealed  
87 that the likelihood of a rat to make an approach was influenced by whether a reward-available or reward-

88 unavailable cue was presented, but was not significantly modulated by whether the rat was presented with a  
89 light or sound cue (Percentage approached: light reward-available = 97%; light reward-unavailable = 34%;  
90 sound reward-available = 91%; sound reward-unavailable 35%; cue identity  $p = .115$ ; cue outcome  $p < .001$ ;  
91 Figure 3B). A similar trend was seen with the length of time taken to complete a trial (Trial length: light  
92 reward-available = 1.85 s; light reward-unavailable = 1.74 s; sound reward-available = 1.91 s; sound reward-  
93 unavailable 1.78 s; cue identity  $p = .106$ ; cue outcome  $p < .001$ ; Figure 3C). Thus, during recording rats  
94 successfully discriminated the cues according to whether or not they signaled the availability of reward at  
95 the reward receptacle.



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



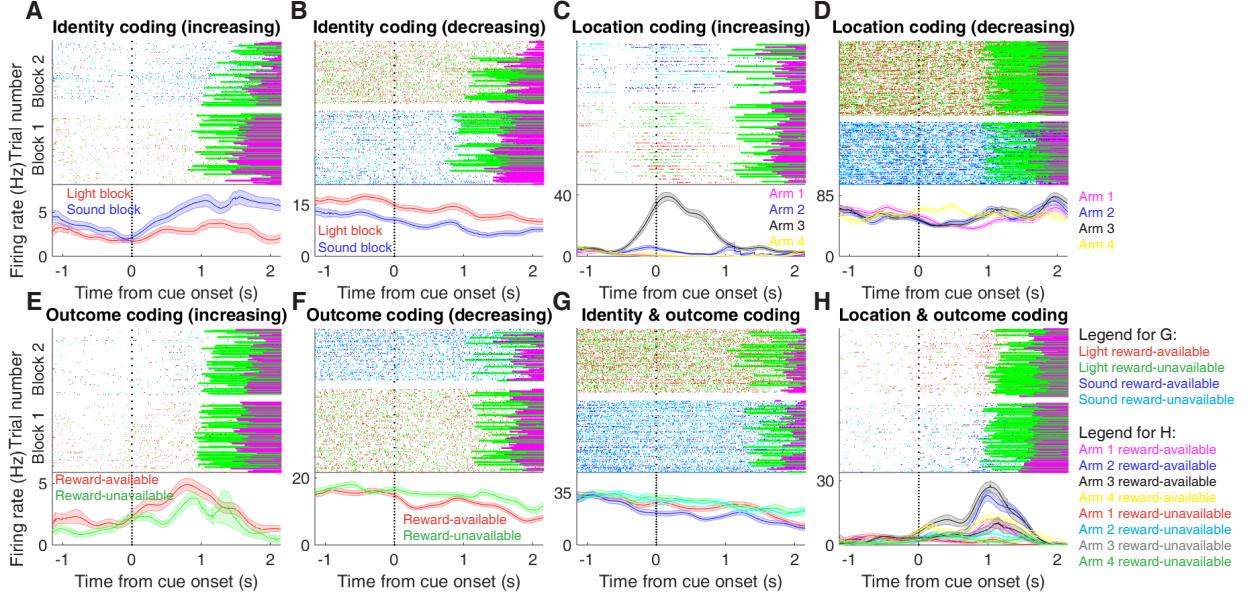
**Figure 3:** Performance on the behavioral task. **A.** Example learning curves across sessions from a single subject (R060) showing the proportion of trials approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light 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. **B-C:** Summary of behavioral performance during recording sessions for each rat. **B:** 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 and reward-unavailable cues, as indicated by the clear difference of proportion approached between reward-available (~90% approached) and reward-unavailable cues (~30% approached), for both blocks (see Results for statistics). **C:** Average trial length for each cue. Note that the time to complete a trial was comparable for light and sound cues.

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

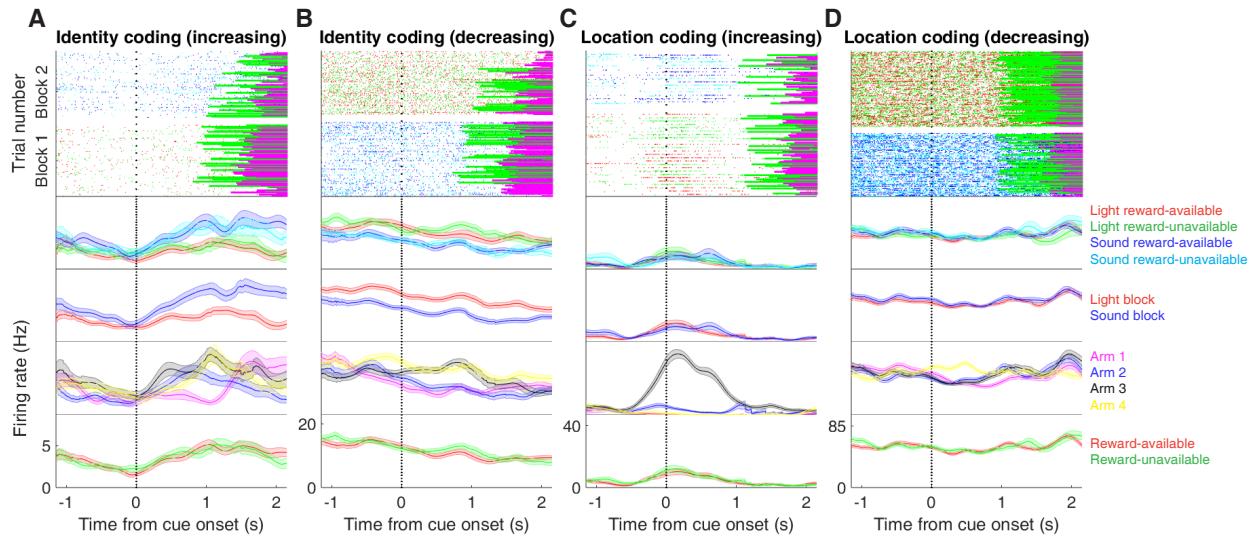
97 We sought to address which parameters of our task were encoded by NAc activity, specifically whether the  
98 NAc encodes aspects of motivationally relevant cues not directly tied to reward, such as the identity and  
99 location of the cue, and whether this coding is accomplished by separate or overlapping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

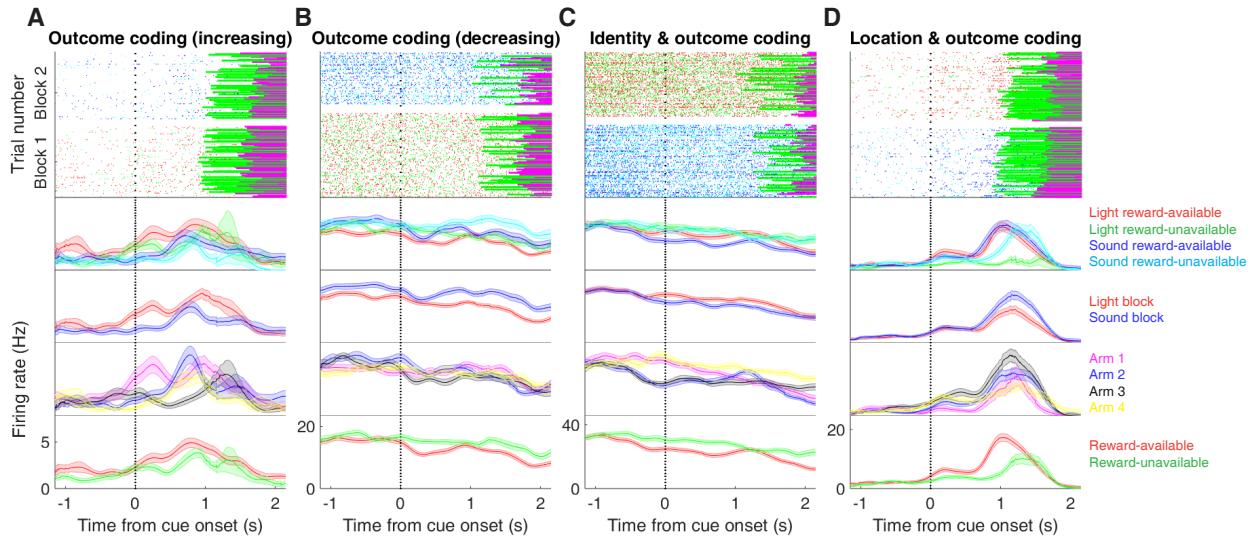
**Table 1:** Overview of recorded NAc units and their relationship to task variables at various time epochs. Percentage is relative to the number of cue-modulated units (n = 133).



**Figure 4:** 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: PETHs showing the average smoothed firing rate for the unit for trials during light (red) and sound (blue) blocks, aligned to cue-onset. Lightly shaded area indicates standard error of the mean. Note this unit showed a larger increase in firing to sound cues. **B:** An example of a unit that was responsive to cue identity as in A, but for a unit that showed a decrease in firing to the cue. Note the sustained higher firing rate during the light block. **C-D:** Cue-modulated units that exhibited location coding. Each color in the PETHs represents average firing response for a different cue location. **C:** The firing rate of this unit only changed on arm 3 of the task. **D:** Firing rate decreased for this unit on all arms but arm 4. **E-F:** Cue-modulated units that exhibited outcome coding, with the PETHs comparing reward-available (red) and reward-unavailable (green) trials. **E:** This unit showed a slightly higher response during presentation of reward-available cues. **F:** This unit showed a dip in firing when presented with reward-available cues. **G-H:** Examples of cue-modulated units that encoded multiple cue features. **G:** This unit showed both identity and outcome coding. **H:** An example of a unit that coded for both identity and location.

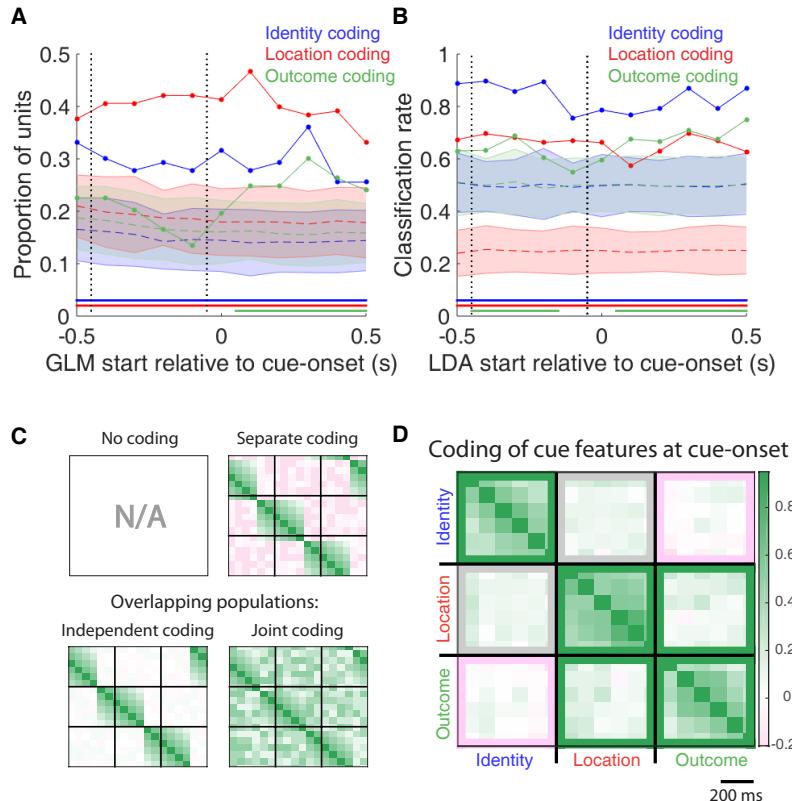


**Figure supplement 1:** Expanded examples of cue-modulated NAc units influenced by different task parameters for Figure 4A-D, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).



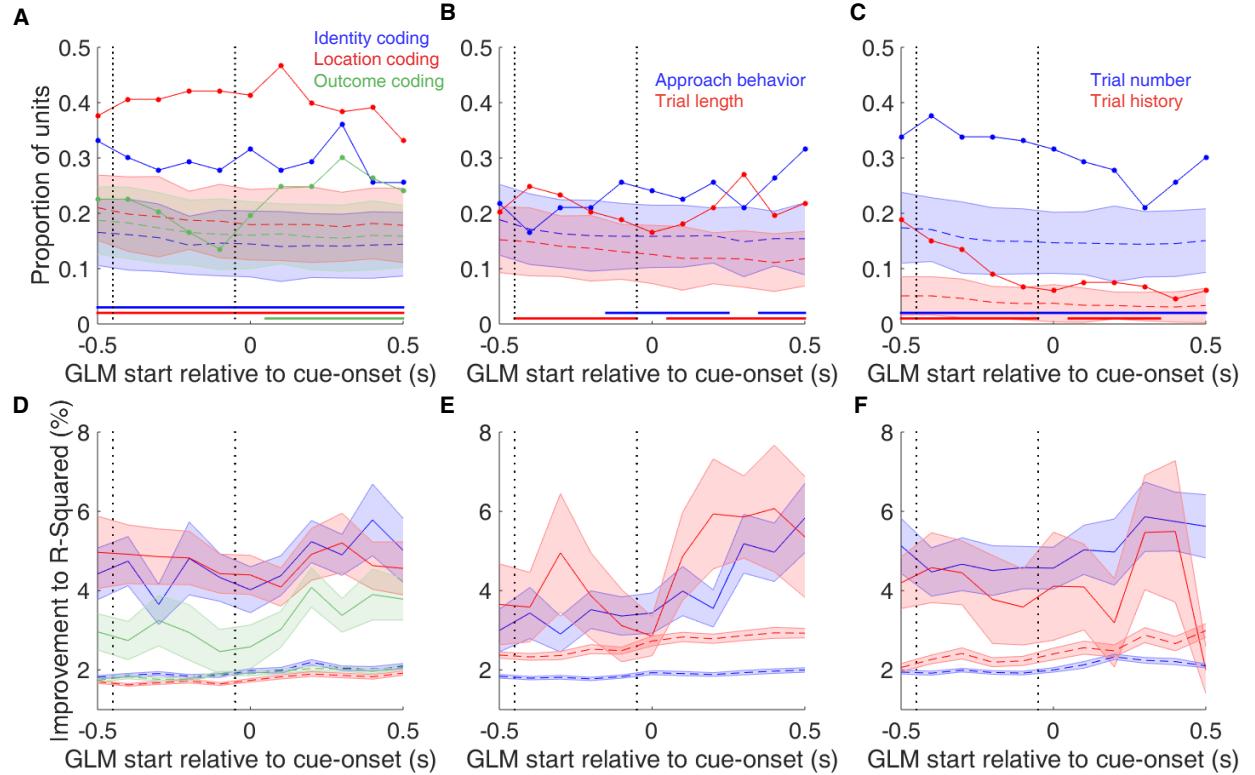
**Figure supplement 2:** Expanded examples of cue-modulated NAc units influenced by different task parameters for Figure 4E-H, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).

112 To characterize more formally whether these cue-modulated responses were influenced by various aspects  
113 of the task, we fit a sliding window generalized linear model (GLM) to the firing rate of each cue-modulated  
114 unit surrounding cue-onset, using a forward selection stepwise procedure for variable selection, a bin size of  
115 500 ms for firing rate and a step size of 100 ms for the sliding window. Fitting GLMs to all trials within a  
116 session revealed that a variety of task parameters accounted for a significant portion of firing rate variance in  
117 NAc cue-modulated units (Figure 5A, supplement 1, supplement 2, Table 1). Notably, a significant proportion  
118 of units discriminated between the light and sound block (*identity coding*: ~30% of cue-modulated units,  
119 accounting for ~5% of firing rate variance) or the arms of the apparatus (*location coding*: ~40% of cue-  
120 modulated units, accounting for ~4% of firing rate variance) throughout the entire window surrounding  
121 cue-onset, whereas a substantial proportion of units discriminating between the common portion of reward-  
122 available and reward-unavailable trials (*outcome coding*: ~25% of cue-modulated units, accounting for ~4%  
123 of firing rate variance) was not observed until after the onset of the cue (z-score > 1.96 when comparing  
124 observed proportion of units to a shuffled distribution obtained when shuffling the firing rates of each unit  
125 across trials before running the GLM), suggesting that the NAc encodes features of outcome-predictive cues  
126 in addition to expected outcome.

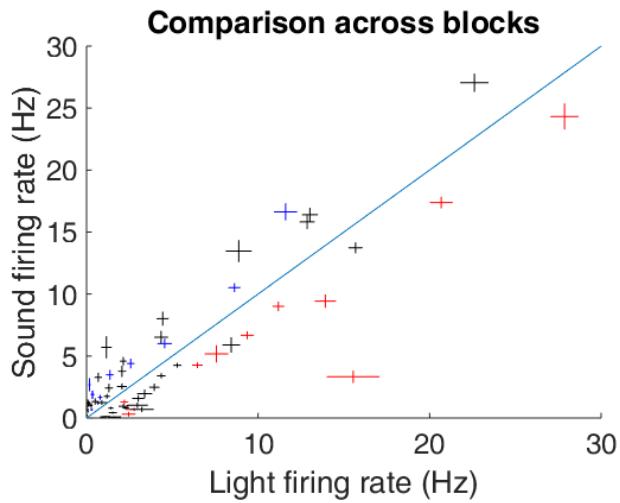


**Figure 5:** Summary of influence of cue features on cue-modulated NAc units at time points surrounding cue-onset. **A:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), location (red solid line), and 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,

**Figure 5:** (Previous page.) with pink indicating separate populations ( $z$ -score  $< -1.96$ ), grey indicating overlapping but independent populations, and green indicating joint overlapping populations ( $z$ -score  $> 1.96$ ).



**Figure 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 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.

127 To assess what information may be encoded at the population level, we trained a classifier on a pseudoensemble  
128 of the 133 cue-modulated units (Figure 5B). Specifically, we used the firing rate of each unit for each  
129 trial as an observation, and different cue conditions as trial labels (e.g. light block, sound block). A lin-  
130 ear discriminant analysis (LDA) classifier with 10-fold cross-validation could correctly predict a trial above  
131 chance levels for the identity and location of a cue across all time points surrounding cue-onset ( $z$ -score  $>$   
132 1.96 when comparing classification accuracy of data versus a shuffled distribution), whereas the ability to  
133 predict whether a trial was reward-available or reward-unavailable (outcome coding) was not significantly  
134 higher than the shuffled distribution for the time point containing 500 ms of pre-cue firing rate, and increased  
135 gradually as a trial progressed, providing evidence that cue information is also present in the pseudoensemble  
136 level.

137 To quantify the overlap of cue feature coding we correlated recoded beta coefficients from the GLMs, as-  
138 signing a value of ‘1’ if a cue feature was a significant predictor for that unit and ‘0’ if not, and compared  
139 the obtained correlations to the shuffled data generated for the GLMs (Figure 1A,C, 5C,D). This revealed  
140 that identity was coded separately from outcome (mean  $r = .009$ ; -6.74 standard deviations from the shuffled  
141 mean), and independently from location (mean  $r = .097$ ; 0.58 standard deviations from the shuffled mean),  
142 while location and outcome were coded by a joint population of units (mean  $r = .119$ ; 2.23 standard devia-  
143 tions from the shuffled mean). Together, these findings show that various cue features are represented in the  
144 NAc at both the single-unit and pseudoensemble level, that identity is coded independently or in separate  
145 populations from location and outcome, respectively, and that location is coded jointly with outcome.

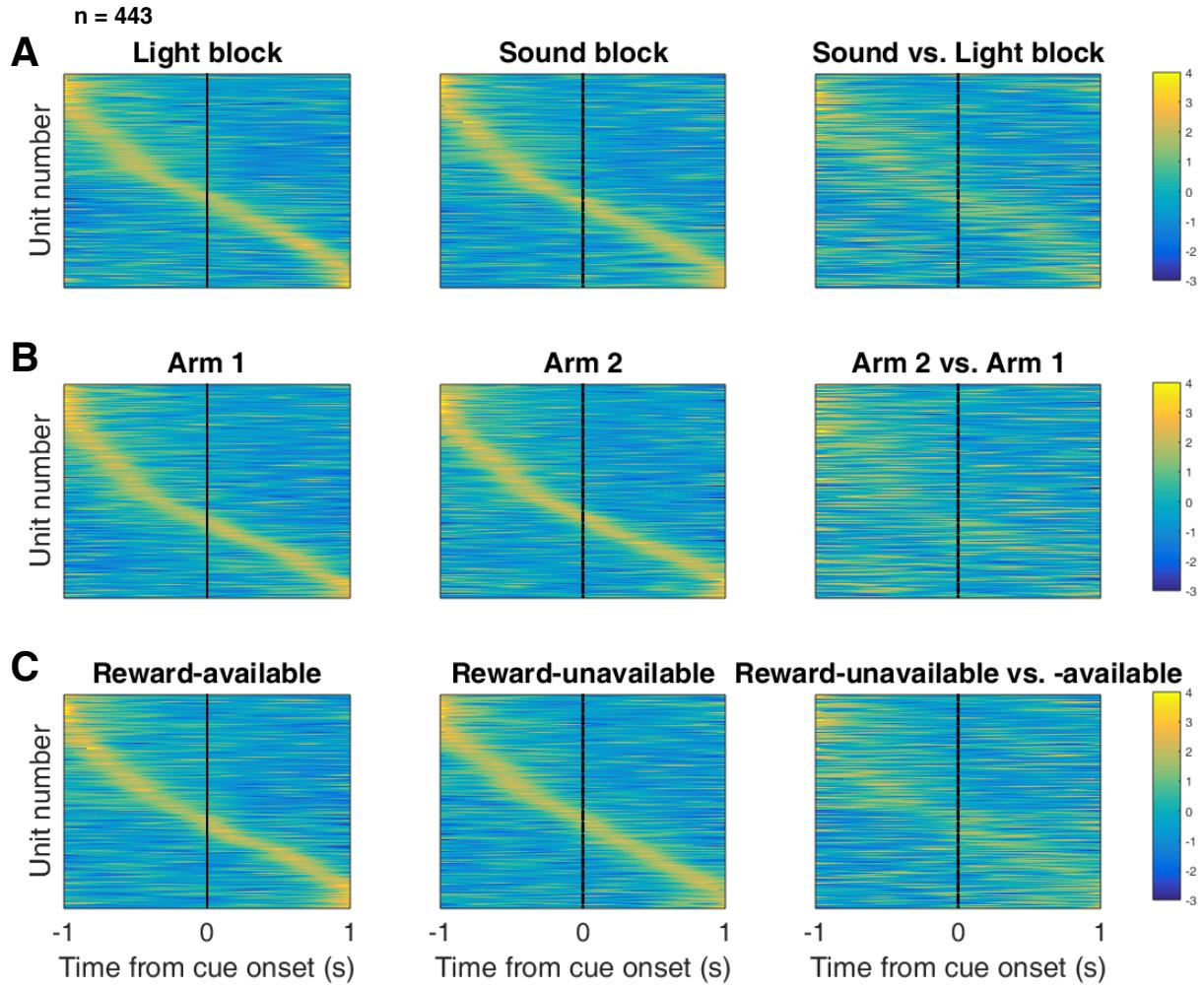
146 **NAc units dynamically segment the task:**

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

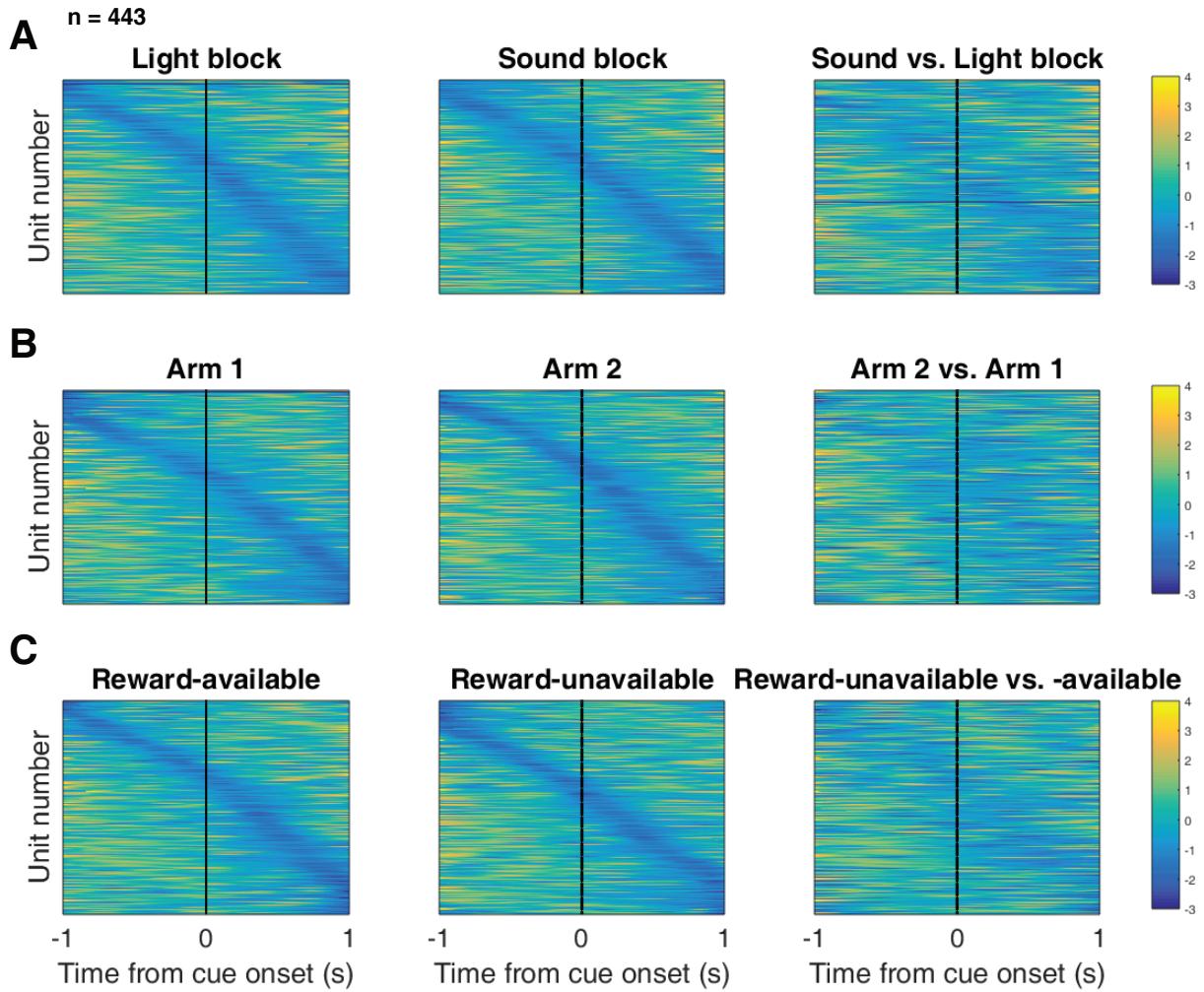
151 of each unit, plotted the normalized firing rates of all units aligned to cue-onset, and sorted them according  
152 to the time of peak firing rate (Figure 6). We did this separately for both the light and sound blocks, and  
153 found a nearly uniform distribution of firing fields in task space that was not limited to alignment to the  
154 cue (Figure 6A). Furthermore, to determine if this population level activity was similar across blocks, we  
155 also organized firing during the sound blocks according to the ordering derived from the light blocks. This  
156 revealed that while there was some preservation of order, the overall firing was qualitatively different across  
157 the two blocks, implying that population activity distinguishes between light and sound blocks.

158 To control for the possibility that any comparison of trials would produce this effect, we divided each block  
159 into two halves and looked at the correlation of the average smoothed firing rates across various combinations  
160 of these halves across our cue-onset centered epoch to see if the across block comparisons were less corre-  
161 lated than the within block correlations. A linear mixed effects model revealed that within block correlations  
162 (e.g. one half of light trials vs other half of light trials) were higher and more similar than across block corre-  
163 lations (e.g. half of light trials vs half of sound trials) suggesting that activity in the NAc discriminates across  
164 light and sound blocks (within block correlations = .383 (light), .379 (sound); across block correlations =  
165 .343, .338, .337, .348; within vs. within block comparison =  $p = .934$ ; within vs. across block comparisons  
166 =  $p < .001$ ). This process was repeated for cue location (Figure 6B; within block correlations = .369 (arm  
167 1), .350 (arm 2); across block correlations = .290, .286, .285, .291; within vs. within block comparison =  $p$   
168 = .071; within vs. across block comparisons =  $p < .001$ ) and cue outcome (Figure 6C; within block corre-  
169 lations = .429 (reward-available), .261 (reward-unavailable); across block correlations = .258, .253, .255, .249;  
170 within vs. within block comparison =  $p < .001$ ; within vs. across block comparisons =  $p < .001$ ). Notably,  
171 the within condition comparison of reward-unavailable trials was less correlated than reward-available trials,  
172 and more similar to the across condition comparisons, potentially due to the greater behavioral variability  
173 for the reward-unavailable trials. Additionally, given that the majority of our units showed an inhibitory re-  
174 sponse to the cue, we also plotted the firing rates according to the lowest time in firing, and again found some  
175 maintenance of order, but largely different ordering across the two blocks (Figure 6 supplement 1). Together,  
176 this discriminative “tiling” suggests that NAc segmentation of the task is qualitatively different even during

<sup>177</sup> those parts of the task not immediately associated with a specific cue, action, or outcome.



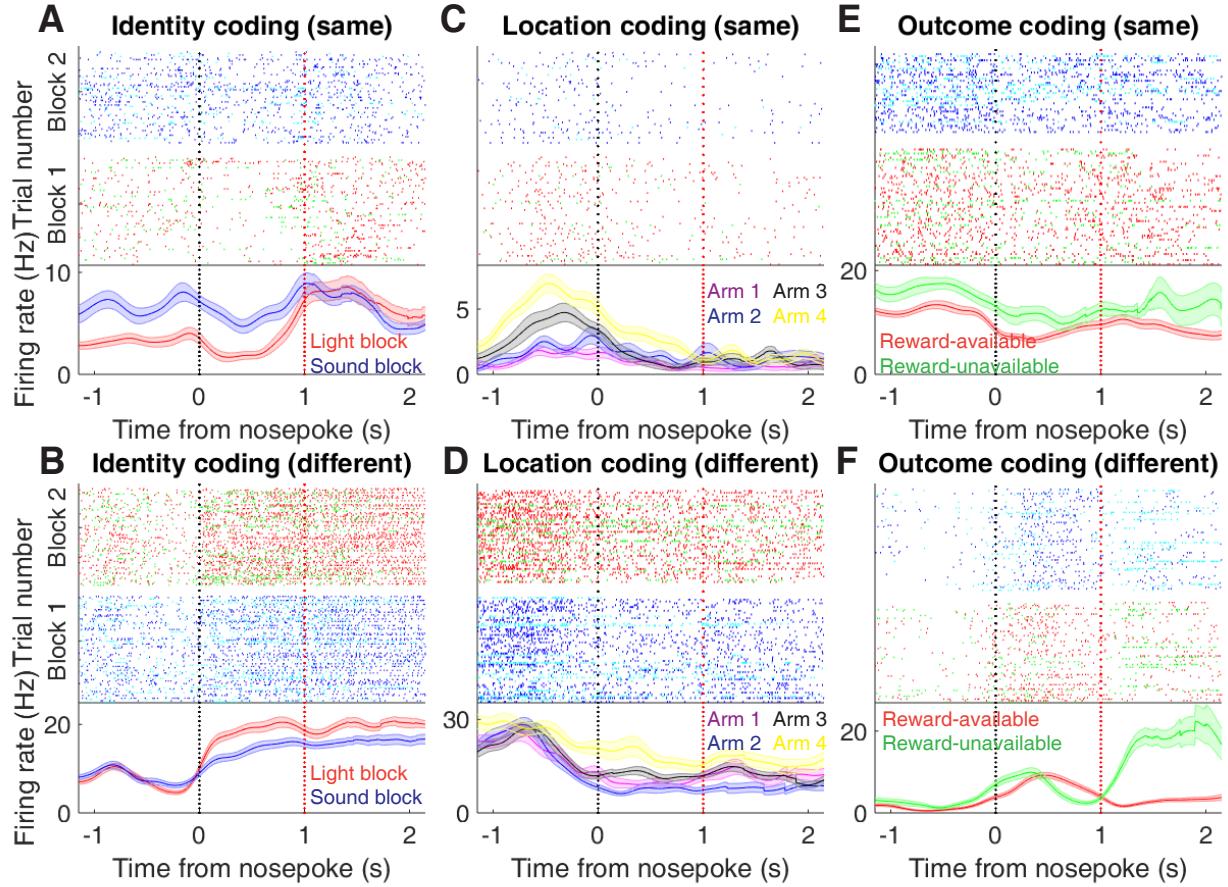
**Figure 6:** Distribution of NAc firing rates across time surrounding cue-onset. Each panel shows normalized (z-score) peak firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue-onset), averaged across all trials for a specific cue type, 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. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C**: Same layout as in A, except that panels now compare reward-available and reward-unavailable trials. Overall, NAc units “tiled” experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.



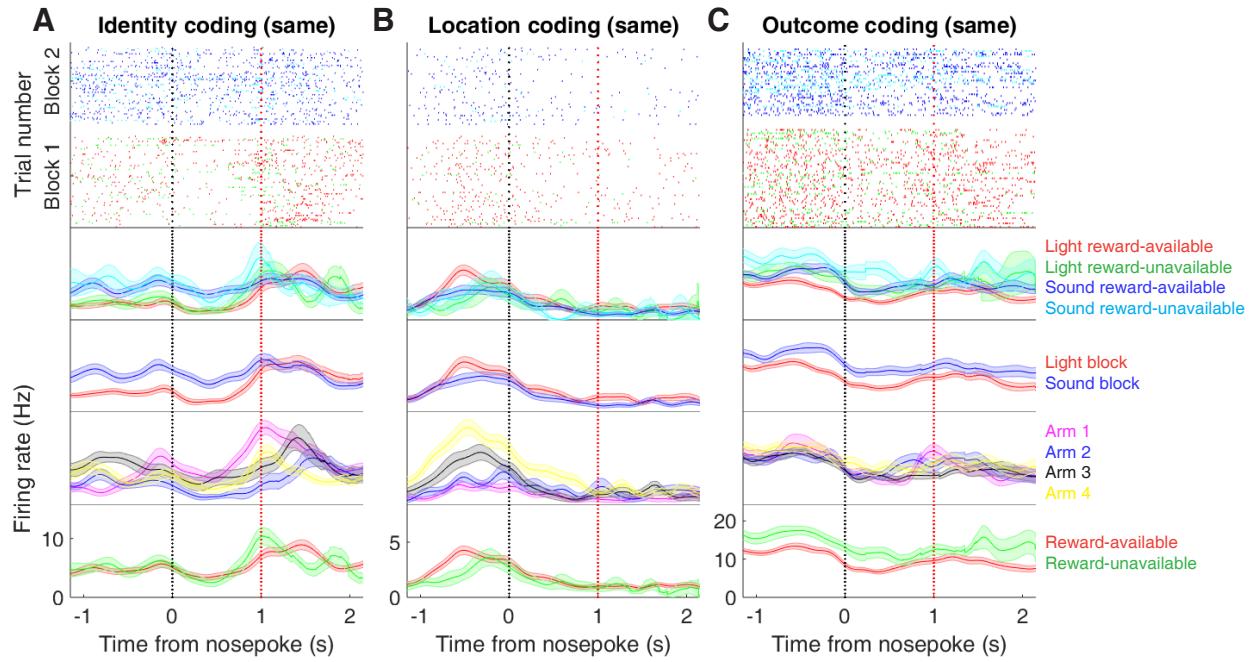
**Figure supplement 1:** Distribution of NAc firing rates across time surrounding cue-onset. Each panel shows normalized (z-score) minimum firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue-onset), averaged across all trials for a specific cue type, indicated by text labels. **A:** Responses during different stimulus blocks as in Figure 6A, but with units ordered according to the time of their minimum firing rate. **B:** Responses during trials on different arms as in Figure 6B, but with units ordered by their minimum firing rate. **C:** Responses during cues signalling different outcomes as in Figure 6C, but with units ordered by their minimum firing rate. Overall, NAc units “tiled” experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

178 **NAc encoding of cue features persists until outcome:**

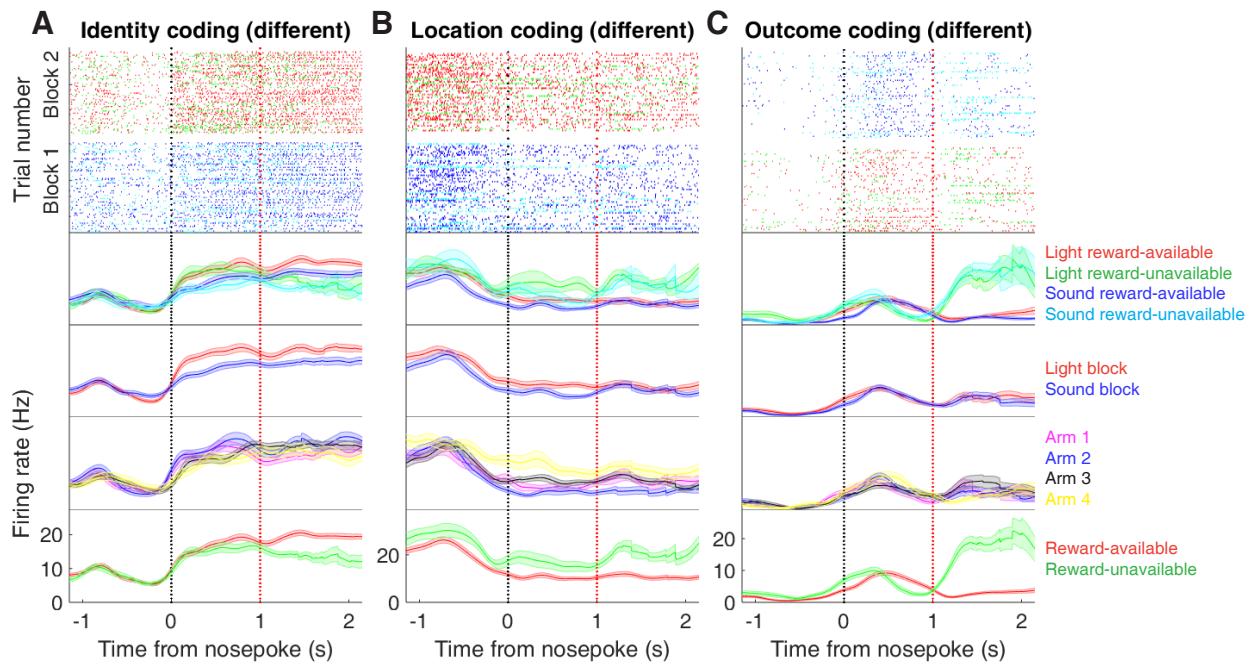
179 In order to be useful for credit assignment in reinforcement learning, a trace of the cue must be maintained  
180 until the outcome, so that information about the outcome can be associated with the outcome-predictive cue  
181 (Figure 1B). Investigation into the post-approach period during nosepoke revealed units that discriminated  
182 various cue features, with some units showing discriminative activity at both cue-onset and nosepoke (Fig-  
183 ure 7, supplement 1, supplement 2). To quantitatively test whether representations of cue features persisted  
184 post-approach until the outcome was revealed, we fit sliding window GLMs to the post-approach firing  
185 rates of cue-modulated units aligned to both the time of nosepoke into the reward receptacle, and after the  
186 outcome was revealed (Figure 8A,B, supplement 1 A-D, Table 1). This analysis showed that a variety of  
187 units discriminated firing according to cue identity (~20% of cue-modulated units), location (~25% of cue-  
188 modulated units), and outcome (~25% of cue-modulated units), but not other task parameters, showing that  
189 NAc activity discriminates various cue conditions well into a trial.



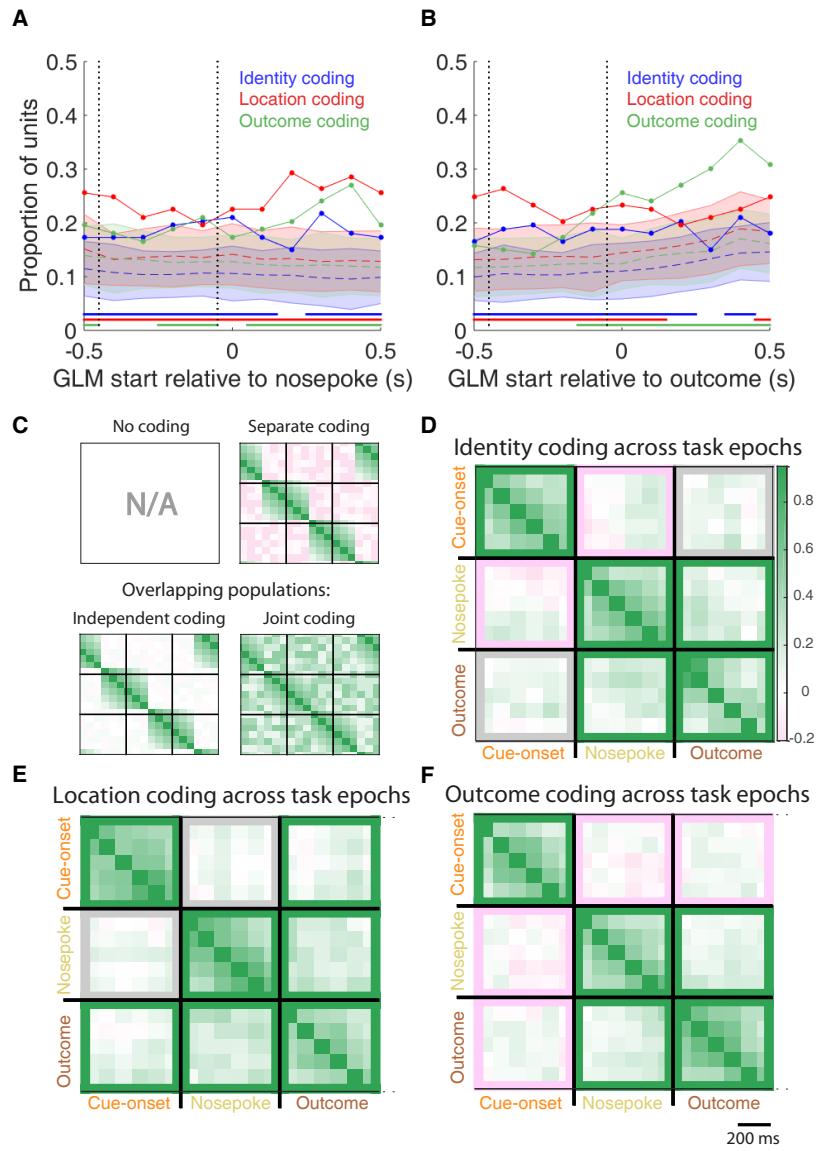
**Figure 7:** Examples of cue-modulated NAc units influenced by cue features at time of nosepoke. **A:** Example of a cue-modulated NAc unit that exhibited identity coding at both cue-onset and during subsequent nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to 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 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 7A,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).

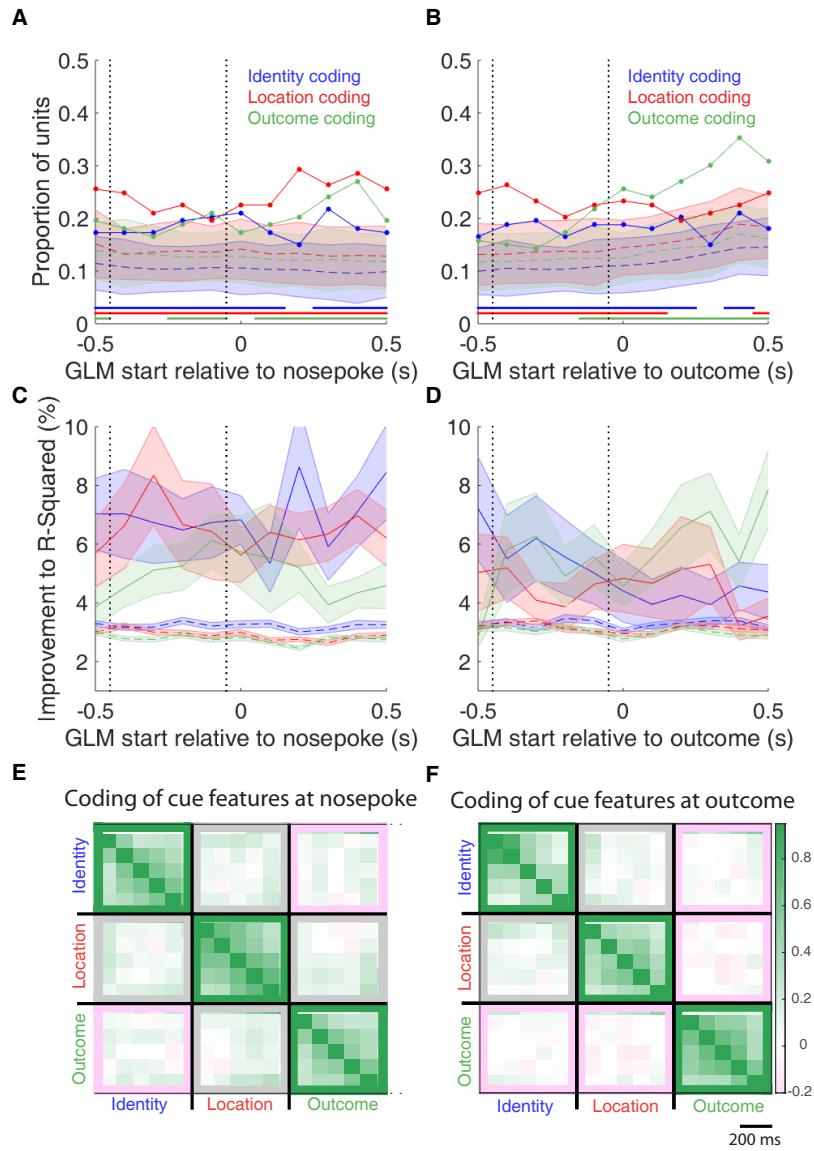


**Figure supplement 2:** Expanded examples of cue-modulated NAc units influenced by different cue features at time of nosepoke for Figure 7B,D,F, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).



**Figure 8:** 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 ( $z$ -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.

**Figure 8:** (Previous page.) 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 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. 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.



**Figure 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 ( $z$ -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 ( $r = 1$ ).

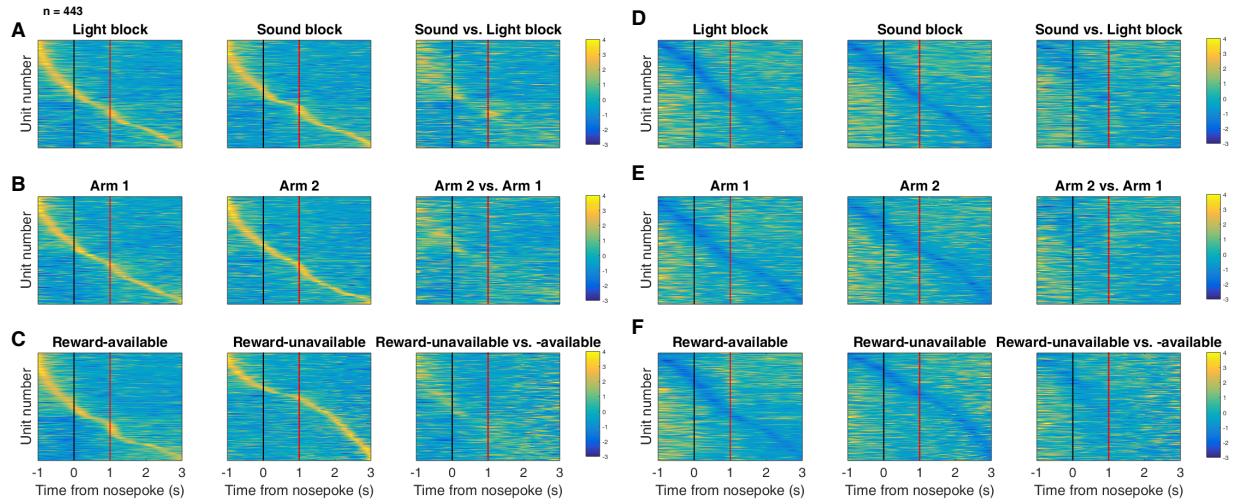
**Figure supplement 1:** (Previous page.) The window of GLMs used in each block is from the onset of the task phase to the 500 ms window post-onset, in 100 ms steps. Each individual value is for a sliding window GLM within that range, with the scale bar contextualizing step size. Colored square borders around each block indicate the result of a comparison of the mean correlation to a shuffled distribution, with pink indicating separate populations ( $z$ -score  $< -1.96$ ), grey indicating overlapping but independent populations, and green indicating joint overlapping populations ( $z$ -score  $> 1.96$ ). **F:** Same as E, but for time bins following outcome receipt. Color bar displays relationship between correlation value and color.

190 To determine whether NAc representations of cue features at nosepoke and outcome were encoded by a  
191 similar pool of units as during cue-onset, we correlated recoded beta coefficients from the GLMs for a cue  
192 feature across time points in the task, and compared the obtained correlations to the shuffled data generated  
193 for the GLMs (Figure 1B,C,8C-F). This revealed that identity coding at cue-onset was separate from coding  
194 at nosepoke (mean  $r = .048$ ; -2.18 standard deviations from the shuffled mean), independent from coding at  
195 outcome (mean  $r = .081$ ; 1.02 standard deviations from the shuffled mean), while joint coding was observed  
196 between nosepoke and outcome (mean  $r = .147$ ; 3.94 standard deviations from the shuffled mean). Applying  
197 this same analysis for cue location showed that location coding at cue-onset was independent from coding  
198 at nosepoke (mean  $r = .058$ ; -1.41 standard deviations from the shuffled mean), while joint coding was  
199 observed with coding at outcome (mean  $r = .093$ ; 2.03 standard deviations from the shuffled mean), as  
200 well as between nosepoke and outcome (mean  $r = .204$ ; 9.35 standard deviations from the shuffled mean).  
201 Applying this analysis to cue outcome revealed that outcome coding at cue-onset was separate from both  
202 coding at nosepoke (mean  $r = -.040$ ; -11.73 standard deviations from the shuffled mean), and coding at  
203 outcome (mean  $r = .025$ ; -5.06 standard deviations from the shuffled mean), while joint coding was observed  
204 between nosepoke and outcome (mean  $r = .148$ ; 3.15 standard deviations from the shuffled mean). Together,  
205 these findings show that the NAc maintains representations of cue features after a behavioral decision has  
206 been made with largely separate or independent pools of units, while a joint overlapping population is seen  
207 post-approach until the rat receives feedback about its decision.

208 To assess overlap among cue features at nosepoke and outcome receipt, we applied the same recoded coef-  
209 ficient analysis (Figure 8 supplement 1 E,F). This revealed that at the time of nosepoke, identity was coded  
210 independently from location (mean  $r = .124$ ; 1.00 standard deviations from the shuffled mean), and sepa-

211 rately from outcome (mean  $r = .052$ ; -3.19 standard deviations from the shuffled mean), and location and  
212 outcome were coded by an overlapping but independent population of units (mean  $r = .097$ ; -1.05 stan-  
213 dard deviations from the shuffled mean); while at outcome, identity was coded independently from location  
214 (mean  $r = .085$ ; -1.19 standard deviations from the shuffled mean), and separately from outcome (mean  $r =$   
215 .039; -4.20 standard deviations from the shuffled mean), and location and outcome were coded by a separate  
216 population of units (mean  $r = .004$ ; -7.22 standard deviations from the shuffled mean).

217 To assess the tiling of units around outcome receipt, we aligned normalized peak firing rates to nosepoke  
218 onset (Figure 8 supplement 2). This revealed a clustering of responses around outcome receipt for all cue  
219 conditions where the rat would have received reward, in addition to the same trend of higher within- vs  
220 across-block correlations for cue identity (Figure 8 supplement 2 A,C; within block correlations = .560  
221 (light), .541 (sound); across block correlations = .487, .481, .483, .486; within vs. within block comparison  
222 =  $p = .112$ ; within vs. across block comparisons =  $p < .001$ ) and cue location (Figure 8 supplement 2 B,E;  
223 within block correlations = .474 (arm 1), .461 (arm 2); across block correlations = .416, .402, .416, .415;  
224 within vs. within block comparison =  $p = .810$ ; within vs. across block comparisons =  $p < .001$ ), but not  
225 cue outcome (Figure 8 supplement 2 C,F; within block correlations = .620 (reward-available), .401 (reward-  
226 unavailable); across block correlations = .418, .414, .390, .408; within vs. within block comparison =  $p <$   
227 .001; within vs. across block comparisons =  $p < .001$ ), further reinforcing that the NAc segments the task  
228 and represents all aspects of task space.



**Figure 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 tiling around outcome receipt for reward-available, but not reward-unavailable trials suggests encoding of reward receipt by NAc units. **D-F:** Heat plots aligned to normalized minimum firing rates. **D:** Responses during different stimulus blocks as in A, but with units ordered according to the time of their minimum firing rate. **E:** Responses during trials on different arms as in B, but with units ordered by their minimum firing rate. **F:** Responses during cues signalling different outcomes as in C, but with units ordered by their minimum firing rate. Overall, NAc units “tiled” experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

229 **Discussion**

230 The main result of the present study is that NAc units encode not only the expected outcome of outcome-  
231 predictive cues, but also the identity of such cues (Figure 1A). The population of units that coded for cue  
232 identity was statistically separate from the population coding for expected outcome at each major time point  
233 within a trial (i.e. overlap less than expected from chance), and independent from the population coding for  
234 the location of the cue (i.e. overlap as expected from chance, Figure 1C). Importantly, this identity coding  
235 was maintained on approach trials by a distinct population of units both during a delay period where the rat  
236 held a nosepoke until the outcome was received, and immediately after outcome receipt (Figure 1B,C). This  
237 information was also present at the population level, with a higher than chance classification accuracy for  
238 predicting the right cue condition using a pseudoensemble. Units that coded different cue features (identity,  
239 outcome, location) exhibited different temporal profiles as a whole, although across all recorded units a tiling  
240 of task structure was observed such that all points within our analyzed task space was accounted for by the  
241 ordered peak firing rates of all units. Furthermore, this tiling differed between various conditions with a cue  
242 feature, such as light versus sound blocks. We discuss these observations and their implications below.

243 **Identity coding:**

244 Our finding that NAc units can discriminate between different outcome-predictive stimuli with similar moti-  
245 vational significance (i.e. encodes cue identity) expands upon an extensive rodent literature examining NAc  
246 correlates of conditioned stimuli (Ambroggi, Ishikawa, Fields, & Nicola, 2008; Atallah et al., 2014; Bis-  
247 sonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Dejean et al., 2017; Goldstein et al., 2012; Ishikawa,  
248 Ambroggi, Nicola, & Fields, 2008; Lansink et al., 2012; McGinty et al., 2013; Nicola, 2004; Roesch et al.,  
249 2009; Roitman et al., 2005; Saddoris et al., 2011; Setlow et al., 2003; Sugam et al., 2014; West & Carelli,  
250 2016; Yun, Wakabayashi, Fields, & Nicola, 2004). Perhaps the most comparable work in rodents comes from  
251 a study that found a subset of NAc units that modulated their firing for an odor when it predicted separate but  
252 equally valued rewards (Cooch et al., 2015). The present work is complementary to such *outcome identity*

253 coding as it shows that NAc units encode *cue identity*, in addition to the reward it predicts (Figure 1A). Sim-  
254 ilarly, Setlow et al. (2003) paired distinct odor cues with appetitive or aversive outcomes, and found separate  
255 populations of units that encoded each cue. Furthermore, during a reversal they found that the majority of  
256 units switched their selectivity, arguing that the NAc units were tracking the motivational significance of  
257 these stimuli. Once again, our study was different in asking how distinct cues encoding the same anticipated  
258 outcome are encoded, suggesting that the NAc dissociates their representations at multiple levels of analysis  
259 (e.g. single-unit and population) even when the motivational significance of these stimuli is identical. A pos-  
260 sible interpretation of this coding of cue identity alongside expected outcome is that these representations are  
261 used to associate reward with relevant features of the environment, so-called credit assignment in the rein-  
262 forcement learning literature (Sutton & Barto, 1998). A burgeoning body of human and non-human primate  
263 work has started to elucidated neural correlates of credit assignment in the PFC (Akaiishi, Kolling, Brown, &  
264 Rushworth, 2016; Asaad, Lauro, Perge, & Eskandar, 2017; Chau et al., 2015; Noonan, Chau, Rushworth, &  
265 Fellows, 2017), and given the importance of cortical inputs in NAc associative representations it is possible  
266 that information related to credit assignment is relayed from the cortex to NAc (Cooch et al., 2015; Ishikawa  
267 et al., 2008).

268 Viewed within the neuroeconomic framework of decision making, functional magnetic resonance imaging  
269 (fMRI) studies have found support for NAc representations of *offer value*, a domain-general common cur-  
270 rency signal that enables comparison of different attributes such as reward identity, effort, and temporal  
271 proximity (Bartra et al., 2013; Levy & Glimcher, 2012; Peters & Büchel, 2009; Sescousse et al., 2015). Our  
272 study adds to a growing body of electrophysiological research that suggests the view of the NAc as a value  
273 center, while informative and capturing major aspects of NAc processing, neglects additional contributions  
274 of NAc to learning and decision making such as the offer (cue) identity signal reported here.

275 A different possible function for cue identity coding is to support contextual modulation of the motivational  
276 relevance of specific cues. A context can be understood as a particular mapping between specific cues and  
277 their outcomes: for instance, in context 1 cue A but not cue B is rewarded, whereas in context 2 cue B but not

cue A is rewarded. Successfully implementing such contextual mappings requires representation of the cue identities. Indeed, Sleszer et al. (2016) recorded NAc responses during the Wisconsin Card Sorting Task, a common set-shifting task used in both the laboratory and clinic, and found units that preferred firing to stimuli when a certain rule, or rule category was currently active. Further support for a modulation of NAc responses by strategy comes from an fMRI study that examined blood-oxygen-level dependent (BOLD) levels during a set-shifting task (FitzGerald et al., 2014). In this task, participants learned two sets of stimulus-outcome contingencies, a visual set and an auditory set. During testing they were presented with both simultaneously, and the stimulus dimension that was relevant was periodically shifted between the two. Here, they found that bilateral NAc activity reflected value representations for the currently relevant stimulus dimension, and not the irrelevant stimulus. The current report of separate identity and outcome coding suggests the possibility that the relevance-gated value representations are generated by the combined activity of several different cell populations.

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 that could be attributed to this confound. However, the possibility remains that instead of, or in addition to, stimulus identity, these units encode a preferred context, or even a macroscale representation of progress through the session. Indeed, encoding of the current strategy could be an explanation for the presence of pre-cue identity coding (Figure 5A), as well as for the differential tiling of task structure across blocks observed in the current study (Figure 6).

A different potential confound is that between outcome and action value coding as the rat was only rewarded for left turns. Our GLM analysis dealt with this by excluding firing rate variance accounted for by a predictor that represented whether the animal approached (left turn) or skipped (right turn) the reward port at the

choice point, thus we were able to identity units that were modulated by the expected outcome of the cue after removing variance for potential action value coding. Another possible caveat is that NAc signals have been shown to be modulated by response vigor (McGinty et al., 2013); to detangle this from our results we included trial length (i.e. latency to arrival at the reward site) as a predictor in our GLMs, and found units with cue feature correlates independent of trial length.

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

### 319 Encoding of position:

Our finding that cue-modulated activity was influenced by cue location supports several previous reports (Lavoie & Mizumori, 1994; Mulder, Shibata, Trullier, & Wiener, 2005; Strait et al., 2016; Wiener et al., 2003). The NAc receives inputs from the hippocampus, and the communication of place-reward information across the two structures suggests that the NAc tracks locations associated with reward (Lansink et al., 2008; Lansink, Goltstein, Lankelma, McNaughton, & Pennartz, 2009; Lansink et al., 2016; Pennartz, 2004; Sjulson, Peyrache, Cumpelik, Cassataro, & Buzsáki, 2017; Tabuchi, Mulder, & Wiener, 2000; van der Meer & Redish, 2011). However, it is notable that in our task, location is explicitly uninformative about reward, yet

327 coding of this uninformative variable persists. The finding of tiling of task space is in alignment with previous  
328 studies showing that NAc units can also signal progress through a sequence of cues and/or actions (Atallah  
329 et al., 2014; Berke, Breck, & Eichenbaum, 2009; Khamassi, Mulder, Tabuchi, Douchamps, & Wiener, 2008;  
330 Lansink et al., 2012; Mulder, Tabuchi, & Wiener, 2004; Shidara, Aigner, & Richmond, 1998), is similar to  
331 observations in the basal forebrain (Tingley & Buzsáki, 2018), and may represent a temporally evolving state  
332 value signal (Hamid et al., 2015; Pennartz, Ito, Verschure, Battaglia, & Robbins, 2011). Given that the current  
333 task was pseudo-random, it is possible that the rats learned the structure of sequential cue presentation, and  
334 the neural activity could reflect this. However, this is unlikely as including a trial history variable in the  
335 GLM analysis did not explain a significant amount of firing rate variance for the vast majority of units. In  
336 any case, NAc units on the present task continued to distinguish between different locations, even though  
337 location, and progress through a sequence, were explicitly irrelevant in predicting reward. We speculate that  
338 this persistent coding of location in NAc may represent a bias in credit assignment, and associated tendency  
339 for rodents to associate motivationally relevant events with the locations where they occur.

340 **Implications:**

341 Maladaptive decision making, as occurs in schizophrenia, addiction, Parkinson's, among others, can re-  
342 sult from dysfunctional RPE and value signals (Frank, Seeberger, & O'Reilly, 2004; Gradin et al., 2011;  
343 Maia & Frank, 2011). This view has been successful in explaining both positive and negative symptoms  
344 in schizophrenia, and deficits in learning from feedback in Parkinson's (Frank et al., 2004; Gradin et al.,  
345 2011). However, the effects of RPE and value updating are contingent upon encoding of preceding action  
346 and cue features, the eligibility trace (Lee et al., 2012; Sutton & Barto, 1998). Value updates can only be  
347 performed on these aspects of preceding experience that are encoded when the update occurs. Therefore,  
348 maladaptive learning and decision making can result from not only aberrant RPEs but also from altered cue  
349 feature encoding. For instance, on this task the environmental stimulus that signaled the availability of re-  
350 ward was conveyed by two distinct cues that were presented in four locations. While in our current study,  
351 the location and identity of the cue did not require any adjustments in the animals behavior, we found coding

352 of these features alongside the expected outcome of the cue that could be the outcome of credit assignment  
353 computations computed upstream (Akaishi et al., 2016; Asaad et al., 2017; Chau et al., 2015; Noonan et al.,  
354 2017). Identifying neural coding related to an aspect of credit assignment is important as inappropriate credit  
355 assignment could be a contributor to conditioned fear overgeneralization seen in disorders with pathologi-  
356 cal anxiety such as generalized anxiety disorder, post traumatic stress disorder, and obsessive-compulsive  
357 disorder (Kaczkurkin et al., 2017; Kaczkurkin & Lissek, 2013; Lissek et al., 2014), and delusions observed  
358 in disorders such as schizophrenia, Alzheimer's and Parkinson's (Corlett, Taylor, Wang, Fletcher, & Krys-  
359 tal, 2010; Kapur, 2003). Thus, our results provide a neural window into the process of credit assignment,  
360 such that the extent and specific manner in which this process fails in syndromes such as schizophrenia,  
361 obsessive-compulsive disorder, etc. can be experimentally accessed.

## 362 **Methods**

### 363 **Subjects:**

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

### 371 **Overall timeline:**

372 Each rat was first handled for seven days during which they were exposed to the experiment room, the  
373 sucrose solution used as a reinforcer, and the click of the sucrose dispenser valves. Rats were then trained  
374 on the behavioral task (described in the next section) until they reached performance criterion. At this point  
375 they underwent hyperdrive implantation targeted at the NAc. Rats were allowed to recover for a minimum  
376 of five days before being retrained on the task, and recording began once performance returned to pre-  
377 surgery levels. Upon completion of recording, animals were glosed, euthanized and recording sites were  
378 histologically confirmed.

379 **Behavioral task and training:**

380 The behavioral apparatus was an elevated, square-shaped track (100 x 100 cm, track width 10 cm) containing  
381 four possible reward locations at the end of track “arms” (Figure 2). Rats initiated a *trial* by triggering a  
382 photobeam located 24 cm from the start of each arm. Upon trial initiation, one of two possible light cues  
383 (L1, L2), or one of two possible sound cues (S1, S2), was presented that signaled the presence (*reward-*  
384 *available trial*, L1+, S1+) or absence (*reward-unavailable trial*, L2-, S2-) of a 12% sucrose water reward  
385 (0.1 mL) at the upcoming reward site. A trial was classified as an *approach trial* if the rat turned left at the  
386 decision point and made a nosepoke at the reward receptacle (40 cm from the decision point), while trials  
387 were classified as a *skip trial* if the rat instead turned right at the decision point and triggered the photobeam  
388 to initiate the next trial. A trial was labeled *correct* if the rat approached (i.e. nosepoked) on reward-available  
389 trials, and skipped (i.e. did not nosepoke) on reward-unavailable trials. On reward-available trials there  
390 was a 1 second delay between a nosepoke and subsequent reward delivery. *Trial length* was determined by  
391 measuring the length of time from cue-onset until nosepoke (for approach trials), or from cue-onset until  
392 the start of the following trial (for skip trials). Trials could only be initiated through clockwise progression  
393 through the series of arms, and each entry into the subsequent arm on the track counted as a trial. Cues were  
394 present until 1 second after outcome receipt on approach trials, and until initiating the following trial on skip  
395 trials.

396 Each session consisted of both a *light block* and a *sound block* with 100 trials each. Within a block, one cue  
397 signaled reward was available on that trial (L1+ or S1+), while the other signaled reward was not available  
398 (L2- or S2-). Light block cues were a flashing white light, and a constant yellow light. Sound block cues  
399 were a 2 kHz sine wave and a 8 kHz sine wave whose amplitude was modulated from 0 to maximum by  
400 a 2 Hz sine wave. Outcome-cue associations were counterbalanced across rats, e.g. for some rats L1+ was  
401 the flashing white light, and for others L1+ was the constant yellow light. The order of cue presentation  
402 was pseudorandomized so that the same cue could not be presented more than twice in a row. Block order  
403 within each day was also pseudorandomized, such that the rat could not begin a session with the same block  
404 for more than two days in a row. Each session consisted of a 5 minute pre-session period on a pedestal (a  
405 terracotta planter filled with towels), followed by the first block, then the second block, then a 5 minute post-  
406 session period on the pedestal. For approximately the first week of training, rats were restricted to running  
407 in the clockwise direction by presenting a physical barrier to running counterclockwise. Cues signaling the  
408 availability and unavailability of reward, as described above, were present from the start of training. Rats  
409 were trained for 200 trials per day (100 trials per block) until they discriminated between the reward-available  
410 and reward-unavailable cues for both light and sound blocks for three consecutive days, according to a chi-  
411 square test rejecting the null hypothesis of equal approaches for reward-available and reward-unavailable  
412 trials, at which point they underwent electrode implant surgery.

413 **Surgery:**

414 Surgical procedures were as described previously (Malhotra, Cross, Zhang, & Van Der Meer, 2015). Briefly,  
415 animals were administered analgesics and antibiotics, anesthetized with isoflurane, induced with 5% in med-  
416 ical grade oxygen and maintained at 2% throughout the surgery (~0.8 L/min). Rats were then chronically  
417 implanted with a “hyperdrive” consisting of 20 independently drivable tetrodes, with 4 designated as refer-  
418 ences tetrodes, and the remaining 16 either all targeted for the right NAc (AP +1.4 mm and ML +1.6 mm  
419 relative to bregma; Paxinos & Watson 1998), or 12 in the right NAc and 4 targeted at the mPFC (AP +3.0  
420 mm and ML +0.6 mm, relative to bregma; only data from NAc tetrodes was analyzed). Following surgery,

421 all animals were given at least five days to recover while receiving post-operative care, and tetrodes were  
422 lowered to the target (DV -6.0 mm) before being reintroduced to the behavioral task.

423 **Data acquisition and preprocessing:**

424 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with  
425 a RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex).  
426 Signals were referenced against a tetrode placed in the corpus callosum above the NAc. Candidate spikes  
427 for sorting into putative single units were obtained by band-pass filtering the data between 600-9000 Hz,  
428 thresholding and aligning the peaks (UltraMegaSort2k, Hill, Mehta, & Kleinfeld, 2011). Spike waveforms  
429 were then clustered with KlustaKwik using energy and the first derivative of energy as features, and manually  
430 sorted into units (MClust 3.5, A.D. Redish et al., <http://redishlab.neuroscience.umn.edu/MClust/MClust.html>).  
431 Isolated units containing a minimum of 200 spikes within a session were included for subsequent analysis.  
432 Units were classified as FSIs by an absence of interspike intervals (ISIs)  $> 2$  s, while MSNs had a combi-  
433 nation of ISIs  $> 2$  s and phasic activity with shorter ISIs (Atallah et al., 2014; Barnes, Kubota, Hu, Jin, &  
434 Graybiel, 2005).

435 **Data analysis:**

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

443 *Neural data.* Given that some of our analyses compare firing rates across time, particularly comparisons  
444 across blocks, we sought to exclude units with unstable firing rates that would generate spurious results  
445 reflecting a drift in firing rate over time unrelated to our task. To do this we ran a Mann-Whitney U test  
446 comparing the cue-modulated firing rates for the first and second half of trials within a block, and excluded  
447 99 of 443 units from analysis that showed a significant change for either block, leaving 344 units for further  
448 analyses by our GLM. To investigate the contribution of different cue features (*cue identity*, *cue location*  
449 and *cue outcome*) on the firing rates of NAc single units, we first determined whether firing rates for a unit  
450 were modulated by the onset of a cue by collapsing across all cues and comparing the firing rates for the 1  
451 s preceding cue-onset with the 1 s following cue-onset. Single units were considered to be *cue-modulated* if  
452 a Wilcoxon signed-rank test comparing pre- and post-cue firing was significant at  $p < .01$ . Cue-modulated  
453 units were then classified as either increasing or decreasing if the post-cue activity was higher or lower than  
454 the pre-cue activity, respectively.

455 To determine the relative contribution of different task parameters to firing rate variance (as in Figure  
456 5A, supplement 1), a forward selection stepwise GLM using a Poisson distribution for the response vari-  
457 able was fit to each cue-modulated unit, using data from every trial in a session. Cue identity (light block,  
458 sound block), cue location (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available, reward-unavailable),  
459 behavior (approach, skip), trial length, trial number, and trial history (reward availability on the previous 2  
460 trials) were used as predictors, with firing rate as the response variable. The GLMs were fit using a sliding  
461 window for firing rate; with a bin size of 500 ms, step size of 100 ms, and an analysis window from 500  
462 ms pre-cue-onset to 500 ms post-cue-onset, such that 11 different GLMs were fit for each unit, tracking the  
463 temporal dynamics of the influence of task parameters on firing rate around the onset of the cue. Units were  
464 classified as being modulated by a given task parameter if addition of the parameter significantly improved  
465 model fit using deviance as the criterion ( $p < .01$ ), and the total proportion of cue-modulated units influ-  
466 enced by a task parameter was counted for each time bin. A comparison of the R-squared value between the  
467 final model and the final model minus the predictor of interest was used to determine the amount of firing  
468 rate variance explained by the addition of that predictor for a given unit. To control for the amount of units

469 that would be affected by a predictor by chance, we shuffled the trial order of firing rates for a particular  
470 unit within a time bin, ran the GLM with the shuffled firing rates, counted the proportion of units encoding  
471 a predictor, and took the average of this value over 100 shuffles. We then calculated how many z-scores the  
472 observed proportion was from the mean of the shuffled distribution, using a z-score of greater than 1.96 as a  
473 marker of significance.

474 To get a sense of the predictive power of these cue feature representations we trained a classifier using firing  
475 rates from a pseudoensemble comprised of our 133 cue-modulated units (Figure 5B). We created a matrix  
476 of firing rates for each time epoch surrounding cue-onset where each row was an observation representing  
477 the firing rate for a trial, and each column was a variable representing the firing rate for a given unit. Trial  
478 labels, or classes, were each condition for a cue feature (e.g. light and sound for cue identity), making sure  
479 to align trial labels across units. We then ran LDA on these matrices, using 10-fold cross validation to train  
480 the classifier on 90% of the trials and testing its predictions on the held out 10% of trials, and repeated this  
481 approach to get the classification accuracy for 100 iterations. To test if the classification accuracy was greater  
482 by chance, we shuffled the order of firing rates for each unit before we trained the classifier. We repeated this  
483 for 100 shuffled matrices for each time point, and calculated how many z-scores the mean classification rate  
484 of the observed data was from the mean of the shuffled distribution, using a z-score of 1.96 as the cut off for  
485 a significant deviation from the shuffled distribution.

486 To determine the degree to which coding of cue identity, cue location, and cue outcome overlapped within  
487 units we correlated the recoded beta coefficients from the GLMs for the cue features (Figure 5C,D). Specific-  
488 ally, we generated an array for each cue feature at each point in time where for all cue-modulated units we  
489 coded a ‘1’ if the cue feature was a significant predictor in the final model, and ‘0’ if it was not. We then  
490 correlated an array of the coded 0s and 1s for one cue feature with a similar array for another cue feature, re-  
491 peating this process for all post cue-onset sliding window combinations. The NAc was determined as coding  
492 a pair of cue features in a) separate populations of units if there was a significant negative correlation ( $r <$   
493 0), b) an independently coded overlapping population of units if there was no significant correlation ( $r = 0$ ),

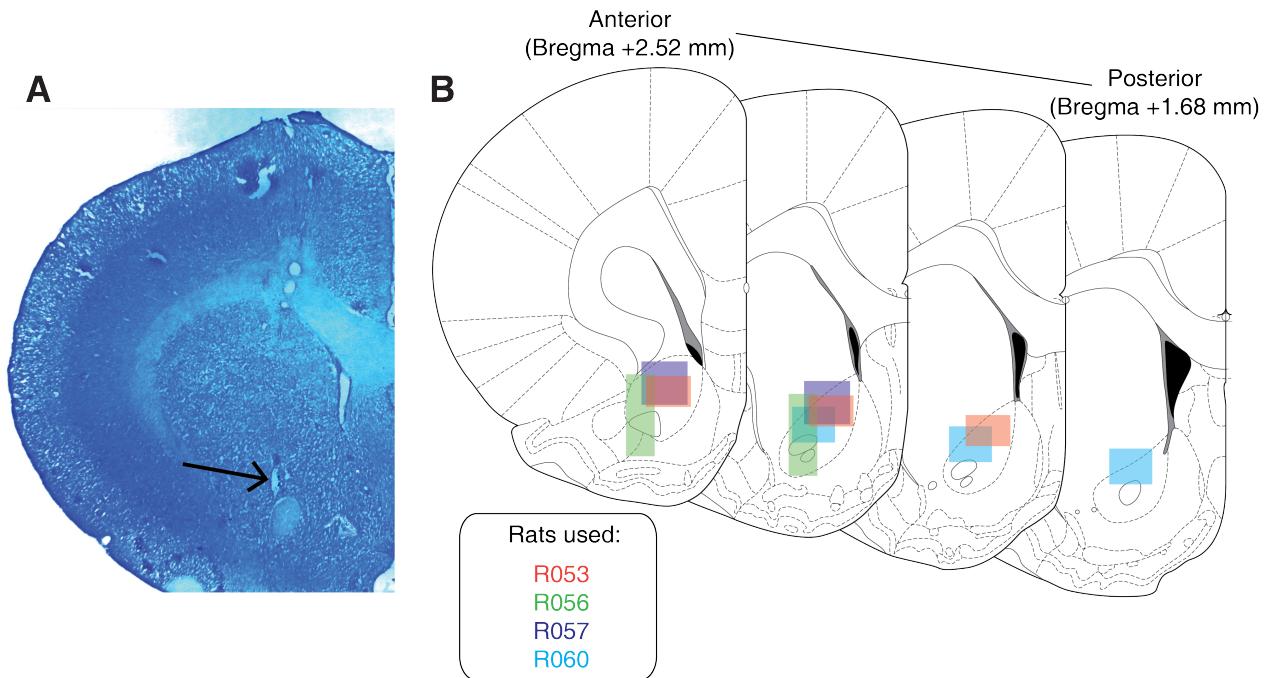
494 or c) a jointly coded overlapping population of units if there was a significant positive correlation ( $r > 0$ ).  
495 To summarize the correlation matrices generated from this analysis, we put the output of our 100 shuffled  
496 GLMs through the same pipeline, took the mean of the 36 correlations for a block comparison for each of the  
497 100 shuffles for an analysis window, and used the mean and standard deviation of these shuffled correlation  
498 averages to compare to the mean of the comparison block for the actual data, by generating a z-score, using a  
499 z-score of greater than 1.96 or less than -1.96 as a marker that that comparison block showed overlap across  
500 cue features greater or less than chance, respectively.

501 To better visualize responses to cues and enable subsequent population level analyses (as in Figures 4,6),  
502 spike trains were convolved with a Gaussian kernel ( $\sigma = 100$  ms), and peri-event time histograms (PETHs)  
503 were generated by taking the average of the convolved spike trains across all trials for a given task condition.  
504 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units  
505 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure  
506 6). To compare representations of task space across cue conditions for a cue feature, the ordering of units  
507 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other  
508 condition (e.g. sound block). To assess whether the tiling patterns were different across cue conditions, we  
509 split each cue condition into two halves, controlling for the effects of time by shuffling trial ordering before  
510 the split, and calculated the correlation of the temporally evolving smoothed firing rate across each of these  
511 halves, giving us 6 correlation values for each unit. We then concatenated these 6 values across all 443 units  
512 to give us an array of 2658 correlation coefficients. We then fit a linear mixed effects model, trying to predict  
513 these block comparison correlations with comparison type (e.g. 1st half of light block vs. 1st half of sound  
514 sound) as a fixed-effect term, and unit id as a random-effect term. Comparison type is nominal, so dummy  
515 variables were created for the various levels of comparison type, and coefficients were generated for each  
516 condition, referenced against one of the within-within comparison types (e.g. 1st half of light block vs. 2nd  
517 half of light block). Additionally, we ran a model comparison between the above model and a null model  
518 with just unit id, to see if adding comparison type improved model fit.

519 To identify the responsivity of units to different cue features at the time of nosepoke into a reward receptacle,  
520 and subsequent reward delivery, the same cue-modulated units from the cue-onset analyses were analyzed at  
521 the time of nosepoke and outcome receipt using identical analysis techniques for all approach trials (Figures  
522 7,8). To compare whether coding of a given cue feature was accomplished by the same or distinct population  
523 of units across time epochs, we ran the recoded coefficient correlation that was used to assess the degree of  
524 overlap among cue features within a time epoch. All analyses were completed in MATLAB R2015a, the  
525 code is available on our public GitHub repository (<http://github.com/vandermeerlab/papers>), and the data  
526 can be accessed through DataLad.

527 **Histology:**

528 Upon completion of the experiment, recording channels were glosed by passing  $10 \mu A$  current for 10 sec-  
529 onds and waiting 5 days before euthanasia, except for rat R057 whose implant detached prematurely. Rats  
530 were anesthetized with 5% isoflurane, then asphyxiated with carbon dioxide. Transcardial perfusions were  
531 performed, and brains were fixed and removed. Brains were sliced in  $50 \mu m$  coronal sections and stained  
532 with thionin. Slices were visualized under light microscopy, tetrode placement was determined, and elec-  
533 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.

534 **References**

- 535 Akaishi, R., Kolling, N., Brown, J. W., & Rushworth, M. (2016, jan). Neural Mechanisms of Credit Assignment in a Multicue  
536 Environment. *Journal of Neuroscience*, 36(4), 1096–1112. doi: 10.1523/JNEUROSCI.3159-15.2016
- 537 Ambroggi, F., Ishikawa, A., Fields, H. L., & Nicola, S. M. (2008, aug). Basolateral Amygdala Neurons Facilitate Reward-Seeking  
538 Behavior by Exciting Nucleus Accumbens Neurons. *Neuron*, 59(4), 648–661. doi: 10.1016/j.neuron.2008.07.004
- 539 Asaad, W. F., Lauro, P. M., Perge, J. A., & Eskandar, E. N. (2017, jul). Prefrontal Neurons Encode a Solution to the Credit-  
540 Assignment Problem. *The Journal of Neuroscience*, 37(29), 6995–7007. doi: 10.1523/JNEUROSCI.3311-16.2017
- 541 Atallah, H. E., McCool, A. D., Howe, M. W., & Graybiel, A. M. (2014, jun). Neurons in the ventral striatum exhibit cell-type-  
542 specific representations of outcome during learning. *Neuron*, 82(5), 1145–1156. doi: 10.1016/j.neuron.2014.04.021
- 543 Averbeck, B. B., & Costa, V. D. (2017, apr). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*,  
544 20(4), 505–512. doi: 10.1038/nrn.4506
- 545 Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005, oct). Activity of striatal neurons reflects dynamic encoding  
546 and recoding of procedural memories. *Nature*, 437(7062), 1158–1161. doi: 10.1038/nature04053
- 547 Bartra, O., McGuire, J. T., & Kable, J. W. (2013, aug). The valuation system: A coordinate-based meta-analysis  
548 of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, 76, 412–427. doi:  
549 10.1016/J.NEUROIMAGE.2013.02.063
- 550 Berke, J. D., Breck, J. T., & Eichenbaum, H. (2009, mar). Striatal Versus Hippocampal Representations During Win-Stay Maze  
551 Performance. *Journal of Neurophysiology*, 101(3), 1575–1587. doi: 10.1152/jn.91106.2008
- 552 Berridge, K. C. (2012, apr). From prediction error to incentive salience: Mesolimbic computation of reward motivation. *European  
553 Journal of Neuroscience*, 35(7), 1124–1143. doi: 10.1111/j.1460-9568.2012.07990.x
- 554 Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., ... Roesch, M. R. (2013, may). Sepa-  
555 rate Populations of Neurons in Ventral Striatum Encode Value and Motivation. *PLoS ONE*, 8(5), e64673. doi: 10.1371/jour-  
556 nal.pone.0064673
- 557 Bouton, M. E. (1993, jul). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological  
558 Bulletin*, 114(1), 80–99. doi: 10.1371/journal.pone.0064673
- 559 Carelli, R. M. (2010). Drug Addiction: Behavioral Neurophysiology. In *Encyclopedia of neuroscience* (pp. 677–682). Elsevier.  
560 doi: 10.1016/B978-008045046-9.01546-1
- 561 Chang, S. E., & Holland, P. C. (2013, nov). Effects of nucleus accumbens core and shell lesions on autoshaped lever-pressing.  
562 *Behavioural Brain Research*, 256, 36–42. doi: 10.1016/j.bbr.2013.07.046
- 563 Chang, S. E., Wheeler, D. S., & Holland, P. C. (2012, may). Roles of nucleus accumbens and basolateral amygdala in autoshaped  
564 lever pressing. *Neurobiology of Learning and Memory*, 97(4), 441–451. doi: 10.1016/j.nlm.2012.03.008

- 565 Chau, B. K. H., Sallet, J., Papageorgiou, G. K., Noonan, M. A. P., Bell, A. H., Walton, M. E., & Rushworth, M. F. S. (2015, sep).  
566 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5),  
567 1106–1118. doi: 10.1016/j.neuron.2015.08.018
- 568 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., & Wightman, R. M. (2007, apr). Coordinated  
569 Accumbal Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2), 237–244. doi:  
570 10.1016/j.neuron.2007.03.021
- 571 Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in  
572 schizophrenia. *Psychological Review*, 99(1), 45.doi: 10.1016/j.neuron.2007.03.021
- 573 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., & Schoenbaum, G. (2015,  
574 jun). Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature  
575 Communications*, 6, 7195. doi: 10.1038/ncomms8195
- 576 Corbit, L. H., & Balleine, B. W. (2011, aug). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are  
577 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33), 11786–11794. doi:  
578 10.1523/JNEUROSCI.2711-11.2011
- 579 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in  
580 Neurobiology*, 92, 345–369. doi: 10.1016/j.pneurobio.2010.06.007
- 581 Cromwell, H. C., & Schultz, W. (2003, may). Effects of Expectations for Different Reward Magnitudes on Neuronal Activity in  
582 Primate Striatum. *Journal of Neurophysiology*, 89(5), 2823–2838. doi: 10.1152/jn.01014.2002
- 583 Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine  
584 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8), 1020–1028. doi: 10.1038/nn1923
- 585 Day, J. J., Wheeler, R. A., Roitman, M. F., & Carelli, R. M. (2006, mar). Nucleus accumbens neurons encode Pavlovian ap-  
586 proach behaviors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5), 1341–1351. doi:  
587 10.1111/j.1460-9568.2006.04654.x
- 588 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., ... Le Moine, C. (2017, apr). Memories of Opiate  
589 Withdrawal Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychophar-  
590 macology*, 42(5), 1157–1168. doi: 10.1038/npp.2016.272
- 591 du Hoffmann, J., & Nicola, S. M. (2014, oct). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the  
592 Nucleus Accumbens. *Journal of Neuroscience*, 34(43), 14349–14364. doi: 10.1523/JNEUROSCI.3492-14.2014
- 593 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental  
594 Psychology*, 32(2), 150–155. doi: 10.1037/h0058316
- 595 Fitzgerald, T. H. B., Schwartenbeck, P., & Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent  
596 and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4), 1271–1279. doi: 10.1523/JNEUROSCI.4389-  
597 13.2014

- 598 Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., ... Akil, H. (2011, jan). A selective role for dopamine  
599 in stimulus-reward learning. *Nature*, 469(7328), 53–59. doi: 10.1038/nature09588
- 600 Floresco, S. B. (2015, jan). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of  
601 Psychology*, 66(1), 25–52. doi: 10.1146/annurev-psych-010213-115159
- 602 Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., & Magyar, O. (2006, mar). Dissociable roles for the nucleus accumbens core and  
603 shell in regulating set shifting. *Journal of Neuroscience*, 26(9), 2449–2457. doi: 10.1523/JNEUROSCI.4431-05.2006
- 604 Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004, dec). By carrot or by stick: Cognitive reinforcement learning in Parkinson-  
605 ism. *Science*, 306(5703), 1940–1943. doi: 10.1126/science.1102941
- 606 Giertler, C., Bohn, I., & Hauber, W. (2004, feb). Transient inactivation of the rat nucleus accumbens does not impair guidance of  
607 instrumental behaviour by stimuli predicting reward magnitude. *Behavioural Pharmacology*, 15(1), 55–63. doi: 10.1126/sci-  
608 ence.1102941
- 609 Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., ... Roesch, M. R. (2012, feb). Ventral  
610 Striatum Encodes Past and Predicted Value Independent of Motor Contingencies. *Journal of Neuroscience*, 32(6), 2027–  
611 2036. doi: 10.1523/JNEUROSCI.5349-11.2012
- 612 Goto, Y., & Grace, A. A. (2008, nov). Limbic and cortical information processing in the nucleus accumbens. *Trends in Neuro-  
613 sciences*, 31(11), 552–558. doi: 10.1016/j.tins.2008.08.002
- 614 Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., ... Steele, J. D. (2011, jun). Expected value and prediction  
615 error abnormalities in depression and schizophrenia. *Brain*, 134(6), 1751–1764. doi: 10.1093/brain/awr059
- 616 Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a  
617 Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. doi: 10.1037/h0059831
- 618 Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., ... Berke, J. D. (2015, jan).  
619 Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1), 117–126. doi: 10.1038/nn.4173
- 620 Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014, jan). Phasic Dopamine Release in the Rat Nucleus  
621 Accumbens Symmetrically Encodes a Reward Prediction Error Term. *The Journal of Neuroscience*, 34(3), 698–704. doi:  
622 10.1523/JNEUROSCI.2489-13.2014
- 623 Hassani, O. K., Cromwell, H. C., & Schultz, W. (2001, jun). Influence of Expectation of Different Rewards on Behavior-Related  
624 Neuronal Activity in the Striatum. *Journal of Neurophysiology*, 85(6), 2477–2489. doi: 10.1152/jn.2001.85.6.2477
- 625 Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: the stimulus-reinforcer relation and directed action*. Psychonomic Society. doi:  
626 10.1152/jn.2001.85.6.2477
- 627 Hill, D. N., Mehta, S. B., & Kleinfeld, D. (2011, jun). Quality Metrics to Accompany Spike Sorting of Extracellular Signals.  
628 *Journal of Neuroscience*, 31(24), 8699–8705. doi: 10.1523/JNEUROSCI.0971-11.2011
- 629 Holland, P. C. (1992, jan). Occasion setting in pavlovian conditioning. *Psychology of Learning and Motivation*, 28(C), 69–125.  
630 doi: 10.1016/S0079-7421(08)60488-0

- 631 Hollerman, J. R., Tremblay, L., & Schultz, W. (1998, aug). Influence of Reward Expectation on Behavior-Related Neuronal Activity  
632 in Primate Striatum. *Journal of Neurophysiology*, 80(2), 947–963. doi: 10.1152/jn.1998.80.2.947
- 633 Honey, R. C., Iordanova, M. D., & Good, M. (2014, feb). Associative structures in animal learning: Dissociating elemental and  
634 configural processes. *Neurobiology of Learning and Memory*, 108, 96–103. doi: 10.1016/j.nlm.2013.06.002
- 635 Hyman, S. E., Malenka, R. C., & Nestler, E. J. (2006, jul). NEURAL MECHANISMS OF ADDICTION: The Role  
636 of Reward-Related Learning and Memory. *Annual Review of Neuroscience*, 29(1), 565–598. doi: 10.1146/an-  
637 nurev.neuro.29.051605.113009
- 638 Ikemoto, S. (2007, nov). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-  
639 olfactory tubercle complex. *Brain Research Reviews*, 56(1), 27–78. doi: 10.1016/j.brainresrev.2007.05.004
- 640 Ishikawa, A., Ambroggi, F., Nicola, S. M., & Fields, H. L. (2008, may). Dorsomedial Prefrontal Cortex Contribution to Behav-  
641 ioral and Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, 28(19), 5088–5098. doi:  
642 10.1523/JNEUROSCI.0253-08.2008
- 643 Joel, D., Niv, Y., & Ruppin, E. (2002, jun). Actor-critic models of the basal ganglia: new anatomical and computational perspectives.  
644 *Neural Networks*, 15(4-6), 535–547. doi: 10.1016/S0893-6080(02)00047-3
- 645 Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., ... Lissek, S. (2017, feb).  
646 Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174(2), 125–134. doi:  
647 10.1176/appi.ajp.2016.15121549
- 648 Kaczkurkin, A. N., & Lissek, S. (2013). Generalization of Conditioned Fear and Obsessive-Compulsive Traits. *Journal of Psychol-  
649 ogy & Psychotherapy*, 7, 3. doi: 10.4172/2161-0487.S7-003
- 650 Kalivas, P. W., & Volkow, N. D. (2005, aug). The neural basis of addiction: A pathology of motivation and choice. *American  
651 Journal of Psychiatry*, 162(8), 1403–1413. doi: 10.1176/appi.ajp.162.8.1403
- 652 Kapur, S. (2003, jan). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology  
653 in schizophrenia. *American Journal of Psychiatry*, 160(1), 13–23. doi: 10.1176/appi.ajp.160.1.13
- 654 Khamassi, M., & Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and  
655 model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, 6, 79. doi: 10.3389/fnbeh.2012.00079
- 656 Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., & Wiener, S. I. (2008, nov). Anticipatory reward signals in ventral striatal  
657 neurons of behaving rats. *European Journal of Neuroscience*, 28(9), 1849–1866. doi: 10.1111/j.1460-9568.2008.06480.x
- 658 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., & Pennartz, C. M. A. (2008, jun).  
659 Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, 28(25),  
660 6372–6382. doi: 10.1523/JNEUROSCI.1054-08.2008
- 661 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., & Pennartz, C. M. (2009, aug). Hippocampus leads ventral  
662 striatum in replay of place-reward information. *PLoS Biology*, 7(8), e1000173. doi: 10.1371/journal.pbio.1000173
- 663 Lansink, C. S., Jackson, J. C., Lankelma, J. V., Ito, R., Robbins, T. W., Everitt, B. J., & Pennartz, C. M. A. (2012, sep). Reward

- 664 Cues in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles. *Journal*  
665 *of Neuroscience*, 32(36), 12444–12459. doi: 10.1523/JNEUROSCI.0593-12.2012
- 666 Lansink, C. S., Meijer, G. T., Lankelma, J. V., Vinck, M. A., Jackson, J. C., & Pennartz, C. M. A. (2016, oct). Reward Ex-  
667 pectancy Strengthens CA1 Theta and Beta Band Synchronization and Hippocampal-Ventral Striatal Coupling. *Journal of*  
668 *Neuroscience*, 36(41), 10598–10610. doi: 10.1523/JNEUROSCI.0682-16.2016
- 669 Lavoie, A. M., & Mizumori, S. J. (1994, feb). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons  
670 of rats. *Brain Research*, 638(1-2), 157–168. doi: 10.1016/0006-8993(94)90645-9
- 671 Lee, D., Seo, H., & Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of*  
672 *Neuroscience*, 35(1), 287–308. doi: 10.1146/annurev-neuro-062111-150512
- 673 Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*,  
674 22, 1027–1038. doi: 10.1016/j.conb.2012.06.001
- 675 Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014, jun). Generalized anxiety disor-  
676 der is associated with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75(11), 909–915. doi:  
677 10.1016/j.biopsych.2013.07.025
- 678 Maia, T. V. (2009, dec). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and*  
679 *Behavioral Neuroscience*, 9(4), 343–364. doi: 10.3758/CABN.9.4.343
- 680 Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature*  
681 *Neuroscience*, 14(2), 154–162. doi: 10.1038/nn.2723
- 682 Malhotra, S., Cross, R. W., Zhang, A., & Van Der Meer, M. A. A. (2015). Ventral striatal gamma oscillations are highly variable  
683 from trial to trial, and are dominated by behavioural state, and only weakly influenced by outcome value. *European Journal*  
684 *of Neuroscience*, 42(10), 2818–2832. doi: 10.1038/nn.2723
- 685 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011, feb). Ventral Striatum and Orbitofrontal Cortex  
686 Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, 31(7), 2700–  
687 2705. doi: 10.1523/JNEUROSCI.5499-10.2011
- 688 McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., & Nicola, S. M. (2013, jun). Invigoration of reward seeking by cue and  
689 proximity encoding in the nucleus accumbens. *Neuron*, 78(5), 910–922. doi: 10.1016/j.neuron.2013.04.010
- 690 Mulder, A. B., Shibata, R., Trullier, O., & Wiener, S. I. (2005, may). Spatially selective reward site responses in tonically active  
691 neurons of the nucleus accumbens in behaving rats. *Experimental Brain Research*, 163(1), 32–43. doi: 10.1007/s00221-  
692 004-2135-3
- 693 Mulder, A. B., Tabuchi, E., & Wiener, S. I. (2004, apr). Neurons in hippocampal afferent zones of rat striatum parse routes into  
694 multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7), 1923–1932. doi: 10.1111/j.1460-  
695 9568.2004.03301.x
- 696 Nicola, S. M. (2004, apr). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a

- 697 Discriminative Stimulus Task. *Journal of Neurophysiology*, 91(4), 1840–1865. doi: 10.1152/jn.00657.2003
- 698 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of  
699 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49), 16585–  
700 16600. doi: 10.1523/JNEUROSCI.3958-10.2010
- 701 Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007, mar). Tonic dopamine: Opportunity costs and the control of response vigor.  
702 *Psychopharmacology*, 191(3), 507–520. doi: 10.1007/s00213-006-0502-4
- 703 Noonan, M. P., Chau, B. K., Rushworth, M. F., & Fellows, L. K. (2017, jul). Contrasting Effects of Medial and Lateral Orbitofrontal  
704 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29), 7023–7035.  
705 doi: 10.1523/JNEUROSCI.0692-17.2017
- 706 Paxinos, G., & Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates* (4th ed.). San Diego: Academic Press. doi:  
707 10.1523/JNEUROSCI.0692-17.2017
- 708 Pennartz, C. M. A. (2004, jul). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation  
709 by Hippocampal Ripples. *Journal of Neuroscience*, 24(29), 6446–6456. doi: 10.1523/JNEUROSCI.0575-04.2004
- 710 Pennartz, C. M. A., Ito, R., Verschure, P., Battaglia, F., & Robbins, T. (2011, oct). The hippocampalstriatal axis in learning,  
711 prediction and goal-directed behavior. *Trends in Neurosciences*, 34(10), 548–559. doi: 10.1016/j.tins.2011.08.001
- 712 Peters, J., & Büchel, C. (2009, dec). Overlapping and distinct neural systems code for subjective value during intertemporal and  
713 risky decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(50), 15727–34.  
714 doi: 10.1523/JNEUROSCI.3489-09.2009
- 715 Rescorla, R. A., & Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and  
716 Instrumental Learning. *Psychological Review*, 74(3), 151–182. doi: 10.1037/h0024475
- 717 Robinson, T. E., & Flagel, S. B. (2009, may). Dissociating the Predictive and Incentive Motivational Properties of  
718 Reward-Related Cues Through the Study of Individual Differences. *Biological Psychiatry*, 65(10), 869–873. doi:  
719 10.1016/j.biopsych.2008.09.006
- 720 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009, oct). Ventral Striatal Neurons Encode the Value  
721 of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42),  
722 13365–13376. doi: 10.1523/JNEUROSCI.2572-09.2009
- 723 Roitman, M. F., Wheeler, R. A., & Carelli, R. M. (2005, feb). Nucleus accumbens neurons are innately tuned for reward-  
724 ing and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4), 587–597. doi:  
725 10.1016/j.neuron.2004.12.055
- 726 Saddoris, M. P., Stamatakis, A., & Carelli, R. M. (2011, jun). Neural correlates of Pavlovian-to-instrumental transfer in the nu-  
727 cleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*,  
728 33(12), 2274–2287. doi: 10.1111/j.1460-9568.2011.07683.x
- 729 Salamone, J. D., & Correa, M. (2012, nov). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3),

- 730 470–485. doi: 10.1016/j.neuron.2012.10.021
- 731 Schultz, W. (2016, feb). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32. doi:  
732 10.1038/nrn.2015.26
- 733 Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992, dec). Neuronal activity in monkey ventral striatum related to the  
734 expectation of reward. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 12(12), 4595–610.  
735 doi: 10.1523/JNEUROSCI.12-12-04595.1992
- 736 Schultz, W., Dayan, P., & Montague, P. R. (1997, mar). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.  
737 doi: 10.1126/science.275.5306.1593
- 738 Sescousse, G., Li, Y., & Dreher, J.-C. (2015, apr). A common currency for the computation of motivational values in the human  
739 striatum. *Social Cognitive and Affective Neuroscience*, 10(4), 467–473. doi: 10.1093/scan/nsu074
- 740 Setlow, B., Schoenbaum, G., & Gallagher, M. (2003, may). Neural encoding in ventral striatum during olfactory discrimination  
741 learning. *Neuron*, 38(4), 625–636. doi: 10.1016/S0896-6273(03)00264-2
- 742 Shidara, M., Aigner, T. G., & Richmond, B. J. (1998, apr). Neuronal signals in the monkey ventral striatum related to progress  
743 through a predictable series of trials. *Journal of Neuroscience*, 18(7), 2613–25. doi: 10.1016/S0896-6273(03)00264-2
- 744 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., & Buzsáki, G. (2017, may). Cocaine place conditioning strengthens  
745 location-specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, 1–10. doi: 10.1101/105890
- 746 Sleezer, B. J., Castagno, M. D., & Hayden, B. Y. (2016, nov). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.  
747 *Journal of Neuroscience*, 36(44), 11223–11237. doi: 10.1523/JNEUROSCI.1766-16.2016
- 748 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., & Hayden, B. Y. (2016, mar). Neuronal selectivity  
749 for spatial positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3), 1098–1111. doi:  
750 10.1152/jn.00325.2015
- 751 Sugam, J. A., Saddoris, M. P., & Carelli, R. M. (2014, may). Nucleus accumbens neurons track behavioral preferences and reward  
752 outcomes during risky decision making. *Biological Psychiatry*, 75(10), 807–816. doi: 10.1016/j.biopsych.2013.09.010
- 753 Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction* (Vol. 9) (No. 5). MIT Press, Cambridge, MA. doi:  
754 10.1109/TNN.1998.712192
- 755 Tabuchi, E. T., Mulder, A. B., & Wiener, S. I. (2000, jan). Position and behavioral modulation of synchronization of hip-  
756 pocampal and accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6), 717–728. doi: 10.1002/1098-  
757 1063(2000)10:6;717::AID-HIPO1009;3.0.CO;2-3
- 758 Takahashi, Y. K., Langdon, A. J., Niv, Y., & Schoenbaum, G. (2016, jul). Temporal Specificity of Reward Prediction Er-  
759 rrors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1), 182–193. doi:  
760 10.1016/j.neuron.2016.05.015
- 761 Tingley, D., & Buzsáki, G. (2018, jun). Transformation of a Spatial Map across the Hippocampal-Lateral Septal Circuit. *Neuron*,  
762 98(6), 1229–1242.e5. doi: 10.1016/J.NEURON.2018.04.028

- 763 van der Meer, M. A. A., & Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward  
764 Information. *Journal of Neuroscience*, 31(8), 2843–2854. doi: 10.1523/JNEUROSCI.4869-10.2011
- 765 West, E. A., & Carelli, R. M. (2016, jan). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after  
766 Reinforcer Devaluation. *Journal of Neuroscience*, 36(4), 1128–1139. doi: 10.1523/JNEUROSCI.2976-15.2016
- 767 Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., & Mulder, A. B. (2003, oct). Spatial and behavioral correlates  
768 in nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,  
769 1250(C), 275–292. doi: 10.1016/S0531-5131(03)00978-6
- 770 Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral  
771 and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, 24(12), 2923–2933. doi:  
772 10.1523/JNEUROSCI.5282-03.2004