

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

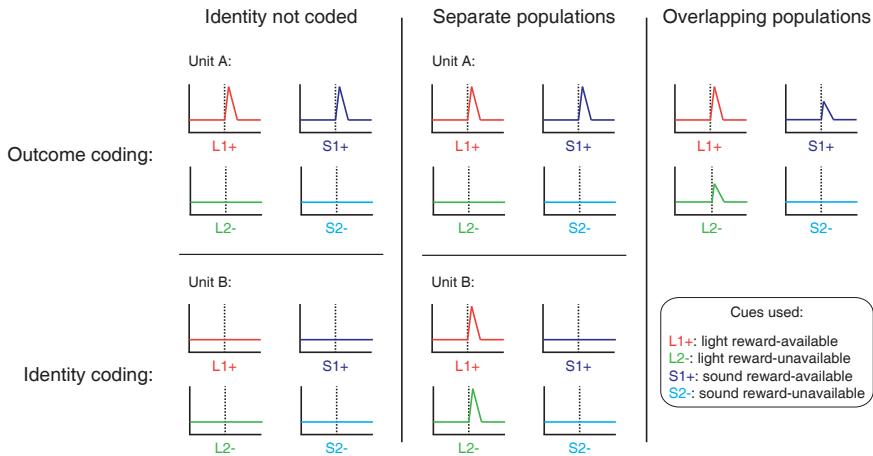
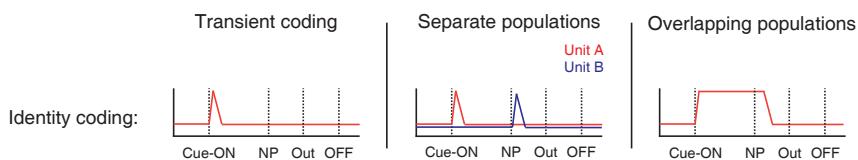
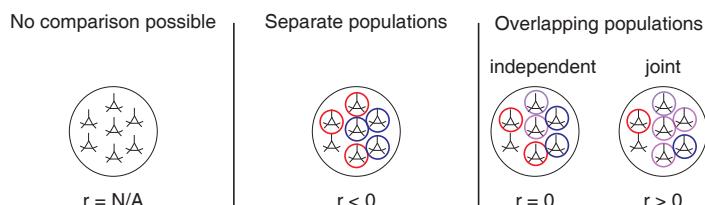
<sup>34</sup> Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues  
<sup>35</sup> (Atallah, McCool, Howe, & Graybiel, 2014; Bissonette et al., 2013; Cooch et al., 2015; Cromwell & Schultz,

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 issue  
42 for at least two reasons. First, in reinforcement learning, motivationally relevant outcomes are typically  
43 temporally delayed relative to the cues that predict them. In order to solve the problem of assigning credit  
44 (or blame) across such temporal gaps, some trace of preceding activity needs to be maintained (Lee et al.,  
45 2012; Sutton & Barto, 1998). For example, if you become ill after eating food X in restaurant A, depending  
46 on if you remember the identity of the restaurant or the food at the time of illness, you may learn to avoid  
47 all restaurants, restaurant A only, food X only, or the specific pairing of X-in-A. Therefore, a complete  
48 understanding of what is learned following feedback requires understanding what trace is maintained. Since  
49 NAc is a primary target of DA signals interpretable as RPEs, and NAc lesions impair RPEs related to timing,  
50 its activity trace will help determine what can be learned when RPEs arrive (Hamid et al., 2015; Hart,  
51 Rutledge, Glimcher, & Phillips, 2014; Ikemoto, 2007; McDannald, Lucantonio, Burke, Niv, & Schoenbaum,  
52 2011; Takahashi et al., 2016). Similarly, in a neuroeconomics framework, NAc is thought to represent a  
53 domain-general subjective value associated with different offers (Bartra, McGuire, & Kable, 2013; Levy &  
54 Glimcher, 2012; Peters & Büchel, 2009; Sescousse, Li, & Dreher, 2015); representation of the offer itself  
55 would provide a potential means 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 po-  
68 sition at the center of motivational systems. We sought to determine whether cue identity is represented in  
69 the NAc, if cue identity is represented alongside other motivationally relevant variables, such as cue value,  
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 encoding 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-predictive 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).

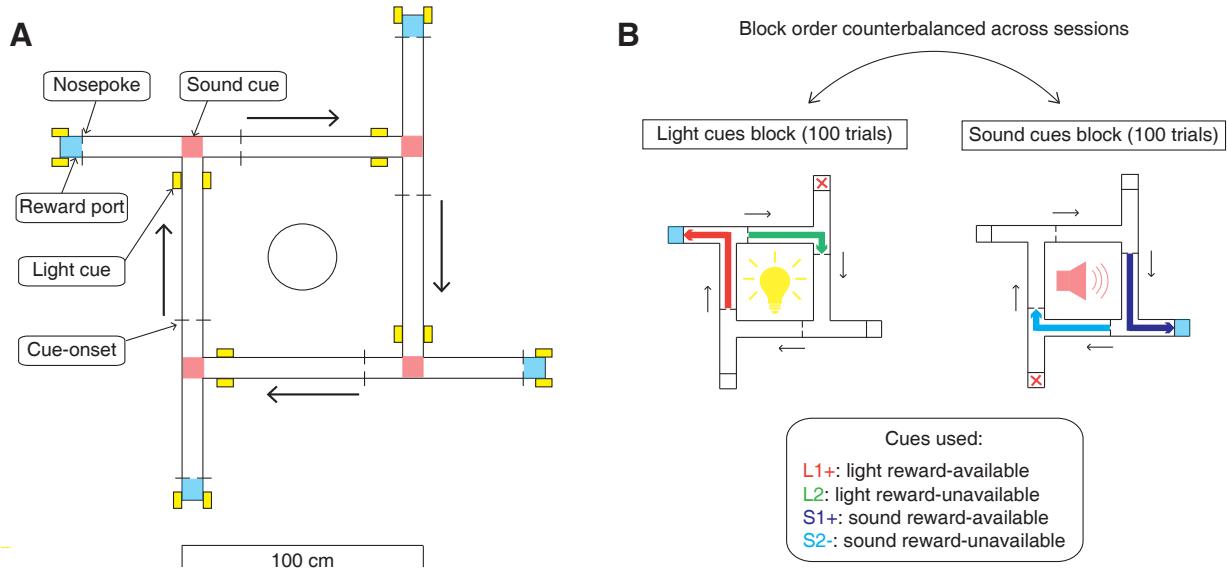
**Figure 1:** (Previous page.) **C:** Schematic pool of NAc units, illustrating the different possible outcomes for each of the analyses.  $R$  values represent the correlation between sets of recoded regression coefficients (see text for analysis details). Left panel: Cue identity is not coded (A: left 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 among 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 ( $r \sim 0$ ) suggests that the overlap of identity and outcome coding, or cue-onset and nosepoke coding, is expected by chance and that the two cue features, or points in time, are coded independently from one another. A positive correlation ( $r > 0$ ) implies a higher overlap than expected by chance, suggesting joint coding. 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

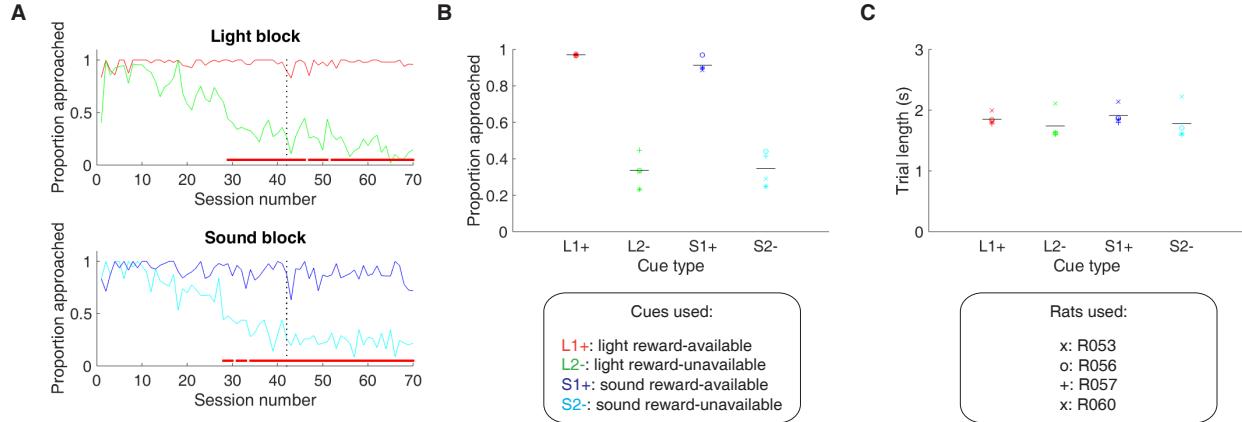
### 74 Behavior

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



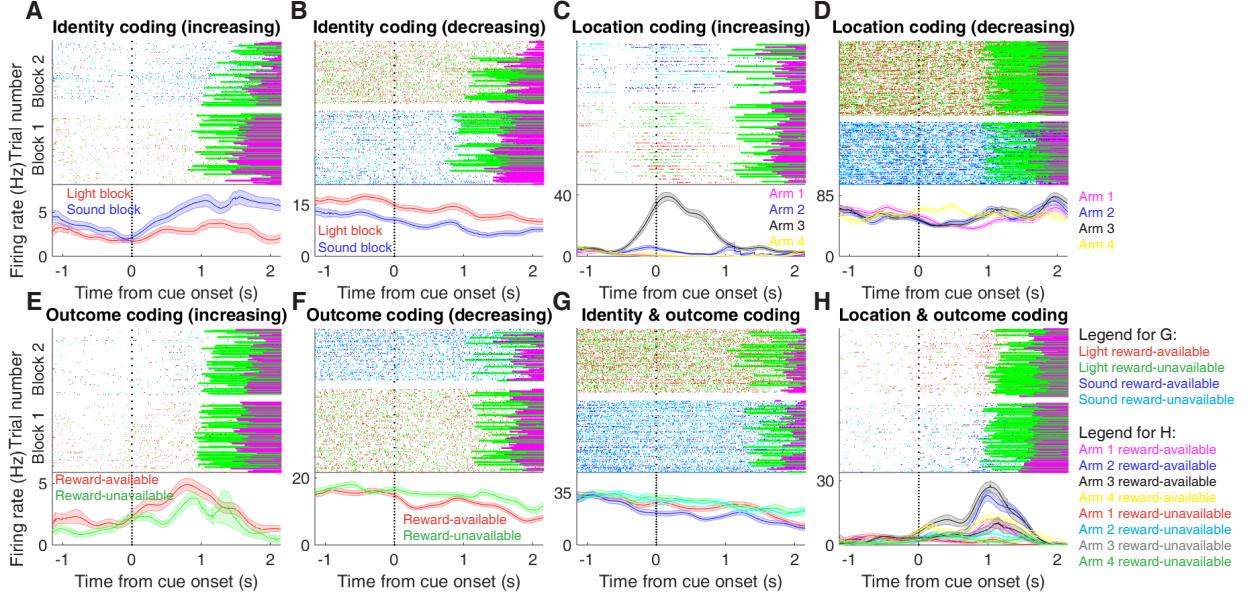
**Figure 3:** Performance on the behavioral task. **A.** Example learning curves across sessions from a single subject (R060) showing the proportion approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light block, light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

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

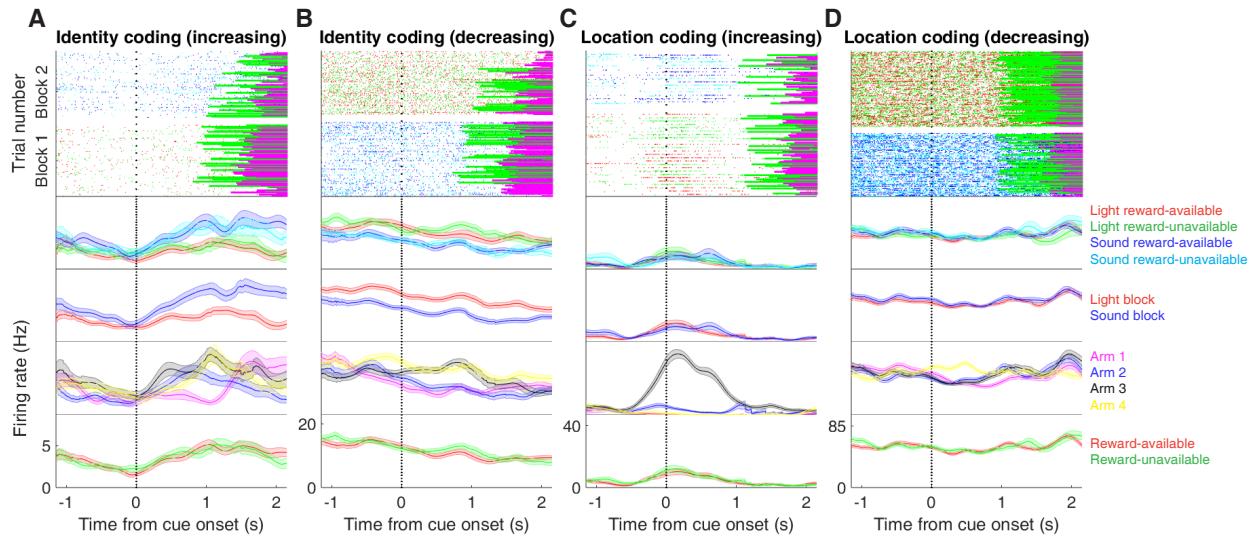
97 We sought to address which parameters of our task were encoded by NAc activity, specifically whether  
98 the NAc encodes aspects of motivationally relevant cues not directly tied to reward, such as the identity  
99 and location of the cue, and whether this coding is accomplished by separate or overlapping populations  
100 (Figure 1A). To do this we recorded a total of 443 units with > 200 spikes in the NAc from 4 rats over 57  
101 sessions (range: 12 - 18 sessions per rat) while they performed a cue discrimination task (Table 1). Units that  
102 exhibited a drift in firing rate over the course of either block, as measured by a Mann-Whitney U comparing  
103 firing 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, with  
105 more showing a decrease in firing ( $n = 103$ ) than an increase ( $n = 30$ ) around the time of cue-onset (Table 1).  
106 Within this group, 24 were classified as FSIs, while 109 were classified as SPNs. Upon visual inspection, we  
107 observed several patterns of firing activity, including units that discriminated firing upon cue-onset across  
108 various cue conditions, showed sustained differences in firing across cue conditions, had transient responses  
109 to the cue, showed a ramping of activity starting at cue-onset, and showed elevated activity immediately  
110 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%)
Previous trial	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%)

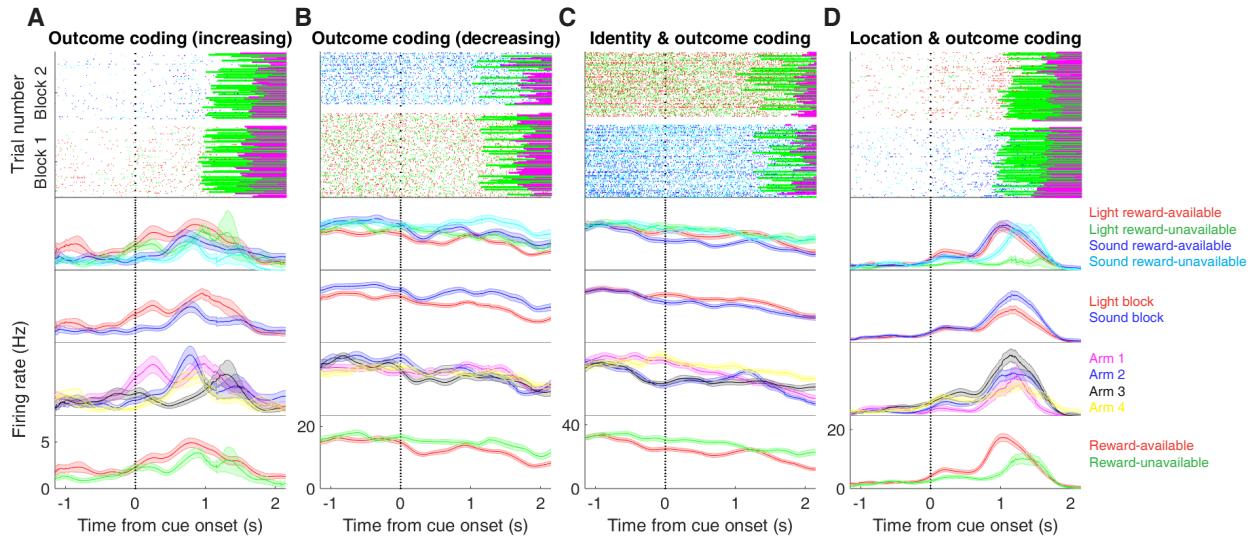
**Table 1:** Overview of recorded NAc units and their relationship to task variables at various time epochs. Percentage is relative to the number of cue-modulated units (n = 133).



**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 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.

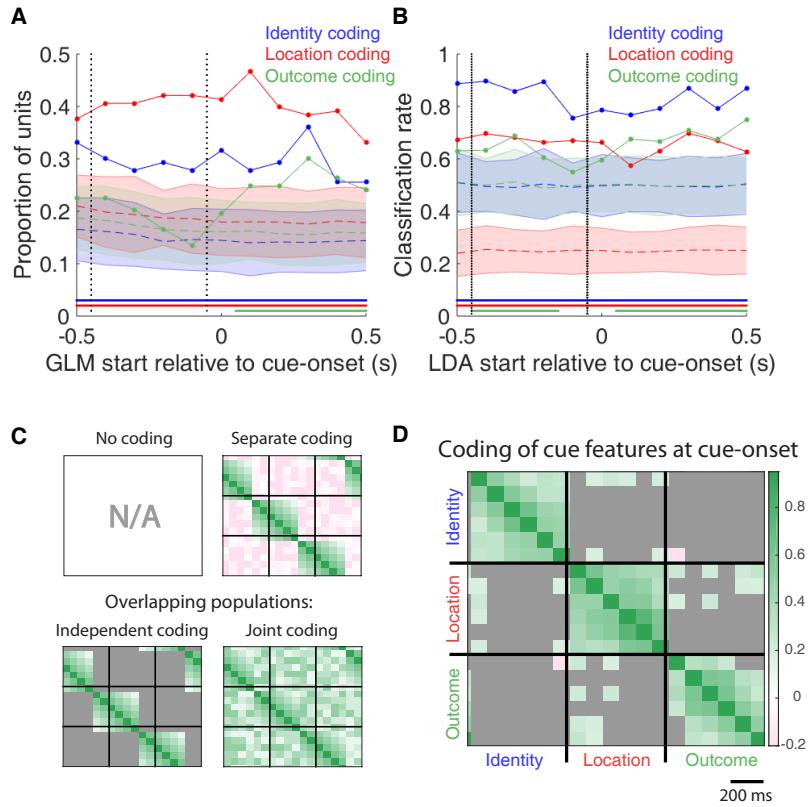


**Figure supplement 1** Expanded examples of cue-modulated NAc units influenced by different task parameters for Figure 4A-D, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).

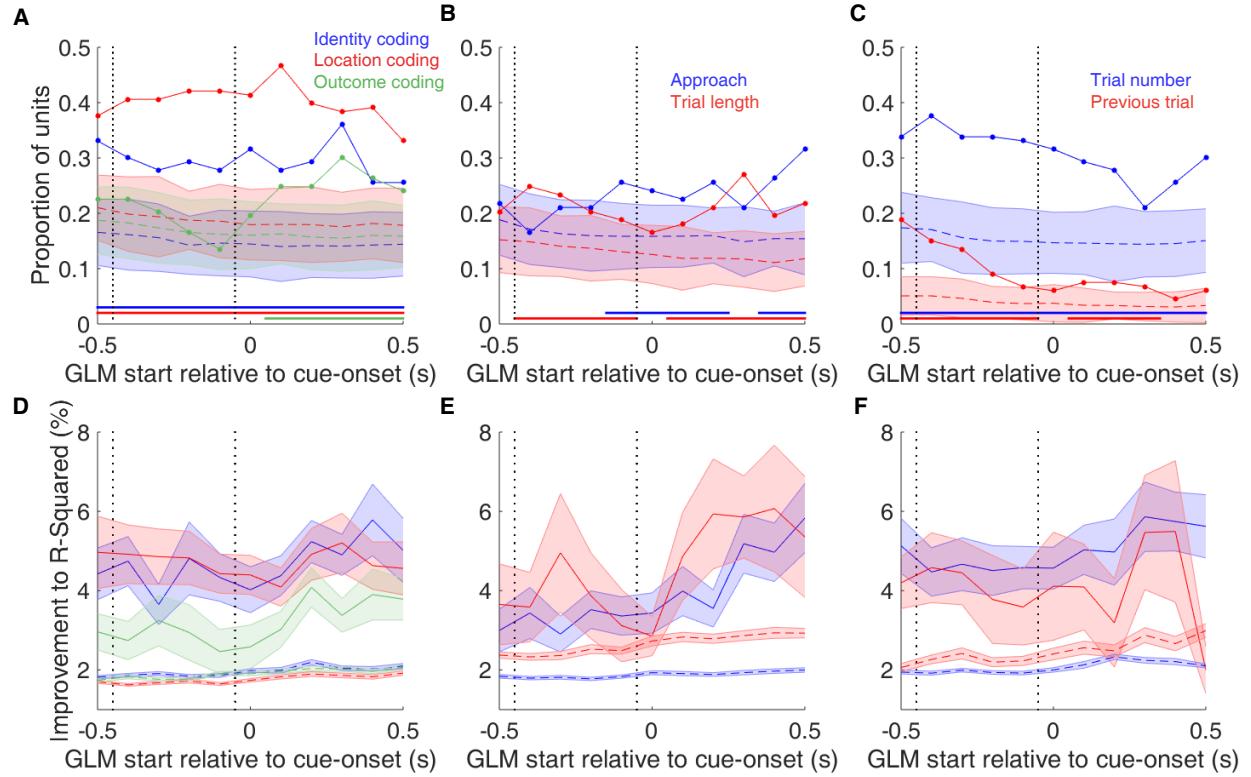


**Figure supplement 2** Expanded examples of cue-modulated NAc units influenced by different task parameters for Figure 4E-H, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).

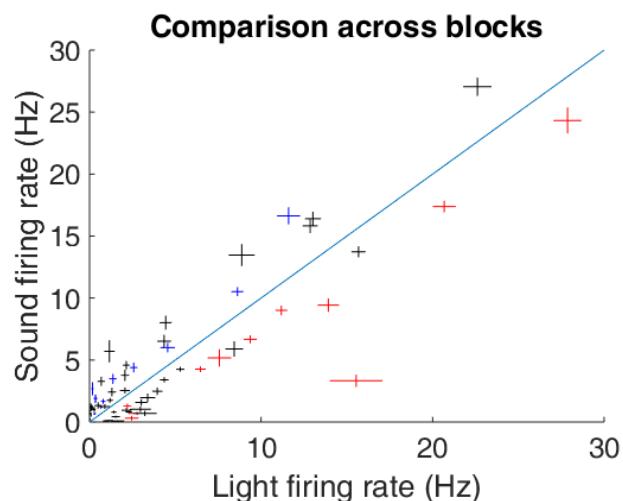
111 To characterize more formally whether these cue-evoked responses were modulated by various aspects of the  
112 task, we fit a sliding window generalized linear model (GLM) using a forward selection stepwise procedure to  
113 the firing rate of each cue-modulated unit surrounding cue-onset, using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



**Figure 5:** Summary of influence of various task parameters on cue-modulated NAc units after 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 samples ( $z$ -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 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 samples ( $z$ -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 different possible alternatives for coding across cue features at cue-onset, by correlating the recoded beta coefficients from the GLMs (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 block representing correlations across the post- cue-onset time bins from the sliding window GLM for two various cue features, with green representing positive correlations ( $r > 0$ ), pink representing negative correlations ( $r < 0$ ), and grey representing no significant correlation ( $r = 0$ ). X- and y-axis are the same, therefore the diagonal represents a correlation of an array against itself. Here the 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 grey in the off-diagonal elements. Bottom right: Joint coding of cue features in an overlapping population, shown here by the green in the off-diagonal elements. **D:** Correlation matrix showing the correlation among identity, location, and outcome coding at cue-onset. Overabundance of grey, and relative lack of green in the off-diagonal elements suggests that coding of cue features occur independently of one another at cue-onset. Scale bar indicates relative bin size of the sliding window.



**Figure supplement 1:** Summary of influence of various task parameters on cue-modulated NAc units after 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 samples ( $z$ -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 previous trial. **D-F:** Average improvement to model fit. **D:** Average percent improvement to R-squared 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 previous trial.



**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 SEM. Colored crosses represents units that had cue identity as a significant predictor of firing rate variance (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 lines indicates point of equal firing across blocks.

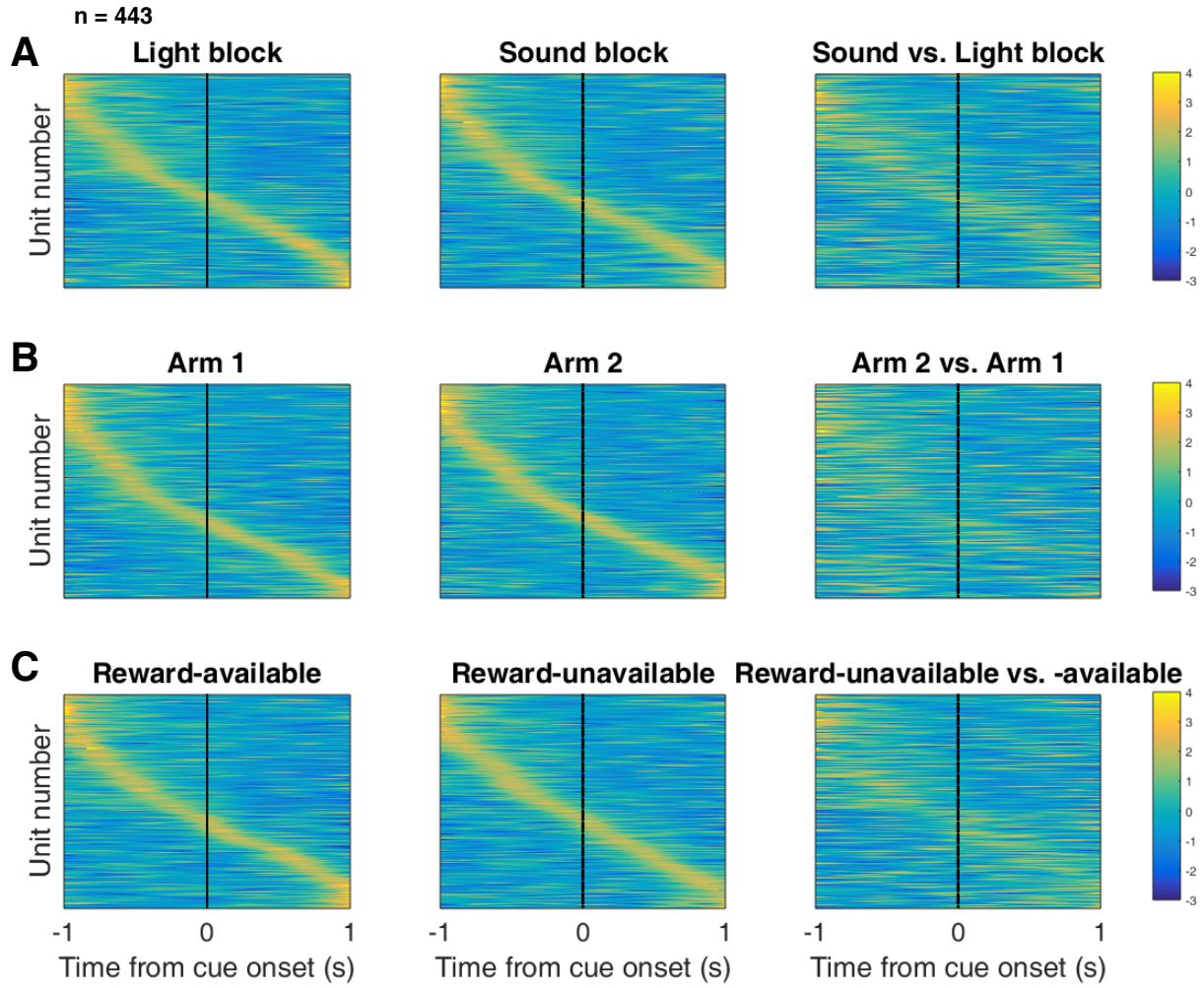
125 To assess what information may be encoded at the population level, we trained a classifier on a pseudoensemble  
126 of the 133 cue-modulated units (Figure 5B). Specifically, we used firing rate across units for each trial  
127 as an observation, and different cue conditions as trial labels (e.g.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

134 To determine whether the overlap of cue feature coding was different than expected by chance we correlated  
135 recoded beta coefficients from the GLMs, and compared the obtained correlations to the shuffled data gen-  
136 erated for the GLMs (Figure 1A,C,5C,D). This revealed that identity was coded separately from outcome  
137 (mean  $r = .009$ ; -6.74 standard deviations from the shuffled mean), and independently from location (mean  $r$   
138 = .097; 0.58 standard deviations from the shuffled mean), while location and outcome were coded by a joint  
139 population of units (mean  $r = .119$ ; 2.23 standard deviations from the shuffled mean). Together, these find-  
140 ings show that various cue features are represented in the NAc at both the single-unit and pseudoensemble  
141 level, that identity is coded independently or in separate populations form location and outcome, respectively,  
142 and that location is coded jointly with outcome.

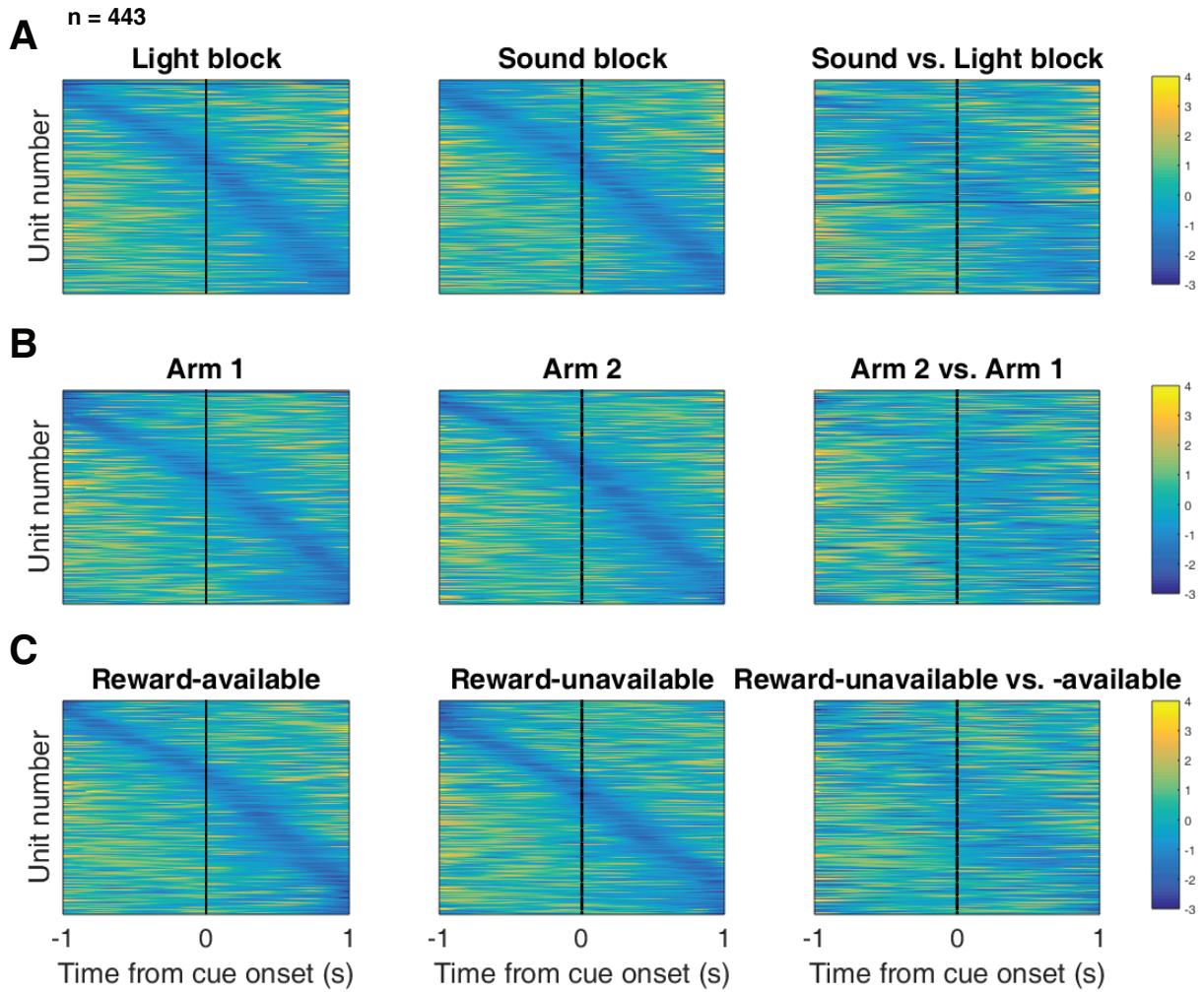
143 **NAc units dynamically segment the task:**

144 Next, we sought to determine how coding of cue features evolved over time. Two main possibilities can  
145 be distinguished (Figure 1B), a unit coding for a feature such as cue outcome could remain persistently  
146 active, or a progression of distinct units could activate in sequence. To visualize the distribution of responses  
147 throughout our task space and test if this distribution is modulated by cue features, we z-scored the firing  
148 rate of each unit and plotted the normalized firing rates of all units aligned to cue-onset and sorted them

according to the time of peak firing rate (Figure 6). We did this separately for both the light and sound blocks, and found a nearly uniform distribution of firing fields in task space that was not limited to alignment to the cue (Figure 6A). Furthermore, to determine if this population level activity was similar across blocks, we also organized firing during the sound blocks according to the ordering derived from the light blocks. This revealed that while there was some preservation of order, the overall firing was qualitatively different across the two blocks, implying that population activity distinguishes between light and sound blocks. To control for the possibility that any comparison of trials would produce this effect, we did a within block comparison, comparing half of the trials in the light block against the other half. This comparison looked similar to our test comparison of sound block trials ordered by light block trials. Additionally, given that the majority of our units showed an inhibitory response to the cue, we also plotted the firing rates according to the lowest time in firing, and again found some maintenance of order, but largely different ordering across the two blocks, and the within block comparison (Figure 6 supplement 1). To further test this, we divided each block into two halves and looked at the correlation of the average smoothed firing rates across various combinations of these halves across our cue-aligned centered epoch. A linear mixed effects model revealed that within block correlations (e.g. one half of light trials vs other half of light trials) were higher and more similar than across block correlations (e.g. half of light trials vs half of sound trials) suggesting that activity in the NAc discriminates across various cue conditions (within block correlations = .383 (light), .379 (sound); across block correlations = .343, .338, .337, .348; within block vs. within block comparison = p = .934; within block vs. across block comparisons = p < .001). This process was repeated for cue location (Figure 6B; within block correlations = .369 (arm 1), .350 (arm 2); across block correlations = .290, .286, .285, .291; within block vs. within block comparison = p = .071; within block vs. across block comparisons = p < .001) and cue outcome (Figure 6C; within block correlations = .429 (reward-available), .261 (reward-unavailable); across block correlations = .258, .253, .255, .249; within block vs. within block comparison = p < .001; within block vs. across block comparisons = p < .001), showing that NAc segmentation of the task is qualitatively different even during those parts of the task not immediately associated with a specific cue, action, or outcome, although the within condition comparison of reward-unavailable trials was less correlated than reward-available trials, and more similar to the across condition comparisons, potentially due to the greater behavioral variability for the reward-unavailable trials.



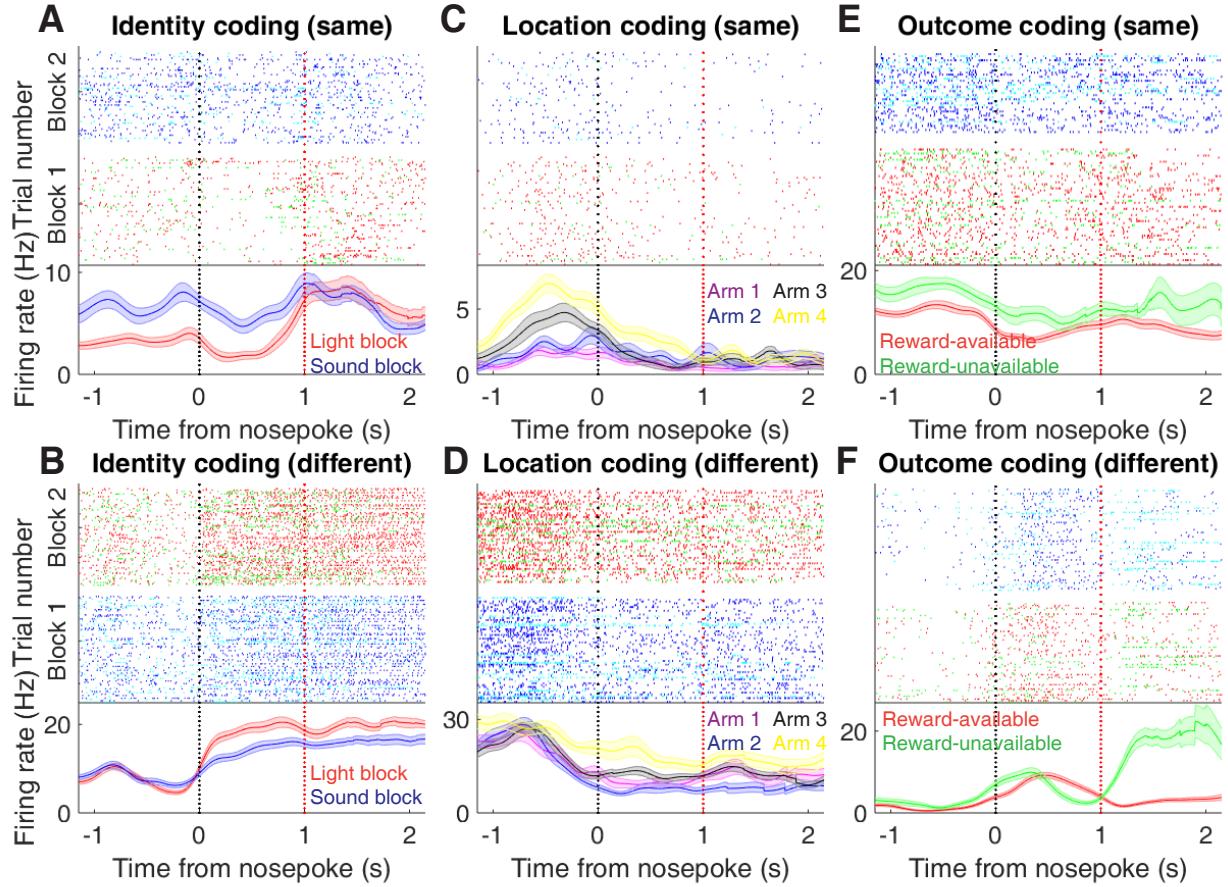
**Figure 6:** Distribution of NAc firing rates across time surrounding cue-onset. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue-onset), averaged across all trials for a specific cue type, indicated by text labels. **A-C:** Heat plots aligned to normalized peak firing rates. **A:** left: Heat plot showing smoothed normalized firing activity of all recorded NAc units ordered according to the time of their peak firing rate during the light block. Each row is a units average activity across time to the light block. Dashed line indicates cue-onset. Notice the yellow band across time, indicating all aspects of visualized task space were captured by the peak firing rates of various units. A, middle: Same units ordered according to the time of the peak firing rate during the sound block. Note that for both blocks, units tile time approximately uniformly with a clear diagonal of elevated firing rates. A, right: Unit firing rates taken from the sound block, ordered according to peak firing rate taken from the light block. Note that a weaker but still discernible diagonal persists, indicating partial similarity between firing rates in the two blocks. A similar pattern exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text). **B:** Same layout as in A, except that the panels now compare two different locations on the track instead of two cue modalities. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C:** Same layout as in A, except that panels now compare reward-available and reward-unavailable trials. 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.



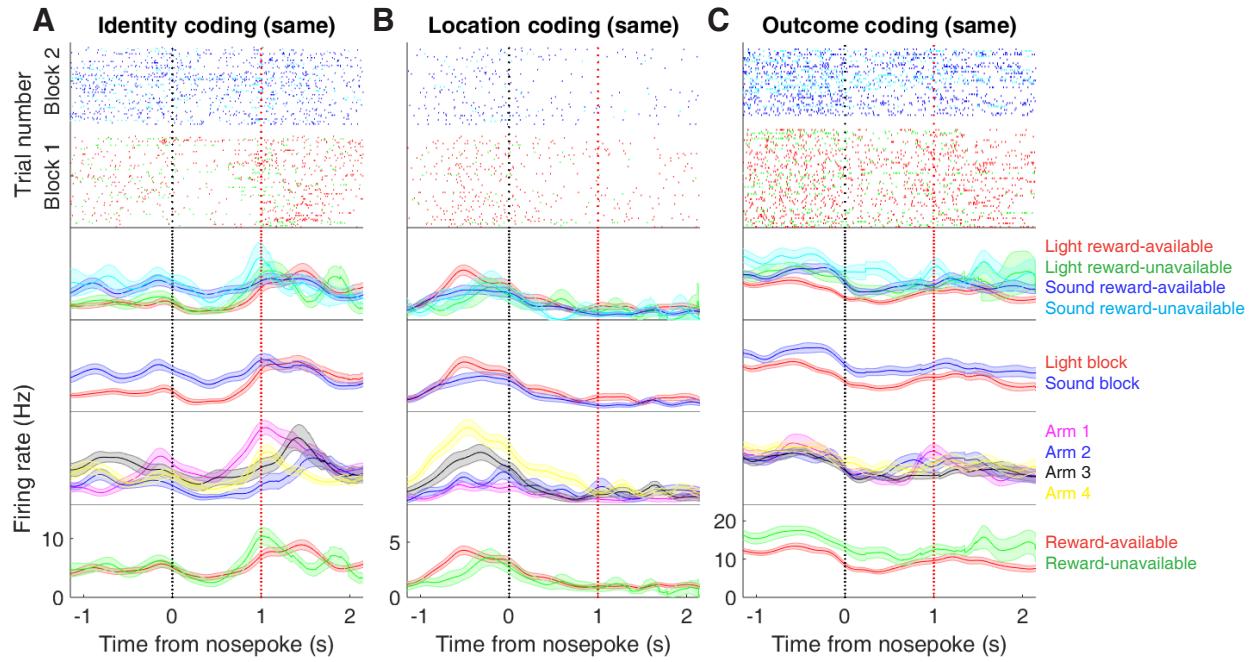
**Figure supplement 1:** Distribution of NAc firing rates across time surrounding cue-onset. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue-onset), averaged across all trials for a specific cue type, indicated by text labels. **A-C:** Heat plots aligned to normalized minimum firing rates. **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.

<sup>177</sup> **NAc encoding of cue features persists until outcome:**

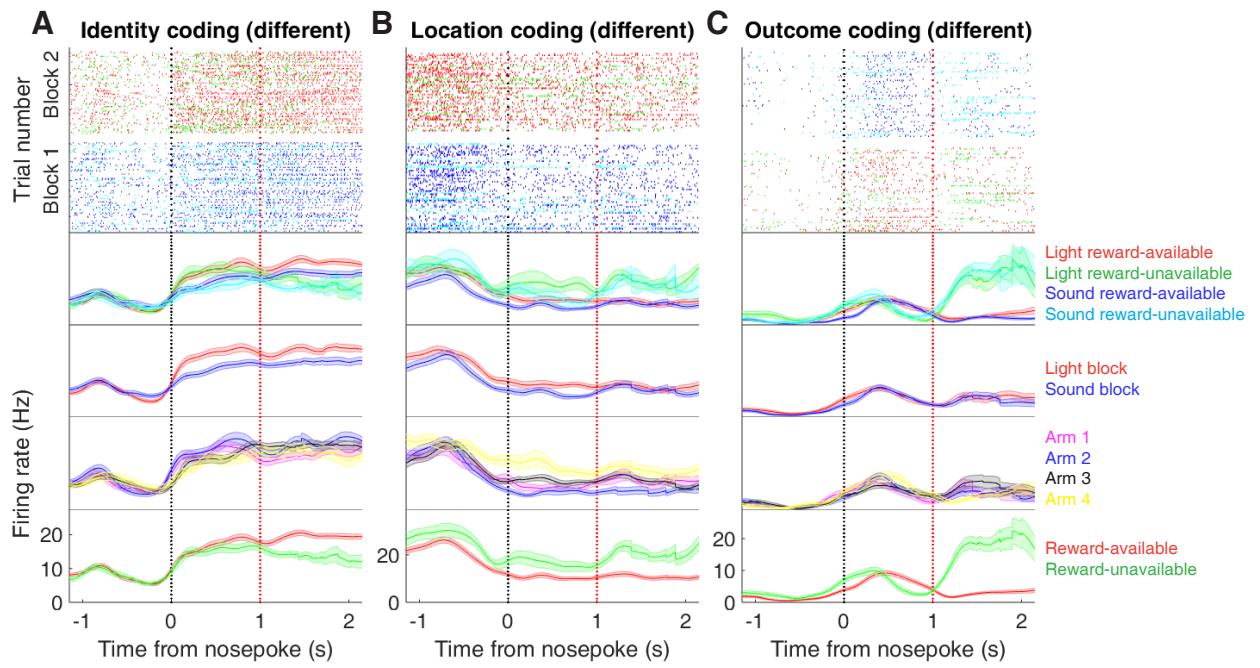
<sup>178</sup> In order to be useful for credit assignment in reinforcement learning, a trace of the cue must be maintained  
<sup>179</sup> until the outcome, so that information about the outcome can be associated with the outcome-predictive cue  
<sup>180</sup> (Figure 1B). Investigation into the post-approach period during nosepoke revealed units that discriminated  
<sup>181</sup> various cue features at cue-onset, and those that did not discriminate prior to nosepoke (Figure 7, supplement  
<sup>182</sup> 1, supplement 2). To quantitatively test whether representations of cue features persisted post-approach until  
<sup>183</sup> the outcome was revealed, we fit sliding window GLMs to the post-approach firing rates of cue-modulated  
<sup>184</sup> units aligned to both the time of nosepoke into the reward receptacle, and after the outcome was revealed  
<sup>185</sup> (Figure 8A,B, supplement 1 A-D, Table 1). This analysis showed that a variety of units still discriminated  
<sup>186</sup> firing according to cue identity (~20% of cue-modulated units), location (~25% of cue-modulated units),  
<sup>187</sup> and outcome (~25% of cue-modulated units), but not other task parameters, showing that NAc activity  
<sup>188</sup> discriminates various cue conditions well into a trial.



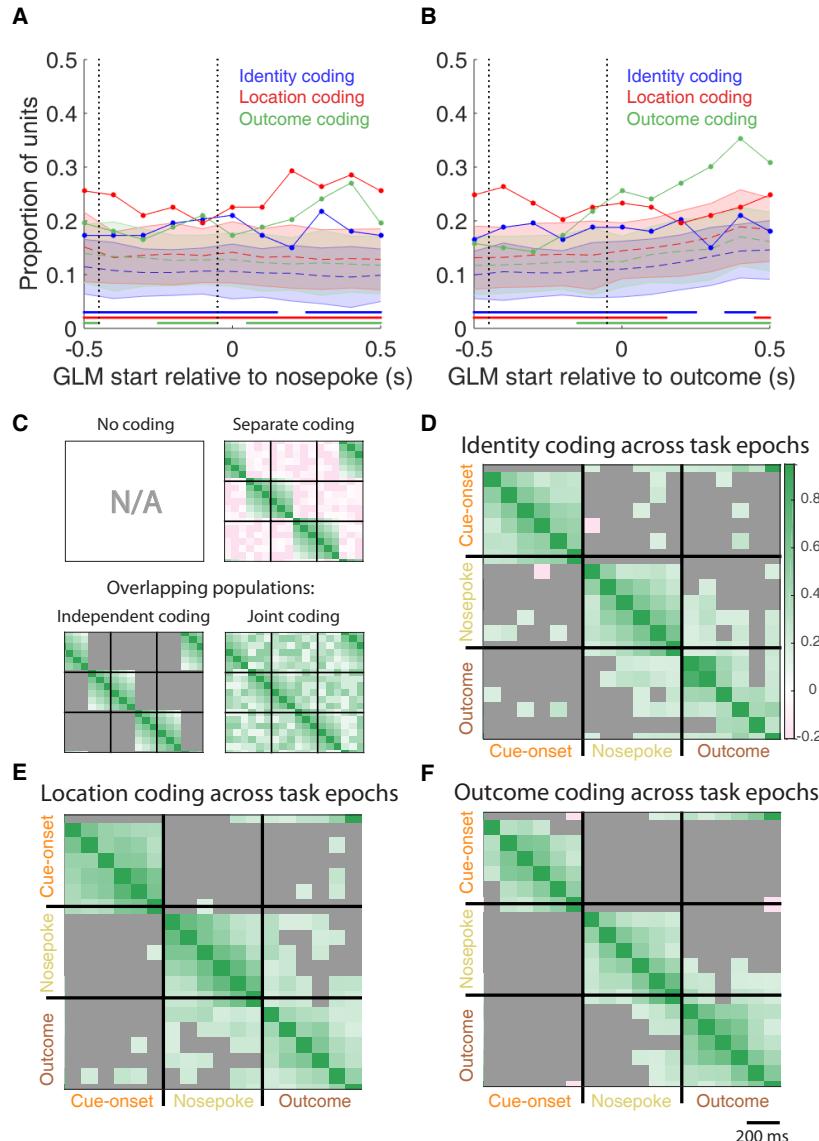
**Figure 7:** Examples of cue-modulated NAc units influenced by various task parameters at time of nosepoke. **A:** Example of a cue-modulated NAc unit that exhibited identity coding at both cue-onset and during nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to nosepoke. Spikes across trials are color coded according to cue type (red: reward-available light; green: reward-unavailable light; navy blue: reward-available sound; light blue: reward-unavailable sound). White space halfway up the rasterplot indicates switching from one block to the next. Black dashed line indicates nosepoke. Red dashed line indicates receipt of outcome. Bottom: PETHs showing the average smoothed firing rate for the unit for trials during light (red) and sound (blue) blocks, aligned to nosepoke. Lightly shaded area indicates standard error of the mean. Note this unit showed a sustained increase in firing to sound cues during the trial. **B:** An example of a unit that was responsive to cue identity at time of nosepoke but not cue-onset. **C-D:** Cue-modulated units that 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.



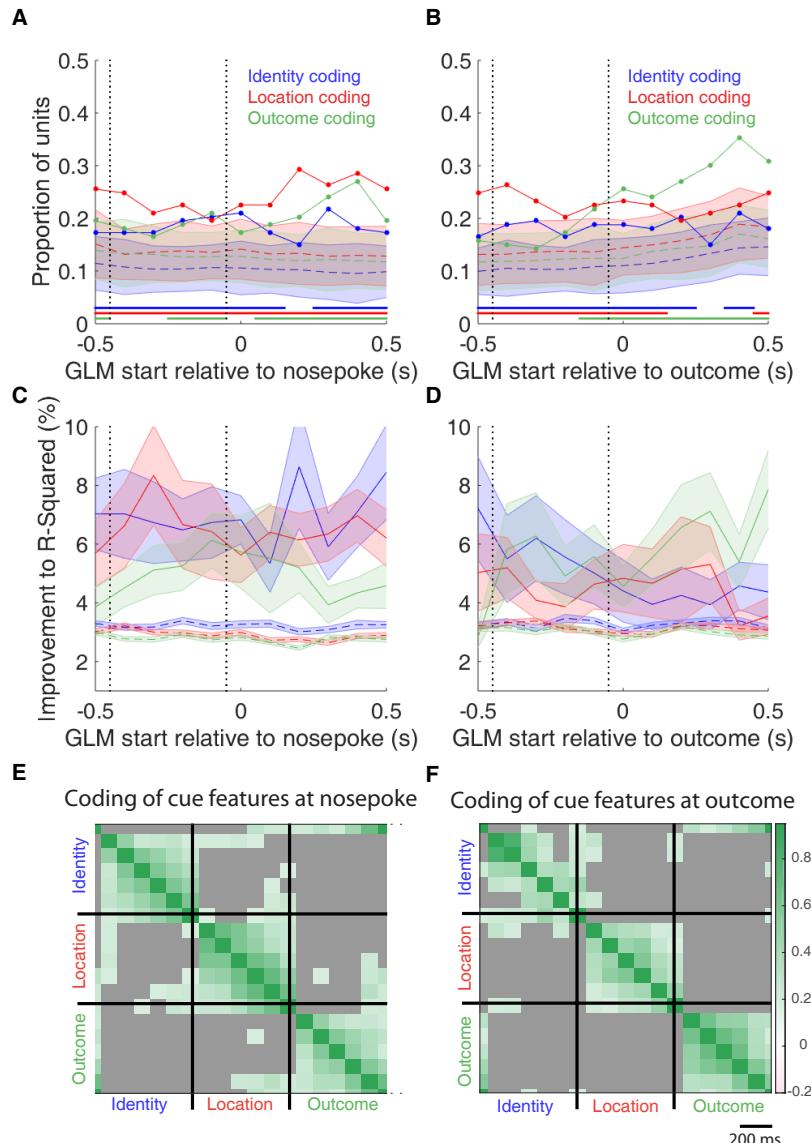
**Figure supplement 1** Expanded examples of cue-modulated NAc units influenced by different task parameters at time of nosepoke for Figure 7A,C,E, showing firing rate breakdown by: cue type (top PETH), cue identity (top-middle PETH), cue location (bottom-middle PETH), and cue outcome (bottom PETH).



**Figure supplement 2** Expanded examples of cue-modulated NAc units influenced by different task parameters 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 various task parameters on cue-modulated NAc units during 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 samples ( $z$ -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 different possible alternatives for coding of a cue feature across various points in a trial, by correlating the recoded beta coefficients from the GLMs (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 block representing correlations across various event-centered time bins from the sliding window GLM for a single cue feature, with green representing positive correlations ( $r > 0$ ), pink negative correlations ( $r < 0$ ), and grey representing no significant correlation ( $r = 0$ ). X- and y-axis are the same, therefore the diagonal represents a correlation of an array against itself. Here the pink in the off-diagonal elements suggests that coding of a cue feature is accomplished by separate populations of units across time. Bottom left: Coding of a cue feature across time occurs in overlapping but independent populations of units, shown here by the abundance of grey in the off-diagonal elements. Bottom right: Joint coding of a cue feature in an overlapping population across time, shown here by the

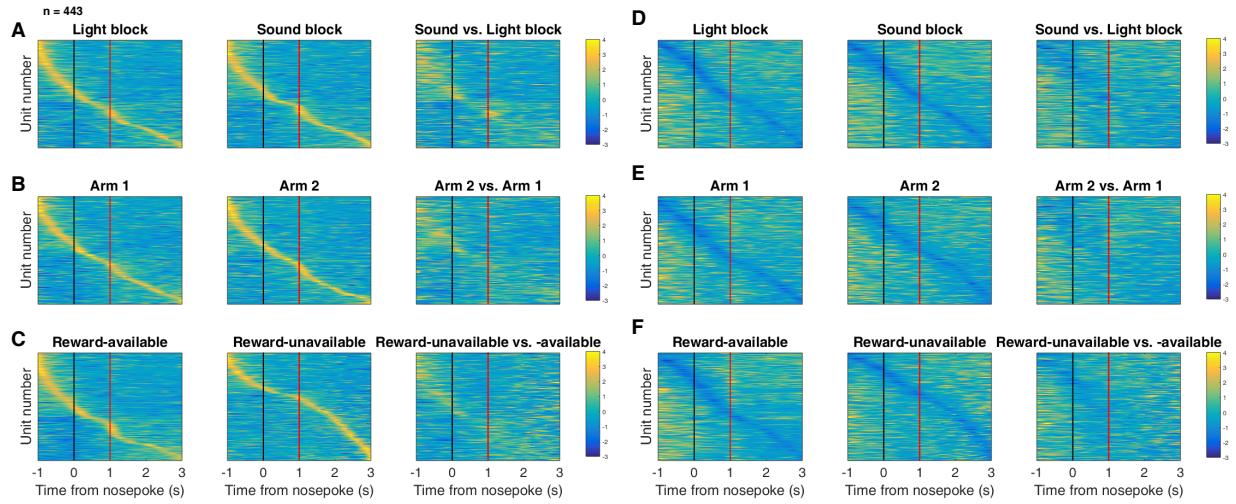


**Figure supplement 1:** Summary of influence of various task parameters on cue-modulated NAc units during nosepoke and subsequent receipt of outcome. **A-B:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of nosepoke (A), and outcome (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 samples ( $z$ -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 block represents correlations across various nosepoke-centered time bins from the sliding window GLM across cue features, with green representing positive correlations ( $r > 0$ ), pink negative correlations ( $r < 0$ ), and grey representing no significant correlation ( $r = 0$ ). X- and y-axis are the same, therefore the diagonal represents a correlation of an array against itself. Scale bar indicates relative bin size of the sliding window. **F:** Same as E, but time bins following outcome.

189 To determine whether NAc representations of cue features at nosepoke and outcome were encoded by a  
190 similar pool of units as during cue-onset, we correlated recoded beta coefficients from the GLMs for a cue  
191 feature across time points in the task, and compared the obtained correlations to the shuffled data generated  
192 for the GLMs (Figure 1B,C,8C-F). This revealed that identity coding at cue-onset was separate from coding  
193 at nosepoke (mean  $r = .048$ ; -2.18 standard deviations from the shuffled mean), independent from coding at  
194 outcome (mean  $r = .081$ ; 1.02 standard deviations from the shuffled mean), while joint coding was observed  
195 between nosepoke and outcome (mean  $r = .147$ ; 3.94 standard deviations from the shuffled mean); location  
196 coding at cue-onset was independent from coding at nosepoke (mean  $r = .058$ ; -1.41 standard deviations  
197 from the shuffled mean), while joint coding was observed with coding at outcome (mean  $r = .093$ ; 2.03  
198 standard deviations from the shuffled mean), as well as between nosepoke and outcome (mean  $r = .204$ ;  
199 9.35 standard deviations from the shuffled mean); and outcome coding at cue-onset was separate from from  
200 both coding at nosepoke (mean  $r = -.040$ ; -11.73 standard deviations from the shuffled mean), and coding at  
201 outcome (mean  $r = .025$ ; -5.06 standard deviations from the shuffled mean), while joint coding was observed  
202 between nosepoke and outcome (mean  $r = .148$ ; 3.15 standard deviations from the shuffled mean). In terms  
203 of overlap among cue features (Figure 8 supplement 1 E,F), at nosepoke identity was coded independently  
204 from location (mean  $r = .124$ ; 1.00 standard deviations from the shuffled mean), and separately from outcome  
205 (mean  $r = .052$ ; -3.19 standard deviations from the shuffled mean), and location and outcome were coded  
206 by a overlapping but independent population of units (mean  $r = .097$ ; -1.05 standard deviations from the  
207 shuffled mean); while at outcome, identity was coded independently from location (mean  $r = .085$ ; -1.19  
208 standard deviations from the shuffled mean), and separately from outcome (mean  $r = .039$ ; -4.20 standard  
209 deviations from the shuffled mean), and location and outcome were coded by a separate population of units  
210 (mean  $r = .004$ ; -7.22 standard deviations from the shuffled mean). Together, these findings show that  
211 the NAc maintains representations of cue features after a behavioral decision has been made with largely  
212 separate or independent pools of units, while a joint overlapping population is seen post-approach until it  
213 receives feedback about its decision.

214 To assess the tiling of units around outcome receipt, we aligning normalized peak firing rates to nosepoke

215 onset (Figure 8 supplement 2). This revealed a clustering of responses around outcome receipt for all cue  
216 conditions where the rat would have received reward, in addition to the same trend of higher within- vs  
217 across-block correlations for cue identity (Figure 8 supplement 2 A,C; within block correlations = .560  
218 (light), .541 (sound); across block correlations = .487, .481, .483, .486; within block vs. within block  
219 comparison =  $p = .112$ ; within block vs. across block comparisons =  $p < .001$ ) and cue location (Figure  
220 8 supplement 2 B,E; within block correlations = .474 (arm 1), .461 (arm 2); across block correlations =  
221 .416, .402, .416, .415; within block vs. within block comparison =  $p = .810$ ; within block vs. across block  
222 comparisons =  $p < .001$ ), but not cue outcome (Figure 8 supplement 2 C,F; within block correlations =  
223 .620 (reward-available), .401 (reward-unavailable); across block correlations = .418, .414, .390, .408; within  
224 block vs. within block comparison =  $p < .001$ ; within block vs. across block comparisons =  $p < .001$ ),  
225 further reinforcing that the NAc segments the task 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. A similar pattern exists for within-block comparisons suggesting that reordering any two sets of trials produces this partial similarity, however correlations within blocks are more similar than correlations across blocks (see text). **B:** Same layout as in A, except that the panels now compare two different locations on the track instead of two cue modalities. As for the different cue modalities, NAc units clearly discriminate between locations, but also maintain some similarity across locations, as evident from the visible diagonal in the right panel. Two example locations were used for display purposes; other location pairs showed a similar pattern. **C:** Same layout as in A, except that panels now compare correct reward-available and incorrect reward-unavailable trials. The disproportionate tiling around outcome receipt for reward-available, but not reward-unavailable trials suggests encoding of reward receipt by NAc units. **D-F:** Heat plots aligned to normalized minimum firing rates. **D:** Responses during different stimulus blocks as in A, but with units ordered according to the time of their minimum firing rate. **E:** Responses during trials on different arms as in B, but with units ordered by their minimum firing rate. **F:** Responses during cues signalling different outcomes as in C, but with units ordered by their minimum firing rate. Overall, NAc units "tiled" experience on the task, as opposed to being confined to specific task events only. Units from all sessions and animals were pooled for this analysis.

226 **Discussion**

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

239 **Identity coding:**

240 Our finding that NAc units can discriminate between different outcome-predictive stimuli with similar moti-  
241 vational significance (i.e. encodes cue identity) expands upon an extensive rodent literature examining NAc  
242 correlates of conditioned stimuli (Ambroggi, Ishikawa, Fields, & Nicola, 2008; Atallah et al., 2014; Bis-  
243 sonette et al., 2013; Cooch et al., 2015; Day et al., 2006; Dejean et al., 2017; Goldstein et al., 2012; Ishikawa,  
244 Ambroggi, Nicola, & Fields, 2008; Lansink et al., 2012; McGinty et al., 2013; Nicola, 2004; Roesch et al.,  
245 2009; Roitman et al., 2005; Saddoris et al., 2011; Setlow et al., 2003; Sugam et al., 2014; West & Carelli,  
246 2016; Yun, Wakabayashi, Fields, & Nicola, 2004). Perhaps the most comparable work in rodents comes  
247 from a study that found distinct coding for an odor when it predicted separate but equally valued rewards  
248 (Cooch et al., 2015). The present work is complementary to such *outcome identity* coding as it shows that  
249 NAc units encode *cue identity*, in addition to the reward it predicts (Figure 1A). Similarly, Setlow et al.

250 (2003) paired distinct odor cues with appetitive or aversive outcomes, and found separate populations of  
251 units that encoded each cue. Furthermore, during a reversal they found that the majority of units switched  
252 their selectivity, arguing that the NAc units were tracking the motivational significance of these stimuli.  
253 Once again, our study was different in asking how distinct cues encoding the same anticipated outcome are  
254 encoded. Such cue identity encoding suggests that even when the motivational significance of these stimuli  
255 is identical, NAc dissociates their representations at the level of the single-units. A possible interpretation  
256 of this coding of cue features alongside expected outcome is that these representations are used to associate  
257 reward with relevant features of the environment, so-called credit assignment in the reinforcement learning  
258 literature (Sutton & Barto, 1998). A burgeoning body of human and non-human primate work has started to  
259 elucidated neural correlates of credit assignment in the PFC, particularly in the lateral orbitofrontal cortex  
260 (Akaishi, Kolling, Brown, & Rushworth, 2016; Asaad, Lauro, Perge, & Eskandar, 2017; Chau et al., 2015;  
261 Noonan, Chau, Rushworth, & Fellows, 2017). Given the importance of cortical inputs in NAc associative  
262 representations, it is possible that information related to credit assignment is relayed from the cortex to NAc  
263 (Coch et al., 2015; Ishikawa et al., 2008).

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

271 A different possible function for cue identity coding is to support contextual modulation of the motivational  
272 relevance of specific cues. A context can be understood as a particular mapping between specific cues and  
273 their outcomes: for instance, in context 1 cue A but not cue B is rewarded, whereas in context 2 cue B  
274 but not cue A is rewarded. Successfully implementing such contextual mappings requires representation of

275 the cue identities. Indeed, Sleezer et al. (2016) recorded NAc responses during the Wisconsin Card Sorting  
276 Task, a common set-shifting task used in both the laboratory and clinic, and found units that preferred firing  
277 to stimuli when a certain rule, or rule category was currently active. Further support for a modulation of  
278 NAc responses by strategy comes from an fMRI study that examined BOLD levels during a set-shifting task  
279 (FitzGerald et al., 2014). In this task, participants learned two sets of stimulus-outcome contingencies, a  
280 visual set and an auditory set. During testing they were presented with both simultaneously, and the stimulus  
281 dimension that was relevant was periodically shifted between the two. Here, they found that bilateral NAc  
282 activity reflected value representations for the currently relevant stimulus dimension, and not the irrelevant  
283 stimulus. The current finding of independent coding of cue identity and expected outcome, suggests that the  
284 fMRI finding is generated by the combined activity of several different functional cell types.

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

294 A different potential confound is that between outcome and action value coding as the rat was only rewarded  
295 for left turns. Our GLM analysis dealt with this by excluding firing rate variance accounted for by a predictor  
296 that represented whether the animal approached (left turn) or skipped (right turn) the reward port at the  
297 choice point, thus we were able to identify units that were modulated by the expected outcome of the cue  
298 after removing variance for potential action value coding. Additionally, NAc signals have been shown to  
299 be modulated by response vigor (McGinty et al., 2013); to detangle this from our results we included trial

300 length (i.e. latency to arrival at the reward site) as a predictor in our GLMs, and found units with cue feature  
301 correlates independent of trial length.

302 An overall limitation of the current study is that rats were never presented with both sets of cues simultaneously,  
303 and were not required to switch strategies between multiple sets of cues (in behavioral pilots, animals took several days of training to successfully switch strategies). Additionally, our recordings were  
304 done during performance on the well-learned behavior, and not during the initial acquisition of the cue-  
305 outcome relationships when an eligibility trace would be most useful. Thus, it is unknown to what extent  
306 the cue identity encoding we observed is behaviorally relevant, although extrapolating data from other work  
307 (Sleezier et al., 2016) suggests that cue identity coding would be modulated by relevance. NAc core lesions  
308 have been shown to impair shifting between different behavioral strategies (Floresco et al., 2006), and it is  
309 possible that selectively silencing the units that prefer responding for a given modality or rule would impair  
310 performance when the animal is required to use that information, or artificial enhancement of those units  
311 would cause them to use the rule when it is the inappropriate strategy.  
312

313 **Encoding of position:**

314 Our finding that cue-evoked activity was modulated by cue location is in alignment with several previous  
315 reports (Lavoie & Mizumori, 1994; Mulder, Shibata, Trullier, & Wiener, 2005; Strait et al., 2016; Wiener et  
316 al., 2003). The NAc receives inputs from the hippocampus, and the communication of place-reward informa-  
317 tion across the two structures suggests that the NAc tracks locations associated with reward (Lansink et al.,  
318 2008; Lansink, Goltstein, Lankelma, McNaughton, & Pennartz, 2009; Lansink et al., 2016; Pennartz, 2004;  
319 Sjulson, Peyrache, Cumpelik, Cassataro, & Buzsáki, 2017; Tabuchi, Mulder, & Wiener, 2000; van der Meer  
320 & Redish, 2011). However, it is notable that in our task, location is explicitly uninformative about reward,  
321 yet coding of this uninformative variable persists. The tiling of task space is in alignment with previous  
322 studies showing that NAc units can also signal progress through a sequence of cues and/or actions (Atallah  
323 et al., 2014; Berke, Breck, & Eichenbaum, 2009; Khamassi, Mulder, Tabuchi, Douchamps, & Wiener, 2008;

324 Lansink et al., 2012; Mulder, Tabuchi, & Wiener, 2004; Shidara, Aigner, & Richmond, 1998), is similar to  
325 observations in the basal forebrain (Tingley & Buzsáki, 2018), and may represent a temporally evolved state  
326 value signal (Hamid et al., 2015; Pennartz, Ito, Verschure, Battaglia, & Robbins, 2011). Given that the cur-  
327 rent task was pseudo-random, it is possible that the rats learned the structure of sequential cue presentation,  
328 and the neural activity could reflect this. However, this is unlikely as including a previous trial variable in  
329 the analysis did not explain a significant amount of firing rate variance in response to the cue for the vast  
330 majority of units. In any case, NAc units on the present task continued to distinguish between different  
331 locations, even though location, and progress through a sequence, were explicitly irrelevant in predicting re-  
332 ward. We speculate that this persistent coding of location in NAc may represent a bias in credit assignment,  
333 and associated tendency for rodents to associate motivationally relevant events with the locations where they  
334 occur.

335 **Implications:**

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

349 as inappropriate credit assignment could be a contributor to conditioned fear overgeneralization seen in dis-  
350 orders with pathological anxiety such as generalized anxiety disorder, post traumatic stress disorder, and  
351 obsessive-compulsive disorder (Kaczkurkin et al., 2017; Kaczkurkin & Lissek, 2013; Lissek et al., 2014),  
352 and delusions observed in disorders such as schizophrenia, Alzheimer's and Parkinson's (Corlett, Taylor,  
353 Wang, Fletcher, & Krystal, 2010; Kapur, 2003). Thus, our results provide a neural window into the process  
354 of credit assignment, such that the extent and specific manner in which this process fails in syndromes such  
355 as schizophrenia, obsessive-compulsive disorder, etc. can be experimentally accessed.

## 356 **Methods**

### 357 **Subjects:**

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

### 365 **Overall timeline:**

366 Each rat was first handled for seven days during which they were exposed to the experiment room, the  
367 sucrose solution used as a reinforcer, and the click of the sucrose dispenser valves. Rats were then trained  
368 on the behavioral task (described in the next section) until they reached performance criterion. At this point

369 they underwent hyperdrive implantation targeted at the NAc. Rats were allowed to recover for a minimum  
370 of five days before being retrained on the task, and recording began once performance returned to pre-  
371 surgery levels. Upon completion of recording, animals were glosed, euthanized and recording sites were  
372 histologically confirmed.

373 **Behavioral task and training:**

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

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

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

406 **Surgery:**

407 Surgical procedures were as described previously (Malhotra, Cross, Zhang, & Van Der Meer, 2015). Briefly,  
408 animals were administered analgesics and antibiotics, anesthetized with isoflurane, induced with 5% in med-  
409 ical grade oxygen and maintained at 2% throughout the surgery ( 0.8 L/min). Rats were then chronically  
410 implanted with a “hyperdrive” consisting of 16 independently drivable tetrodes, either all 16 targeted for the  
411 right NAc (AP +1.4 mm and ML +1.6 mm relative to bregma; Paxinos & Watson 1998), or 12 in the right  
412 NAc and 4 targeted at the mPFC (AP +3.0 mm and ML +0.6 mm, relative to bregma; only data from NAc  
413 tetrodes was analyzed). Following surgery, all animals were given at least five days to recover while receiv-  
414 ing post-operative care, and tetrodes were lowered to the target (DV -6.0 mm) before being reintroduced to  
415 the behavioral task.

416 **Data acquisition and preprocessing:**

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

428 **Data analysis:**

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

436 *Neural data.* Given that some of our analyses compare firing rates across time, particularly comparisons  
437 across blocks, we sought to exclude units with unstable firing rates that would generate spurious results  
438 reflecting a drift in firing rate over time unrelated to our task. To do this we ran a Mann-Whitney U test  
439 comparing the cue-evoked firing rates for the first and second half of trials within a block, and excluded 99  
440 of 443 units from analysis that showed a significant change for either block, leaving 344 units for further

441 analyses by our GLM. To investigate the contribution of different cue features (*cue identity*, *cue location*  
442 and *cue outcome*) on the firing rates of NAc single units, we first determined whether firing rates for a unit  
443 were modulated by the onset of a cue by collapsing across all cues and comparing the firing rates for the 1  
444 s preceding cue-onset with the 1 s following cue-onset. Single units were considered to be *cue-modulated* if  
445 a Wilcoxon signed-rank test comparing pre- and post-cue firing was significant at  $p < .01$ . Cue-modulated  
446 units were then classified as either increasing or decreasing if the post-cue activity was higher or lower than  
447 the pre-cue activity, respectively.

448 To determine the relative contribution of different task parameters to firing rate variance (as in Figures 4,5A),  
449 a forward selection stepwise GLM using a Poisson distribution for the response variable was fit to each cue-  
450 modulated unit, using data from every trial in a session. Cue identity (light block, sound block), cue location  
451 (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available, reward-unavailable), behavior (approach,  
452 skip), trial length, trial number, and trial history (reward availability on the previous 2 trials) were used as  
453 predictors, with firing rate as the response variable. The GLMs were fit using a sliding window for firing  
454 rate; with a bin size of 500 ms, step size of 100 ms, and an analysis window from 500 ms pre-cue onset to  
455 500 ms post-cue onset, such that 11 different GLMs were fit for each unit, tracking the temporal dynamics  
456 of the influence of task parameters on firing rate around the onset of the cue. Units were classified as being  
457 modulated by a given task parameter if addition of the parameter significantly improved model fit using  
458 deviance as the criterion ( $p < .01$ ), and the total proportion of cue-modulated units influenced by a task  
459 parameter was counted for each time bin. A comparison of the R-squared value between the final model  
460 and the final model minus the predictor of interest was used to determine the amount of firing rate variance  
461 explained by the addition of that predictor for a given unit. To control for the amount of units that would  
462 be affected by a predictor by chance, we shuffled the trial order of firing rates for a particular unit within a  
463 time bin, and took the average of this value over 100 shuffles. We then calculated how many z-scores the  
464 observed proportion was from the mean of the shuffled samples, using a z-score of greater than 1.96 as a  
465 marker of significance.

466 To get a sense of the predictive power of these cue feature representations we trained a classifier using firing  
467 rates from a pseudoensemble of our 133 cue-modulated units (Figure 5B). We created a matrix of firing  
468 rates for each time epoch surrounding cue-onset where each row was an observation representing the firing  
469 rate for a trial, and each column was a variable representing the firing rate for a given unit. Trial labels, or  
470 classes, were each condition for a cue feature (e.g. light and sound for cue identity) when trying to predict  
471 each cue feature, making sure to align trial labels across units. We then ran LDA on these matrices, using  
472 10-fold cross validation to train the classifier on 90% of the trials and testing its predictions on the held out  
473 10% of trials, and repeated this approach to get the classification accuracy for 100 iterations. To test if the  
474 classification accuracy was greater by chance, we shuffled the order of firing rates for each unit before we  
475 trained the classifier. We repeated this for 100 shuffled matrices for each time point, and calculated how  
476 many z-scores the mean classification rate of the observed data was from the mean of the shuffled samples,  
477 using a z-score of 1.96 as the cut off for a significant deviation from the shuffled distribution.

478 To determine the degree to which coding of cue identity, cue location, and cue outcome overlapped within  
479 units we computed pearsons r on recoded beta coefficients from the GLMs for the cue features (Figure  
480 5C,D). Specifically, for all cue-modulated units we coded a 1 if the cue feature was a significant predictor  
481 in the final model, and 0 if it was not, and calculated pearsons r comparing an array of the coded 0s and  
482 1s for one cue feature with a similar array for another cue feature, repeating this process for all post cue-  
483 onset sliding window combinations. The NAc was determined as coding a pair of cue features in a) a  
484 separate populations of units if there was a significant negative correlation ( $r < 0$ ), b) an independently  
485 coded overlapping population of units if there was no significant correlation ( $r = 0$ ), and c) a jointly coded  
486 overlapping population of units if there was a significant positive correlation ( $r > 0$ ). To summarize the  
487 correlation matrices generated from this analyses, we put the output of our 100 shuffled GLMs through  
488 the same pipeline, took the mean of all correlations for a block comparison for each of the 100 shuffles  
489 for an analysis window, and used the mean and standard deviation of these shuffled correlation averages to  
490 compare to the mean of the comparison block of the correlation matrix for the actual data, and generate a  
491 z-score, using a z-score of greater than 1.96 or less than -1.96 as a marker that that comparison block showed

492 overlap across cue features greater or less than chance, respectively.

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

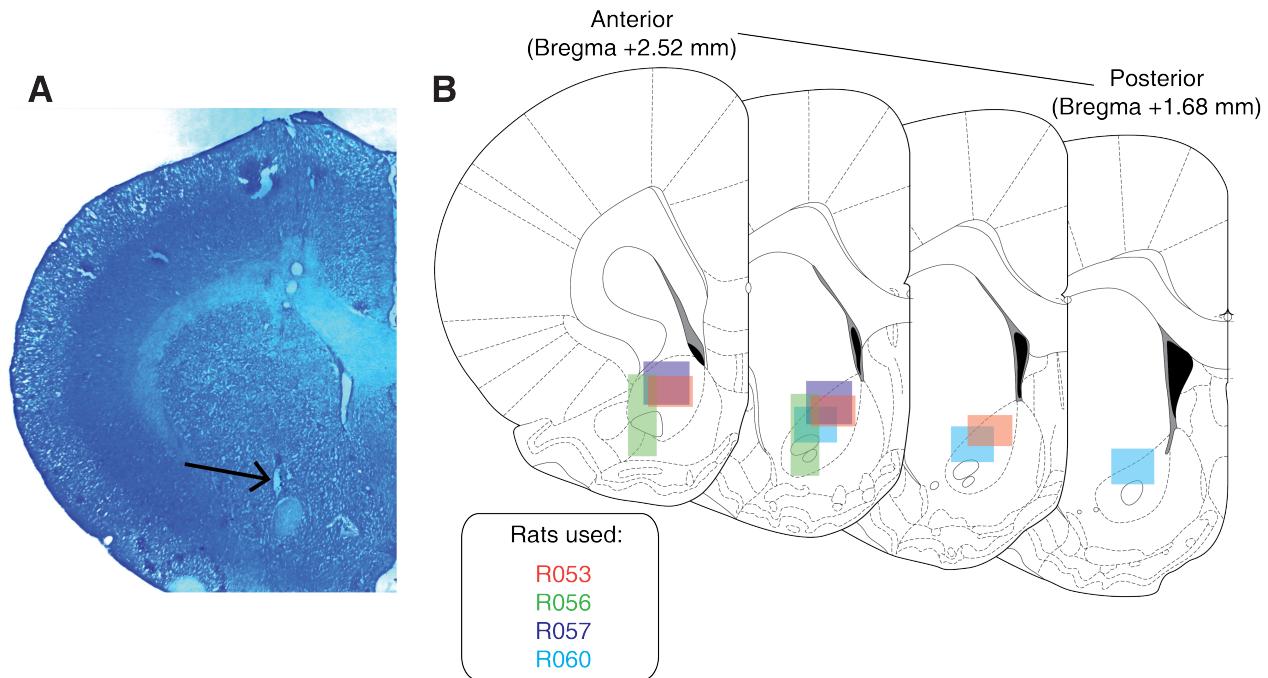
512 To identify the responsivity of units to different cue features at the time of nosepoke into a reward receptacle,  
513 and subsequent reward delivery, the same cue-responsive units from the cue-onset analyses were analyzed at  
514 the time of nosepoke and outcome receipt using identical analysis techniques for all approach trials (Figures  
515 7,8). To compare whether coding of a given cue feature was accomplished by the same or distinct population  
516 of units across time epochs, we ran the recoded coefficient correlation that was used to assess the degree of

517 overlap among cue features within a time epoch.

518 All analyses were completed in MATLAB R2015a, the code is available on our public GitHub repository  
519 (<http://github.com/vandermeelab/papers>), and the data can be accessed through DataLad.

520 **Histology:**

521 Upon completion of the experiment, recording channels were glosed by passing  $10 \mu A$  current for 10 sec-  
522 onds and waiting 5 days before euthanasia, except for rat R057 whose implant detached prematurely. Rats  
523 were anesthetized with 5% isoflurane, then asphyxiated with carbon dioxide. Transcardial perfusions were  
524 performed, and brains were fixed and removed. Brains were sliced in  $50 \mu m$  coronal sections and stained  
525 with thionin. Slices were visualized under light microscopy, tetrode placement was determined, and elec-  
526 trodes with recording locations in the NAc were analyzed (Figure 9).



**Figure 9:** 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.

527 **References**

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

- 558 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).  
559 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5),  
560 1106–1118. doi: 10.1016/j.neuron.2015.08.018
- 561 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., & Wightman, R. M. (2007, apr). Coordinated  
562 Accumbal Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2), 237–244. doi:  
563 10.1016/j.neuron.2007.03.021
- 564 Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in  
565 schizophrenia. *Psychological Review*, 99(1), 45.doi: 10.1016/j.neuron.2007.03.021
- 566 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., & Schoenbaum, G. (2015,  
567 jun). Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature  
568 Communications*, 6, 7195. doi: 10.1038/ncomms8195
- 569 Corbit, L. H., & Balleine, B. W. (2011, aug). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are  
570 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33), 11786–11794. doi:  
571 10.1523/JNEUROSCI.2711-11.2011
- 572 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in  
573 Neurobiology*, 92, 345–369. doi: 10.1016/j.pneurobio.2010.06.007
- 574 Cromwell, H. C., & Schultz, W. (2003, may). Effects of Expectations for Different Reward Magnitudes on Neuronal Activity in  
575 Primate Striatum. *Journal of Neurophysiology*, 89(5), 2823–2838. doi: 10.1152/jn.01014.2002
- 576 Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine  
577 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8), 1020–1028. doi: 10.1038/nn1923
- 578 Day, J. J., Wheeler, R. A., Roitman, M. F., & Carelli, R. M. (2006, mar). Nucleus accumbens neurons encode Pavlovian ap-  
579 proach behaviors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5), 1341–1351. doi:  
580 10.1111/j.1460-9568.2006.04654.x
- 581 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., ... Le Moine, C. (2017, apr). Memories of Opiate  
582 Withdrawal Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychophar-  
583 macology*, 42(5), 1157–1168. doi: 10.1038/npp.2016.272
- 584 du Hoffmann, J., & Nicola, S. M. (2014, oct). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the  
585 Nucleus Accumbens. *Journal of Neuroscience*, 34(43), 14349–14364. doi: 10.1523/JNEUROSCI.3492-14.2014
- 586 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental  
587 Psychology*, 32(2), 150–155. doi: 10.1037/h0058316
- 588 Fitzgerald, T. H. B., Schwartenbeck, P., & Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent  
589 and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4), 1271–1279. doi: 10.1523/JNEUROSCI.4389-  
590 13.2014

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

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

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

- 690 Discriminative Stimulus Task. *Journal of Neurophysiology*, 91(4), 1840–1865. doi: 10.1152/jn.00657.2003
- 691 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of  
692 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49), 16585–  
693 16600. doi: 10.1523/JNEUROSCI.3958-10.2010
- 694 Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007, mar). Tonic dopamine: Opportunity costs and the control of response vigor.  
695 *Psychopharmacology*, 191(3), 507–520. doi: 10.1007/s00213-006-0502-4
- 696 Noonan, M. P., Chau, B. K., Rushworth, M. F., & Fellows, L. K. (2017, jul). Contrasting Effects of Medial and Lateral Orbitofrontal  
697 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29), 7023–7035.  
698 doi: 10.1523/JNEUROSCI.0692-17.2017
- 699 Paxinos, G., & Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates* (4th ed.). San Diego: Academic Press. doi:  
700 10.1523/JNEUROSCI.0692-17.2017
- 701 Pennartz, C. M. A. (2004, jul). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation  
702 by Hippocampal Ripples. *Journal of Neuroscience*, 24(29), 6446–6456. doi: 10.1523/JNEUROSCI.0575-04.2004
- 703 Pennartz, C. M. A., Ito, R., Verschure, P., Battaglia, F., & Robbins, T. (2011, oct). The hippocampalstriatal axis in learning,  
704 prediction and goal-directed behavior. *Trends in Neurosciences*, 34(10), 548–559. doi: 10.1016/J.TINS.2011.08.001
- 705 Peters, J., & Büchel, C. (2009, dec). Overlapping and distinct neural systems code for subjective value during intertemporal and  
706 risky decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(50), 15727–34.  
707 doi: 10.1523/JNEUROSCI.3489-09.2009
- 708 Rescorla, R. A., & Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and  
709 Instrumental Learning. *Psychological Review*, 74(3), 151–182. doi: 10.1037/h0024475
- 710 Robinson, T. E., & Flagel, S. B. (2009, may). Dissociating the Predictive and Incentive Motivational Properties of  
711 Reward-Related Cues Through the Study of Individual Differences. *Biological Psychiatry*, 65(10), 869–873. doi:  
712 10.1016/j.biopsych.2008.09.006
- 713 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009, oct). Ventral Striatal Neurons Encode the Value  
714 of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42),  
715 13365–13376. doi: 10.1523/JNEUROSCI.2572-09.2009
- 716 Roitman, M. F., Wheeler, R. A., & Carelli, R. M. (2005, feb). Nucleus accumbens neurons are innately tuned for reward-  
717 ing and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4), 587–597. doi:  
718 10.1016/j.neuron.2004.12.055
- 719 Saddoris, M. P., Stamatakis, A., & Carelli, R. M. (2011, jun). Neural correlates of Pavlovian-to-instrumental transfer in the nu-  
720 cleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*,  
721 33(12), 2274–2287. doi: 10.1111/j.1460-9568.2011.07683.x
- 722 Salamone, J. D., & Correa, M. (2012, nov). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3),

- 723 470–485. doi: 10.1016/j.neuron.2012.10.021
- 724 Schultz, W. (2016, feb). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32. doi:  
725 10.1038/nrn.2015.26
- 726 Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992, dec). Neuronal activity in monkey ventral striatum related to the  
727 expectation of reward. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 12(12), 4595–610.  
728 doi: 10.1523/JNEUROSCI.12-12-04595.1992
- 729 Schultz, W., Dayan, P., & Montague, P. R. (1997, mar). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.  
730 doi: 10.1126/science.275.5306.1593
- 731 Sescousse, G., Li, Y., & Dreher, J.-C. (2015, apr). A common currency for the computation of motivational values in the human  
732 striatum. *Social Cognitive and Affective Neuroscience*, 10(4), 467–473. doi: 10.1093/scan/nsu074
- 733 Setlow, B., Schoenbaum, G., & Gallagher, M. (2003, may). Neural encoding in ventral striatum during olfactory discrimination  
734 learning. *Neuron*, 38(4), 625–636. doi: 10.1016/S0896-6273(03)00264-2
- 735 Shidara, M., Aigner, T. G., & Richmond, B. J. (1998, apr). Neuronal signals in the monkey ventral striatum related to progress  
736 through a predictable series of trials. *Journal of Neuroscience*, 18(7), 2613–25. doi: 10.1016/S0896-6273(03)00264-2
- 737 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., & Buzsáki, G. (2017, may). Cocaine place conditioning strengthens  
738 location-specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, 1–10. doi: 10.1101/105890
- 739 Sleezer, B. J., Castagno, M. D., & Hayden, B. Y. (2016, nov). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.  
740 *Journal of Neuroscience*, 36(44), 11223–11237. doi: 10.1523/JNEUROSCI.1766-16.2016
- 741 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., & Hayden, B. Y. (2016, mar). Neuronal selectivity  
742 for spatial positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3), 1098–1111. doi:  
743 10.1152/jn.00325.2015
- 744 Sugam, J. A., Saddoris, M. P., & Carelli, R. M. (2014, may). Nucleus accumbens neurons track behavioral preferences and reward  
745 outcomes during risky decision making. *Biological Psychiatry*, 75(10), 807–816. doi: 10.1016/j.biopsych.2013.09.010
- 746 Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction* (Vol. 9) (No. 5). MIT Press, Cambridge, MA. doi:  
747 10.1109/TNN.1998.712192
- 748 Tabuchi, E. T., Mulder, A. B., & Wiener, S. I. (2000, jan). Position and behavioral modulation of synchronization of hip-  
749 pocampal and accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6), 717–728. doi: 10.1002/1098-  
750 1063(2000)10:6;717::AID-HIPO1009;3.0.CO;2-3
- 751 Takahashi, Y. K., Langdon, A. J., Niv, Y., & Schoenbaum, G. (2016, jul). Temporal Specificity of Reward Prediction Er-  
752 rrors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1), 182–193. doi:  
753 10.1016/j.neuron.2016.05.015
- 754 Tingley, D., & Buzsáki, G. (2018, jun). Transformation of a Spatial Map across the Hippocampal-Lateral Septal Circuit. *Neuron*,  
755 98(6), 1229–1242.e5. doi: 10.1016/J.NEURON.2018.04.028

- 756 van der Meer, M. A. A., & Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward  
757 Information. *Journal of Neuroscience*, 31(8), 2843–2854. doi: 10.1523/JNEUROSCI.4869-10.2011
- 758 West, E. A., & Carelli, R. M. (2016, jan). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after  
759 Reinforcer Devaluation. *Journal of Neuroscience*, 36(4), 1128–1139. doi: 10.1523/JNEUROSCI.2976-15.2016
- 760 Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., & Mulder, A. B. (2003, oct). Spatial and behavioral correlates  
761 in nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,  
762 1250(C), 275–292. doi: 10.1016/S0531-5131(03)00978-6
- 763 Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral  
764 and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, 24(12), 2923–2933. doi:  
765 10.1523/JNEUROSCI.5282-03.2004