

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

Authors: Jimmie M. Gmaz¹, James E. Carmichael¹, Matthijs A. A. van der Meer^{1*}

¹Department of Psychological and Brain Sciences, Dartmouth College, Hanover NH 03755

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

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

¹² Introduction

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

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

³³ Surprisingly, although substantial work has been done on the coding of outcomes predicted by such cues
³⁴ *Schultz et al.* [1992]; *Hollerman and Tremblay* [1998]; *Hassani et al.* [2001]; *Cromwell and Schultz* [2003];
³⁵ *Setlow et al.* [2003]; *Nicola* [2004]; *Roitman et al.* [2005]; *Day et al.* [2006]; *Roesch et al.* [2009]; *Saddoris*

36 *et al.* [2011]; *Goldstein et al.* [2012]; *Lansink et al.* [2012]; *Bissonette et al.* [2013]; *McGinty et al.*
37 [2013]; *Atallah et al.* [2014]; *Sugam et al.* [2014]; *Cooch et al.* [2015]; *West and Carelli* [2016], much
38 less is known about how outcome-predictive cues themselves are encoded in the NAc but see; *Sleezer et al.*
39 [2016]. This is an important issue for at least two reasons. First, in reinforcement learning, motivationally
40 relevant outcomes are typically temporally delayed relative to the cues that predict them. In order to solve
41 the problem of assigning credit (or blame) across such temporal gaps, some trace of preceding activity needs
42 to be maintained *Sutton and Barto* [1998]; *Lee et al.* [2012]. For example, if you become ill after eating
43 food X in restaurant A, depending on if you remember the identity of the restaurant or the food at the time
44 of illness, you may learn to avoid all restaurants, restaurant A only, food X only, or the specific pairing of
45 X-in-A. Therefore, a complete understanding of what is learned following feedback requires understanding
46 what trace is maintained. Since NAc is a primary target of DA signals interpretable as RPEs, and NAc lesions
47 impair RPEs related to timing, its activity trace will help determine what can be learned when RPEs arrive
48 *Ikemoto* [2007]; *McDannald et al.* [2011]; *Hart et al.* [2014]; *Hamid et al.* [2015]; *Takahashi et al.* [2016].
49 Similarly, in an economic framework, in order to maximize utility a rational agent selects the offer with
50 the largest subjective value *Samuelson* [1947]. NAc has been implicated as containing representations of
51 this domain-general subjective value *Peters and Bü* [2009]; *Levy and Glimcher* [2012]; *Bartra et al.* [2013];
52 *Sescousse et al.* [2015], and having representations of an offer alongside its subjective value could be one
53 way in which the two are linked.

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

62 allow for context-dependent mapping of outcomes predicted by specific cues.

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

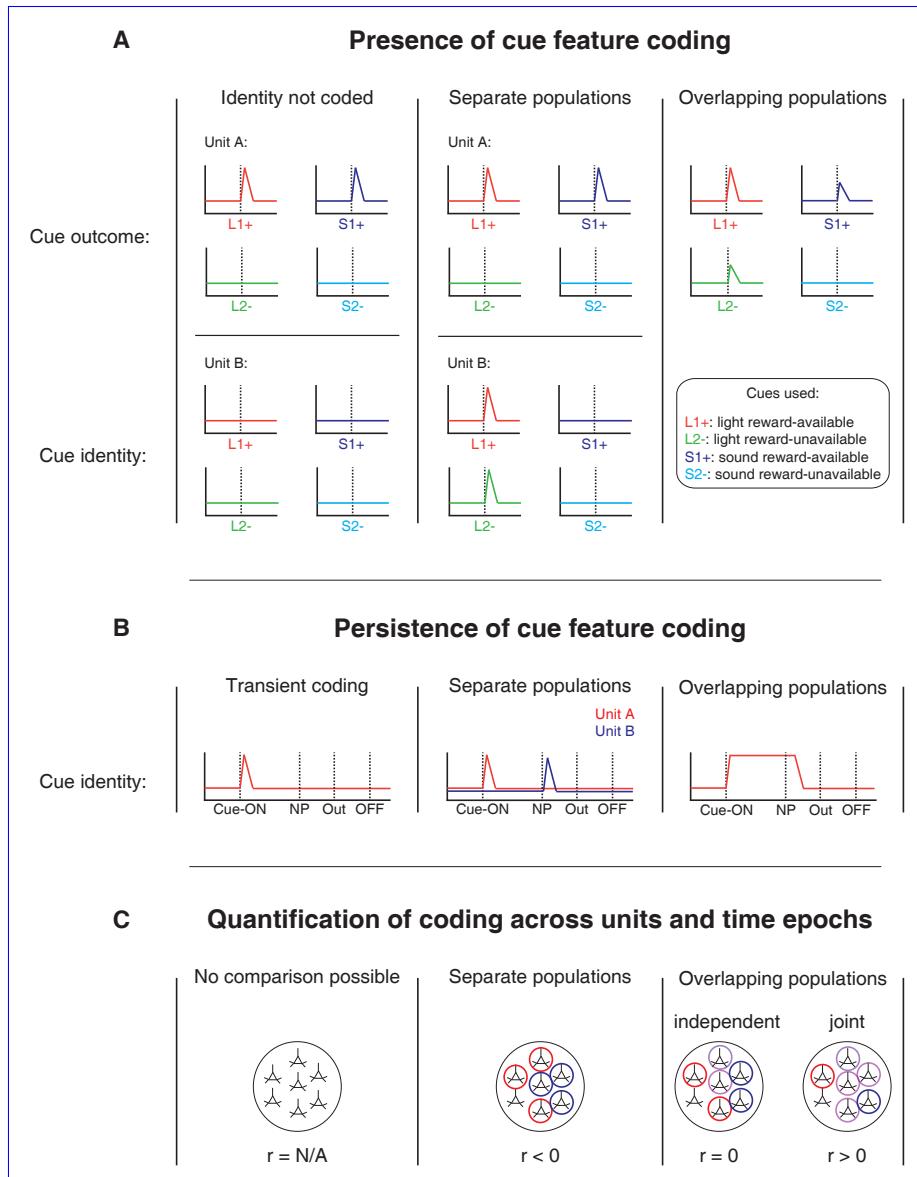


Figure 1: Schematic of potential coding strategies for cue identity (light, sound) and cue outcome (reward-available, reward-unavailable) employed by single units in the NAc across different units (A) and in time (B).

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

- H1 (left Left panel):** Coding of cue identity is absent in the NAc. Top: Unit A encodes a motivationally relevant variable, such as expected outcome, similarly across other cue features, such as cue identity or physical location. Hypothetical plot is firing rate across time. L1+ (red) signifies a reward-available light cue, S1+ (navy blue) a reward-available sound cue, L2- (green) a reward-unavailable light cue, S2- (light blue) a reward-unavailable sound cue. Dashed line indicates onset of cue. Bottom: No units within the NAc discriminate their firing according to cue identity.
- H2 (middle Middle panel):** Coding of cue identity occurs independently in a separate population of units from encoding of motivationally relevant variables such as expected outcome or subsequent vigor. Top: Same as H1 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.
- H3 (right Right panel):** Coding of cue identity is integrated occurs in an overlapping population of cells with coding of other motivationally relevant variables. Hypothetical example demonstrating a unit that responds to outcome-predictive cues, but firing rate is also modulated by cue identity, firing most for the reward-available light cue.

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

- H1 (left Left panel):** Cue-onset triggers a transient response to a unit that codes for cue identity. Dashed lines indicate time of a behavioral or environmental event. 'Cue-ON' signifies onset of cue, 'NP' signifies when the rat holds a nosepoke at a reward receptacle, 'Out' signifies when the outcome is revealed, 'OFF' signifies when the cue turns off.
- H2 (middle Middle and right panel):** Coding of cue identity persists at other time points, shown here during a nosepoke hold period until outcome is revealed.

Figure 1: (Previous page.) Coding can either be maintained by a sequence of units (middle panel) or by the same unit as during cue-onset (right panel). **C:** Schematic pool of NAc units, illustrating the different outcomes for each of the analyses. R values represent the correlation between sets of 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 cue identity and cue outcome are represented by separate populations of units (A: middle panel), or cue identity is represented by distinct units among different points in a trial (B: middle panel). Red circles represents coding for one cue feature or epoch, blue circles for the other cue feature or epoch. Right panel: Cue identity and cue outcome (A: right panel), or cue identity 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 cue identity and cue outcome, or cue-onset and nosepoke coding, is expected by chance and that the two cue features 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.

70 Results

71 Behavior

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

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

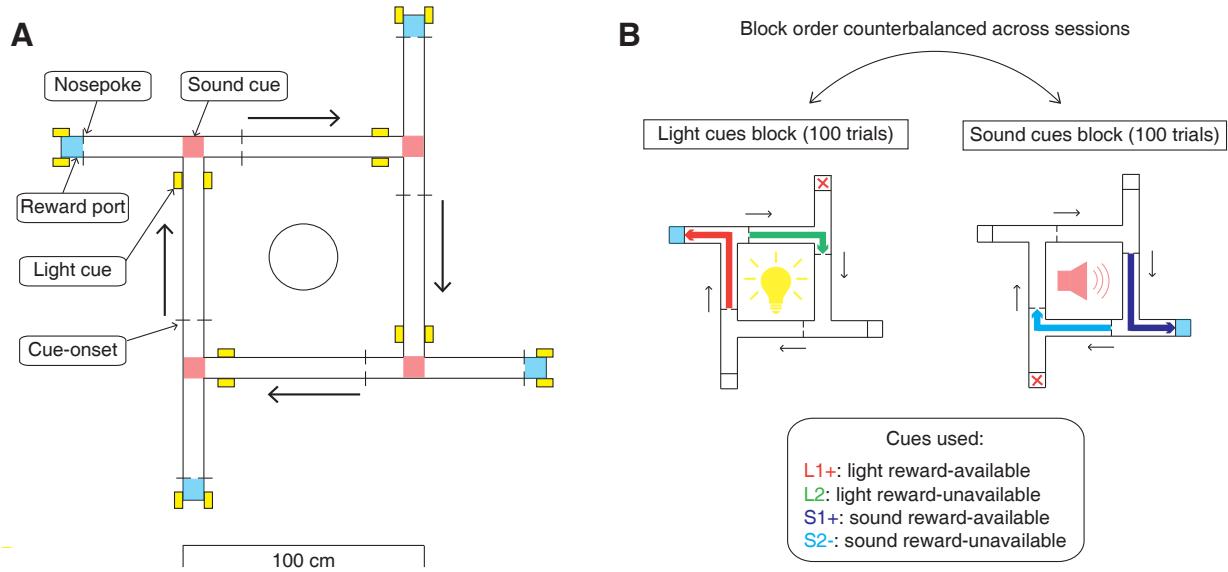


Figure 2: Schematic of behavioral task. **A:** To scale depiction of square track consisting of multiple identical T-choice points. At each choice point, the availability of 12% sucrose reward at the nearest reward receptacle (light blue fill) was signaled by one of four possible cues, presented when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

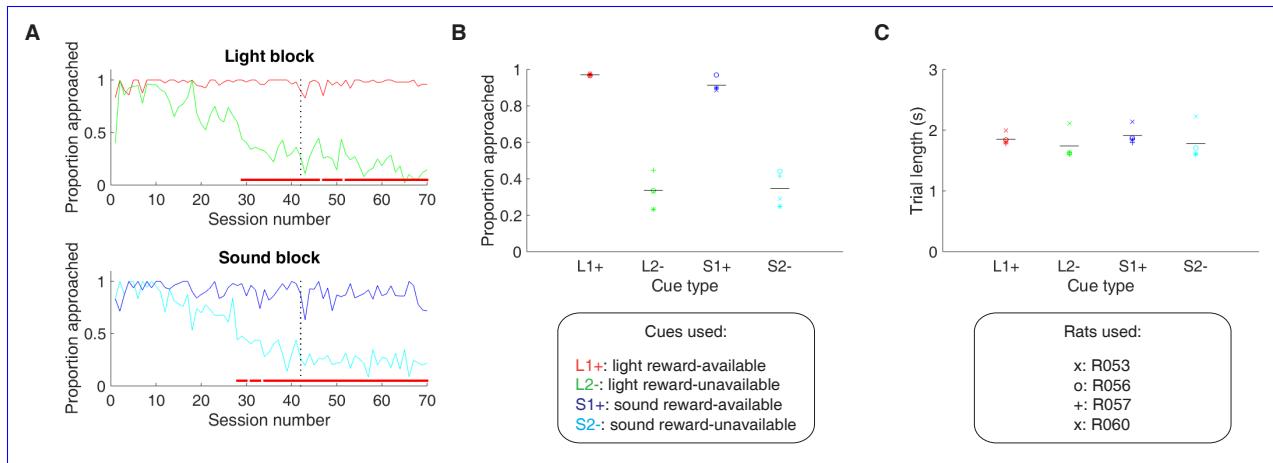


Figure 3: Performance on the behavioral task. **A.** Example learning curves across sessions from a single subject (R060) showing the proportion approached for reward-available (red line for light block, navy blue line for sound block) and reward-unavailable trials (green line for light block, light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

93 NAc units encode behaviorally relevant and irrelevant cue features

94 Single unit responses discriminate between cue features:

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

110 To characterize more formally whether these cue-evoked responses were modulated by various aspects of
111 the task, we fit a GLM-forward selection stepwise generalized linear model (GLM) to each cue-modulated
112 unit. Fitting GLMs to all trials within a session revealed that a variety of task parameters accounted for a
113 significant portion of firing rate variance in NAc cue-modulated units (Figure 5, Table 1). Notably, there
114 were units that discriminated between whether the rat was performing in the light or sound block (28% of
115 cue-modulated units, accounting for 6% of variance on average), which arm the rat was currently on (38%
116 of cue-modulated units, accounting for 6% of variance on average), and whether the rat was engaged in the

common portion of a reward-available or reward-unavailable trial (26% of cue-modulated units, accounting for 4% of variance on average), suggesting that the NAc encodes features of reward-predictive cues ~~separate from~~ in addition to expected outcome (Figure 4A-F). Furthermore, ~~???~~. To determine whether overlap of coding of cue features ~~within units was not~~ was different than expected by chance according to chi-square tests we correlated recoded beta coefficients from the GLMs (Figure 5C). This revealed that cue identity was coded independently from both cue outcome ($r = -.018$; $p = .840$) and cue location ($r = .107$; $p = .220$), as determined by non-significant correlations, but that cue location and cue outcome were encoded jointly ($r = .221$; $p = .011$), as signified by a significant positive correlation. Chi-square tests depicted a similar pattern of results (cue identity and cue location = $\chi^2 = .687$, $p = .407$; cue identity and cue outcome = $\chi^2 = .022$, $p = .882$; cue location and cue outcome = $\chi^2 = 3.02$, $p = .082$). The GLM fitting was repeated for the first half of the trials in each stimulus block. This revealed a qualitatively similar pattern of results, suggesting that this selectivity is present early within a block (first half of block vs. all trials: 37 vs. 37, suggesting for integrated coding across various aspects of a cue (Figure 4G, H45 vs. 50, and 25 vs. 34 units for cue identity, cue location, and cue outcome, respectively). To control for differences that may have arisen due to different behavior at the choice point, we also analyzed our data for approach trials only. Comparison of the GLMs using only approach trials versus all trials was qualitatively similar (approach trials vs. all trials: 33 vs. 37, 45 vs. 50, and 28 vs. 34 units for cue identity, cue location, and cue outcome, respectively). Additionally, a sliding window GLM centered on cue-onset revealed that cue identity and cue location contributed to the activity of a significant proportion of cue-modulated units throughout this epoch, whereas an increase in units encoding cue outcome became apparent after cue-onset (z-score > 1.96 when comparing proportion of units to that observed when shuffling the firing rates before running the GLM, Figure 5D). Together, these findings show that various cue features are represented in the NAc ~~, and that this coding is both integrated and separate from expected outcome in overlapping populations, that cue modality is coded independently from cue outcome and cue location, and that cue outcome and cue location are coded jointly~~ (Figure 1; H2,H3).

Task parameter	Total	\uparrow MSN	\downarrow MSN	\uparrow FSI	\downarrow FSI
All units	443	155	216	27	45
<i>Rat ID</i>					
R053	145	51	79	4	11
R056	70	12	13	17	28
R057	136	55	75	3	3
R060	92	37	49	3	3
Analyzed units	344	117	175	18	34
Cue modulated units	133	24	85	6	18
<i>GLM aligned to cue-onset</i>					
Cue identity	37 (28%)	7 (29%)	21 (25%)	1 (17%)	8 (44%)
Cue location	50 (38%)	13 (54%)	27 (32%)	3 (50%)	7 (39%)
Cue outcome	34 (26%)	10 (42%)	18 (21%)	0 (-)	6 (33%)
Approach behavior	31 (23%)	8 (33%)	18 (21%)	1 (17%)	4 (22%)
Trial length	25 (19%)	5 (21%)	18 (21%)	0 (-)	2 (11%)
Trial number	32 (24%)	11 (46%)	12 (14%)	1 (17%)	8 (44%)
Previous trial	5 (4%)	0 (-)	5 (6%)	0 (-)	0 (-)
<i>GLM aligned to nosepoke</i>					
Cue identity	66 (50%)	14 (58%)	36 (42%)	2 (33%)	14 (78%)
Cue location	66 (50%)	14 (58%)	40 (47%)	3 (50%)	9 (50%)
Cue outcome	42 (32%)	8 (33%)	29 (34%)	0 (-)	5 (28%)
<i>GLM aligned to outcome</i>					
Cue outcome	10 (8%)	0 (-)	6 (7%)	0 (-)	4 (22%)

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 ($n = 133$).

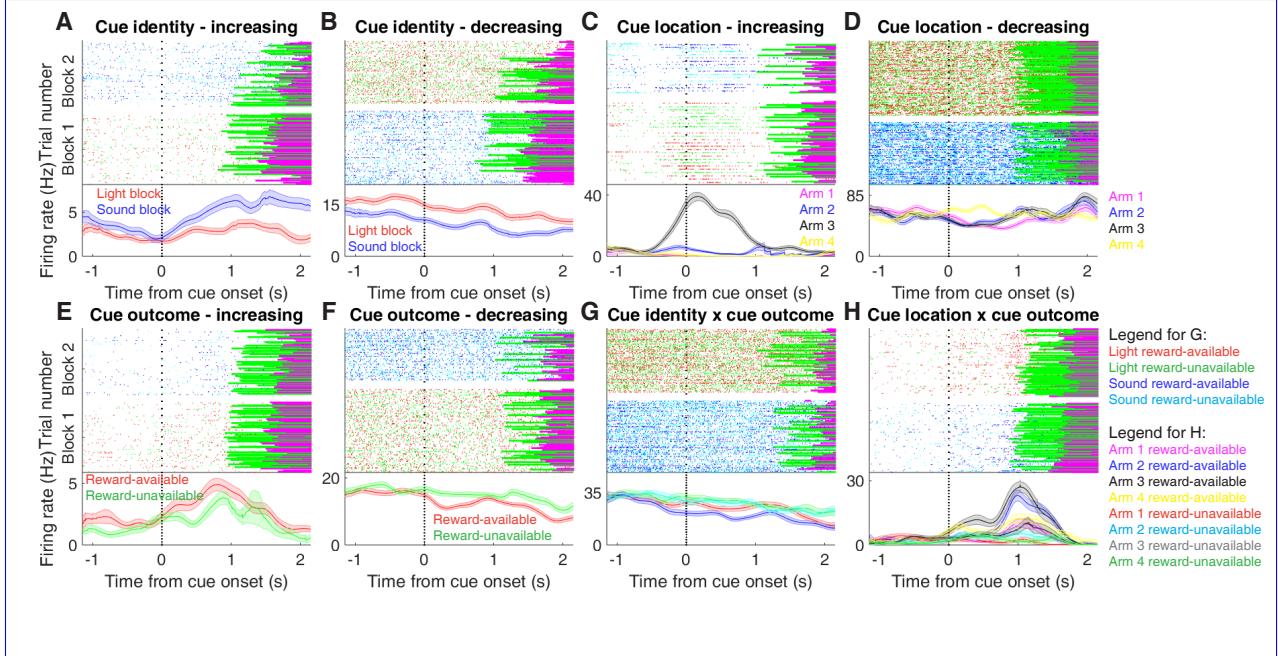


Figure 4: Examples of **different** cue-modulated NAc units influenced by **various** different task parameters. **A:** Example of a cue-modulated NAc unit that showed an increase in firing following the cue, and encoded cue identity. Top: rasterplot showing the spiking activity across all trials aligned to cue-onset. Spikes across trials are color-coded according to cue type (red: reward-available light; green: reward-unavailable light; navy blue: reward-available sound; light blue: reward-unavailable sound). Green and magenta bars indicate trial termination when a rat initiated the next trial or made a nosepoke, respectively. White space halfway up the rasterplot indicates switching from one block to the next. Dashed line indicates cue onset. Bottom: PETHs showing the average smoothed firing rate for the unit for trials during light (red) and sound (blue) blocks, aligned to cue-onset. Lightly shaded area indicates standard error of the mean. Note this unit showed a larger increase in firing to sound cues. **B:** An example of a unit that was responsive to cue identity as in A, but for a unit that showed a decrease in firing to the cue. Note the sustained higher firing rate during the light block. **C-D:** Cue-modulated units that encoded cue location, each color in the PETHs represents average firing response for a different cue location. **C:** The firing rate of this unit only changed on arm 3 of the task. **D:** Firing decreased for this unit on all arms but arm 4. **E-F:** Cue-modulated units that encoded cue outcome, with the PETHs comparing reward-available (red) and reward-unavailable (green) trials. **E:** This unit showed a slightly higher response during presentation of reward-available cues. **F:** This unit showed a dip in firing when presented with reward-available cues. **G-H:** Examples of cue-modulated units that encoded multiple cue features. **G:** This unit integrated cue identity and outcome. **H:** An example of a unit that integrated cue identity and location.

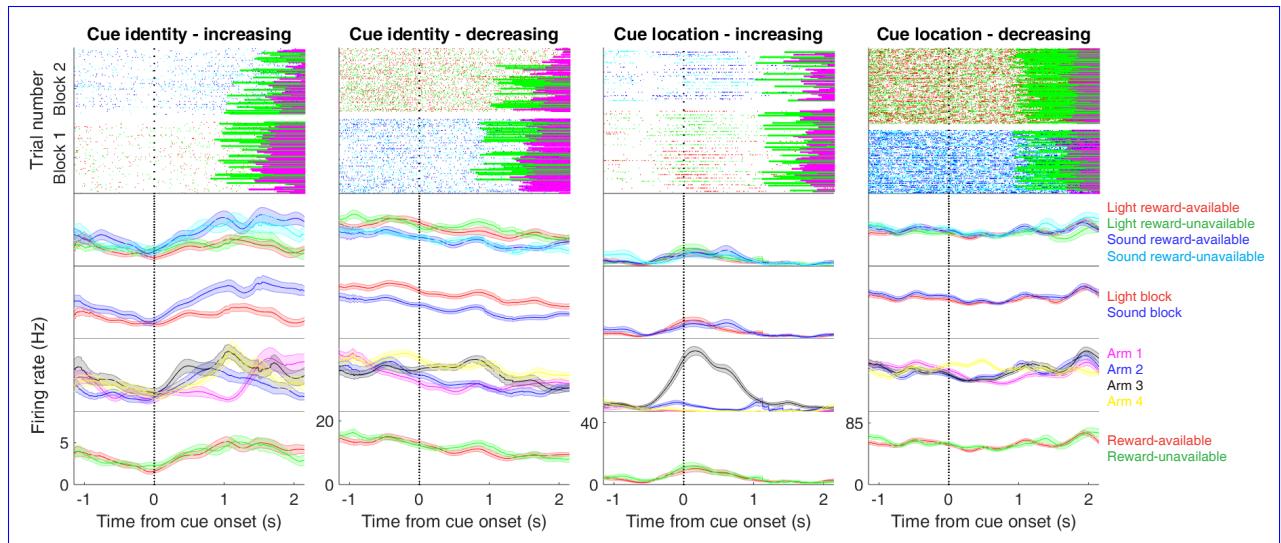


Figure supplement 1 Expanded neural examples.

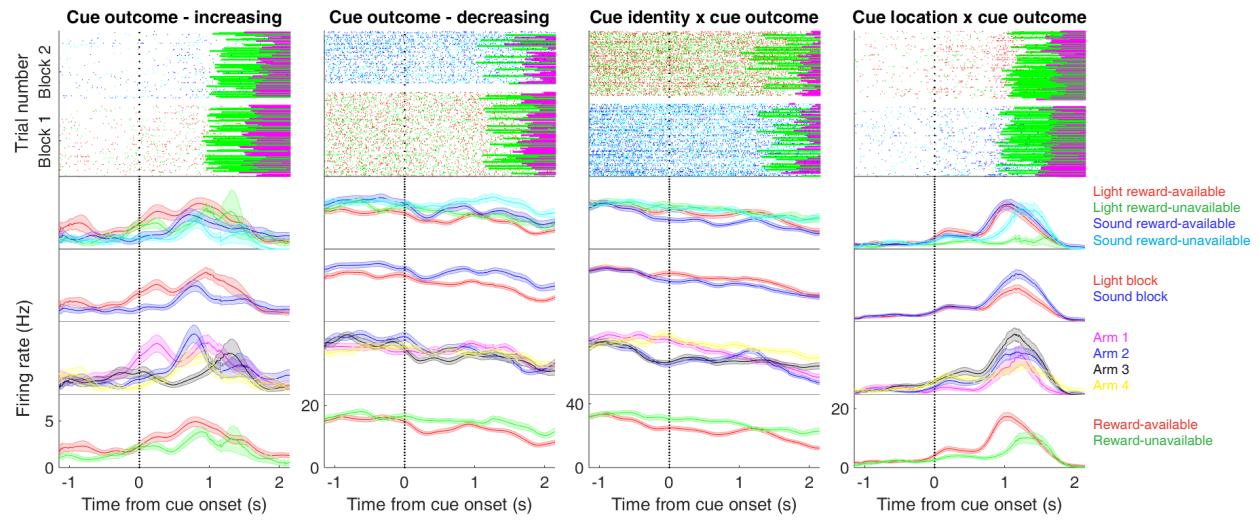


Figure supplement 2 Expanded neural examples.

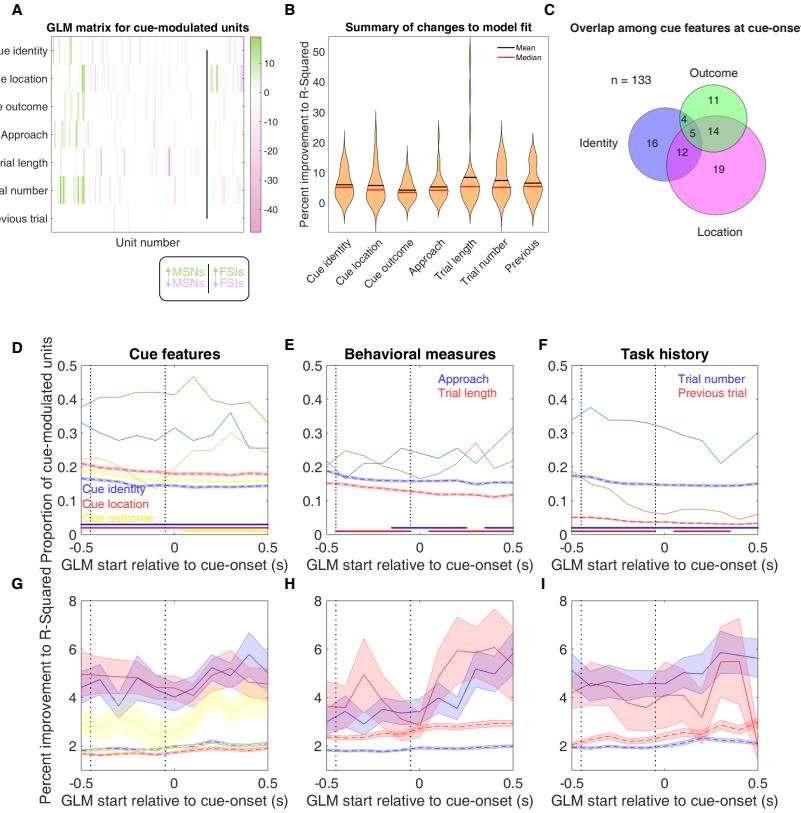


Figure 5: Summary of influence of various task parameters on cue-modulated NAc units after cue-onset. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank test. Each row represents a given task parameter, and each column corresponds to a single unit. Colors indicate how much of the firing rate variance an individual predictor contributed to the model, as measured by differences in R-squared between the final model and the model minus the predictor of interest. Ordering from left to right: MSNs that increased firing in response to the cue (green, left of line), MSNs with a decreasing response (red, left of line), FSIs with an increasing response (green, right of line), FSIs with a decreasing response (red, right of line). Darker shades indicate more firing rate variance explained by a given predictor. Black line indicates separation of MSNs and FSIs. Scale bar indicates range of improvements to model fit for units with an increasing (green) and decreasing (red) response to the cue. **B:** Violin plots demonstrating changes in R-squared values with the addition of each of the individual predictors. The mean, median, and distribution of changes in R-squared values is plotted for each of the seven task parameters used in the GLM. **C:** Venn diagram illustrating the number of cue-modulated units encoding cue identity (blue circle), cue location (green circle), cue outcome (pink circle), as well as the overlap among units that encoded multiple cue features. **D-F:** Sliding window GLM illustrating the proportion of cue-modulated units influenced by various predictors around time of cue-onset. **D:** Sliding window GLM (bin size: 500 ms; step size: 100 ms) demonstrating the proportion of cue-modulated units where cue identity (blue solid line), cue location (red solid line), and cue outcome (yellow solid line) significantly contributed to the model at various time epochs relative to cue-onset. Dashed colored lines indicate the average of shuffling the firing rate order that went into the GLM 100 times. **E:** Same as D, but for approach behavior and trial length. **F:** Same as D, but for trial number and previous trial. **G-I:** Average improvement to model fit. **G:** Average percent improvement to R-squared for units where cue identity, cue location, or cue outcome were significant contributors to the final model for time epochs surrounding cue onset. Shaded area around mean represents the standard error of the mean. **H:** Same as G, but for approach behavior and trial length. **I:** Same G, but for trial number and previous trial.

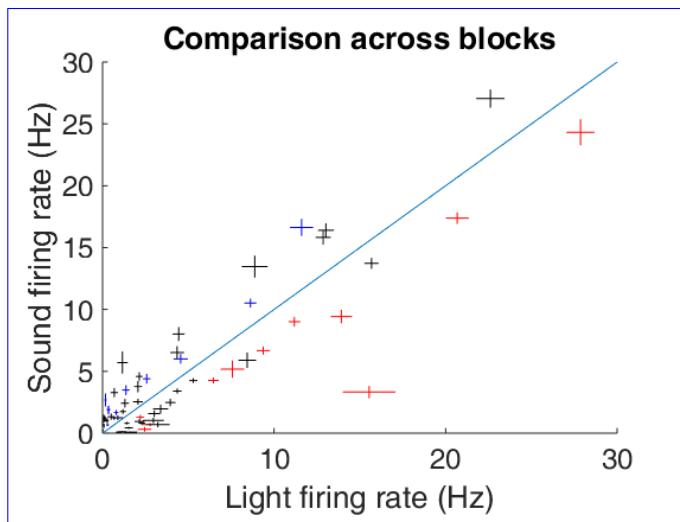


Figure supplement 1: Scatter plot depicting comparison of firing rates for cue-modulated units across light and sound blocks. Crosses are centered on the mean firing rate, range represents the SEM. Colored crosses represents units that had cue identity as a significant predictor of firing rate variance (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.

141 **Population level averages reveal characteristic response profiles**
142 **Pseudoensembles reveal how cue feature
information is encoded at the population level:**

143 To assess what information may be encoded at the population level, we trained a classifier on a pseudoensemble
144 of the 133 cue-modulated units (Figure 6). To do this we used the firing rate for each trial as an observation,
145 with the firing rate of each unit as a variable, and used 10-fold cross validation to train a linear discriminant
146 analysis (LDA) classifier, and see how well we could correctly classify a test trial with the correct label
147 (e.g./ a light or sound trial). This approach revealed that we could correctly predict a trial above chances
148 levels for all cue features for most time epochs (z -score > 1.96 when comparing classification accuracy of
149 data versus shuffling the firing rates). Notably, the ability to predict whether a trial was reward-available or
150 reward-unavailable (cue outcome) was low pre-cue and increased as a trial progressed (Figure 6A).

151 We observed a variety of single unit response profiles around the time of cue onset (Figure 4). To investigate
152 whether these firing rate patterns were related to what cue features were encoded, we plotted the population
153 level averages for units that were modulated by each feature. To do this, we normalized firing activity for
154 each unit that was modulated by a given cue feature, such as light block, then generated the cue-onset aligned
155 population average firing rate for each of the cue features (Figure ??). Overall, this analysis revealed that cells
156 that showed an increase upon cue presentation had stronger responses for the preferred cue condition (Figure
157 ??A,C,E). Interestingly, units that were classified as decreasing in response to the cue showed a biphasic
158 response at the population level, with a small peak at a time in alignment with entry into the arm, followed
159 by a sustained dip after cue-onset (Figure ??B,D,F). Units that were modulated by cue identity showed a
160 stronger increase in response to the preferred task block, as well as a higher tonic firing rate to the preferred
161 task block, most notably in units that decreased in firing rate to the cue (Figure ??A,B). Units that were
162 modulated by cue location showed a graded response to locations of decreasing preference, with peak firing
163 occurring around cue-onset (Figure ??C,D). Units that were modulated by cue outcome showed a ramping
164 of activity after cue-onset for their preferred cue type. Additionally, units that exhibited a decrease in firing
165 in response to the cue and whose activity was modulated by cue outcome, showed a sustained discriminatory

166 response to reward-available and reward-unavailable cues that extended beyond cue-onset (Figure ??F).
 167 Together, these visualizations of the averaged population responses revealed nuanced differences in the way
 168 NAc units are modulated by cue conditions across cue features.

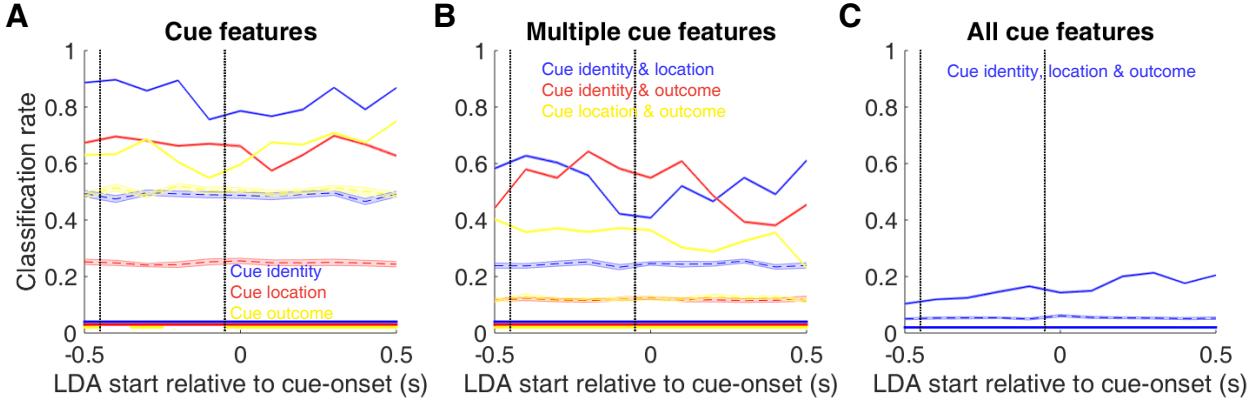


Figure 6: Classification performance of pseudoensembles for cue features. **A:** Sliding window LDA (bin size: 500 ms; step size: 100 ms) demonstrating the classification rate for cue identity (blue solid line), cue location (red solid line), and cue outcome (yellow solid line) using a pseudoensemble 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 25 times. Solid lines at the bottom indicate when the classifier performance greater than the shuffled samples. Points in between the two vertical dashed lines indicate bins where both pre- and post-cue-onset time periods were used in the classifier. **B:** Same as A, but for predicting combinations of cue features, specifically cue identity and cue location (blue), cue identity and cue outcome (red), and cue location and cue outcome (yellow). **C:** Same as A, but for predicting trial types that use all three cue features.

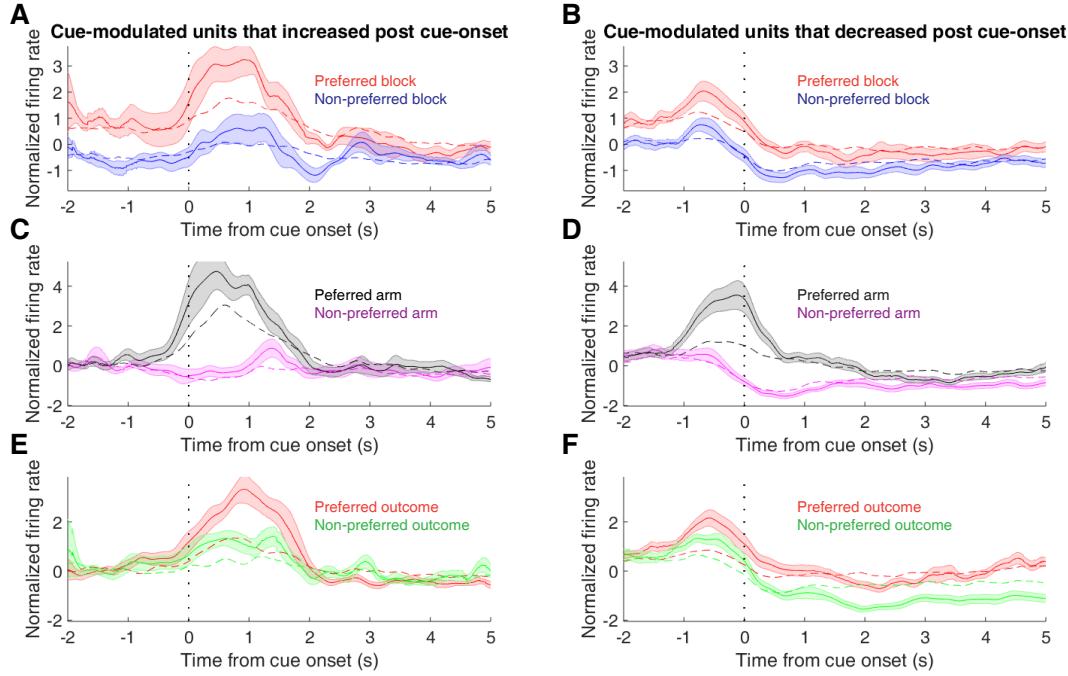


Figure supplement 1: Population-level averages of cue feature sensitive NAc units. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to cue-onset. Activity is plotted for preferred stimulus block (red) and nonpreferred stimulus block (blue). Dashed vertical line indicates onset of cue. Dashed color lines indicate the result of shuffling the identity of the units used for this average 1000 times. Lightly shaded area indicates standard error of the mean. Note larger increase to preferred stimulus block over nonpreferred stimulus block. Black lines indicate the average of 1000 rounds of random sampling of units from the non-drifting population for the preferred and non-preferred blocks. **B:** Same as A but for units that decreased in firing. Note population level activity reveals units classified as decreasing in response to cue show a biphasic response at the population level, with a transient increase around the time the rat starts on the arm, followed by a minimum after cue onset. Also, note the sustained difference in firing between the two blocks. **C-D:** Same as A-B for cue location. Activity is plotted for most preferred arm (black) and least preferred arm (magenta). **E-F:** Same as A-B for cue outcome. Activity is plotted for preferred expected outcome (red), and nonpreferred outcome (green). Note the larger increase to the cue representing the units preferred outcome (E), and the sustained decrease to the nonpreferred outcome (F).

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

170 Given the varied time courses and response profiles of NAc units to various aspects of the cue, the NAc may
171 be computing a temporally evolving state value signal (Pennartz., 2011). Pennartz et al. [2011]. If this
172 is the case, then the recruitment of NAc units should vary alongside changes in the environment. To look
173 at the distribution of responses throughout our task space and see if this distribution is modulated by cue
174 features, we z-scored the firing rate of each unit and plotted the normalized firing rates of all units aligned to
175 cue-onset and sorted them according to the time of peak firing rate (Figure 7). We did this separately for both
176 the light and sound blocks, and found a nearly uniform distribution of firing fields in task space that was not
177 limited to alignment to the cue (Figure 7A). Furthermore, to determine if this population level activity was
178 similar across blocks, we also organized firing during the sound blocks according to the ordering derived
179 from the light blocks. This revealed that while there was some preservation of order, the overall firing was
180 qualitatively different across the two blocks, implying that population activity distinguishes between light
181 and sound blocks. To control for the possibility that any comparison of trials would produce this effect,
182 we did a within block comparison, comparing half of the trials in the light block against the other half.
183 This comparison looked similar to our test comparison of sound block trials ordered by light block trials.
184 Additionally, given that the majority of our units showed an inhibitory response to the cue, we also plotted
185 the firing rates according to the lowest time in firing, and again found some maintenance of order, but largely
186 different ordering across the two blocks, and the within block comparison (Figure 7D). To further test this, we
187 divided each block into two halves and looked at the correlation of the average smoothed firing rates across
188 various combinations of these halves across our cue-aligned centered epoch. A linear mixed effects model
189 revealed that within block correlations (e.g. one half of light trials vs other half of light trials) were higher
190 and more similar than across block correlations (e.g. half of light trials vs half of sound trials) suggesting
191 that activity in the NAc discriminates across various cue conditions (within block correlations = .383 (light),
192 .379 (sound); across block correlations = .343, .338, .337, .348; within block vs. within block comparison =
193 p = .934; within block vs. across block comparisons = p < .001). This process was repeated for cue location
194 (Figure 7B,E; within block correlations = .369 (arm 1), .350 (arm 2); across block correlations = .290, .286,

.285, .291; within block vs. within block comparison = $p = .071$; within block vs. across block comparisons
= $p < .001$) and cue outcome (Figure 7C,F; within block correlations = .429 (reward-available), .261 (reward-
unavailable); across block correlations = .258, .253, .255, .249; within block vs. within block comparison
= $p < .001$; within block vs. across block comparisons = $p < .001$), showing that NAc segmentation of the
task is qualitatively different even during those parts of the task not immediately associated with a specific
cue, action, or outcome, although the within condition comparison of reward-unavailable trials was less
correlated than reward-available trials, and more similar to the across condition comparisons, potentially due
to the greater behavioral variability for the reward-unavailable trials.

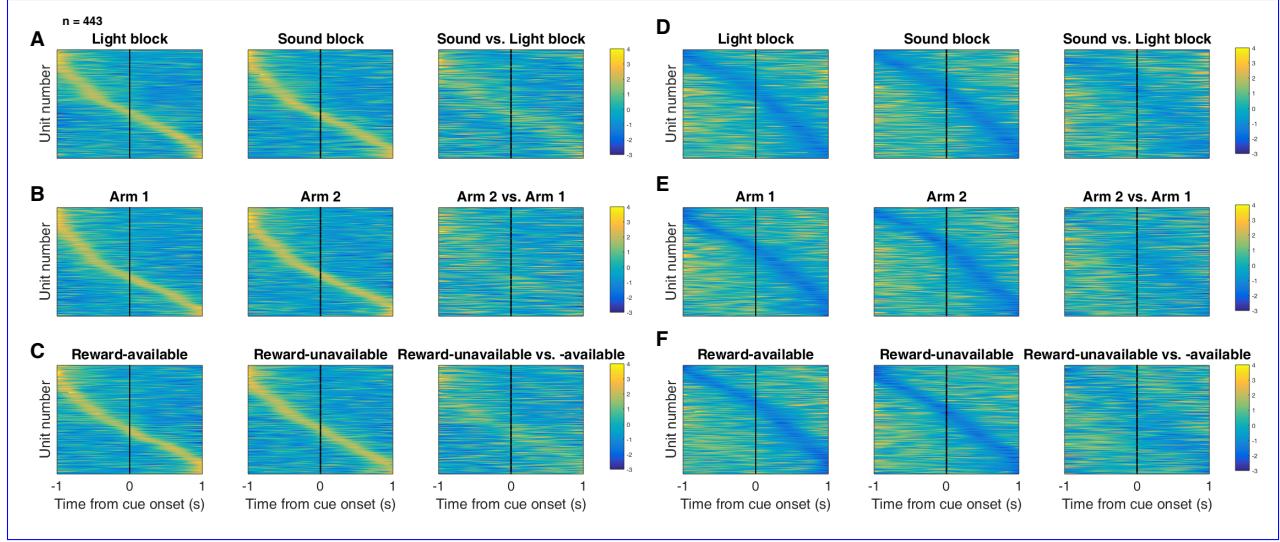


Figure 7: Distribution of NAc firing rates across time surrounding cue onset. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates cue onset), averaged across all trials for a specific cue type, indicated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

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

204 In order to be useful for credit assignment in reinforcement learning, a trace of the cue must be maintained
205 until the outcome, so that information about the outcome can be associated with the outcome-predictive cue.
206 To test whether representations of cue features persisted post-approach until the outcome was revealed, we
207 fit a GLM to the post-approach firing rates of cue-modulated units aligned to the time of nosepoke into the
208 reward receptacle. This analysis showed that a variety of units still discriminated firing according to various
209 cue features, but not other task parameters, showing that NAc activity discriminates various cue conditions
210 well into a trial (Table 1, Figures 8,[??](#),[??](#),9). Additionally, these units were a mix between most of the
211 units that encoded cue features at cue-onset (~~observed overlap greater than expected by chance according to~~
212 ~~chi-square tests~~ $\chi^2 = .386, .781, .758$, across ~~cue-onset and nosepoke for cue identity, cue location, and cue~~
213 ~~outcome, respectively, all p values < .001~~), and those that did not previously have a cue feature as a predictor
214 (29, 48, and 30 out of 133 cue-modulated units encoded both time points for cue identity, cue location, and
215 cue outcome, respectively). [Chi-square tests confirmed these results \(cue identity = \$\chi^2 = 6.165\$, \$p = .013\$; cue](#)
216 [location = \$\chi^2 = 21.670\$, \$p < .001\$; cue outcome = \$\chi^2 = 34.560\$, \$p < .001\$ \)](#). Population level averages for units
217 that increased to cue-onset showed a ramping up of activity that peaked upon nosepoke, whereas units that
218 decreased to cue-onset showed a gradual reduction of firing activity that reached a minimum upon nosepoke
219 (Figure ??). Additionally, a peak is seen for preferred cue outcome in decreasing units at 1 second post
220 cue-onset when reward was received, demonstrating an integration of expected and received reward (Figure
221 ??F). ~~Furthermore, aligning~~

222 [Aligning](#) normalized peak firing rates to nosepoke onset, revealed a clustering of responses around outcome
223 receipt for all cue conditions where the rat would have received reward (Figure ??), in addition to the same
224 trend of higher within- vs across-block correlations for cue identity (Figure ??A,C; within block correlations
225 = .560 (light), .541 (sound); across block correlations = .487, .481, .483, .486; within block vs. within block
226 comparison = $p = .112$; within block vs. across block comparisons = $p < .001$) and cue location (Figure
227 ??B,E; within block correlations = .474 (arm 1), .461 (arm 2); across block correlations = .416, .402, .416,

228 .415; within block vs. within block comparison = $p = .810$; within block vs. across block comparisons = $p <$
229 .001), but not cue outcome (Figure ??C,F; within block correlations = .620 (reward-available), .401 (reward-
230 unavailable); across block correlations = .418, .414, .390, .408; within block vs. within block comparison =
231 $p < .001$; within block vs. across block comparisons = $p < .001$).

232 To determine whether coding of cue features persisted after the outcome was revealed, a GLM was fit to the
233 firing rates of cue-modulated units at the time of outcome receipt, during which the cue was still present.
234 Fitting a GLM revealed 10 units (8%) where cue outcome accounted for an average of 32% of firing rate
235 variance (Table 1, data not shown). An absence of cue identity or cue location coding at this level of analysis
236 was observed, but looking at the data more closely with a sliding window GLM revealed that cue identity,
237 cue location, and cue outcome were encoded throughout time epochs surround cue-onset, nosepoke hold,
238 and outcome receipt, suggesting that the NAc maintains a representation of these cue features once the rat
239 receives behavioral feedback for its decision (Figure 9E-J).

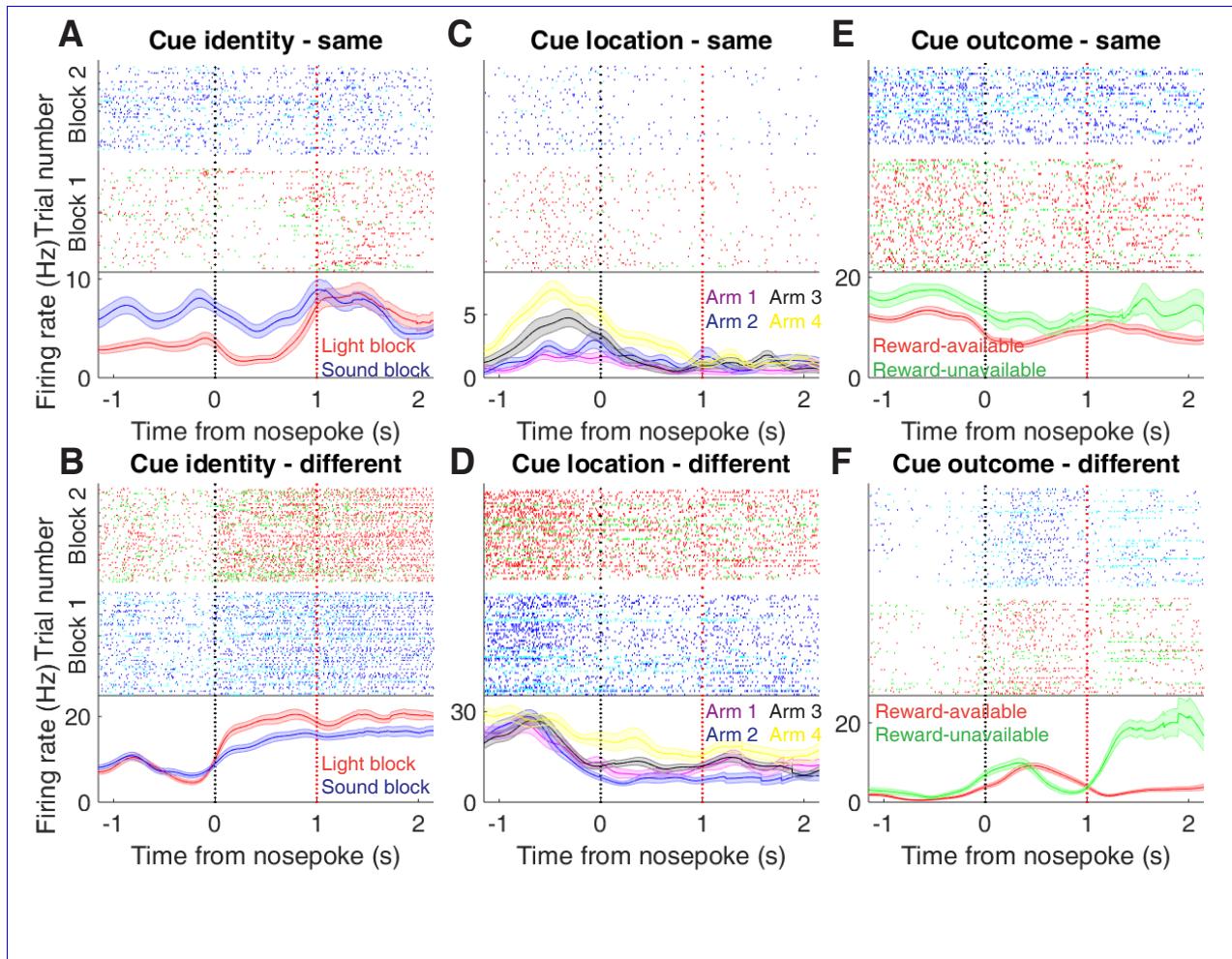


Figure 8: Examples of cue-modulated NAc units influenced by various task parameters at time of nosepoke. **A:** Example of a cue-modulated NAc unit that encoded cue identity at both cue-onset and during nosepoke hold. Top: rasterplot showing the spiking activity across all trials aligned to nosepoke. Spikes across trials are color coded according to cue type (red: reward-available light; green: reward-unavailable light; navy blue: reward-available sound; light blue: reward-unavailable sound). White space halfway up the rasterplot indicates switching from one block to the next. Black dashed line indicates nosepoke. Red dashed line indicates receipt of outcome. Bottom: PETHs showing the average smoothed firing rate for the unit for trials during light (red) and sound (blue) blocks, aligned to nosepoke. Lightly shaded area indicates standard error of the mean. Note this unit showed a sustained increase in firing to sound cues during the trial. **B:** An example of a unit that was responsive to cue identity at time of nosepoke but not cue-onset. **C-D:** Cue-modulated units that encoded cue location, at both cue-onset and nosepoke (C), and only nosepoke (D). Each color in the PETHs represents average firing response for a different cue location. **E-F:** Cue-modulated units that encoded cue outcome, at both cue-onset and nosepoke (E), and only nosepoke (F), with the PETHs comparing reward-available (red) and reward-unavailable (green) trials.

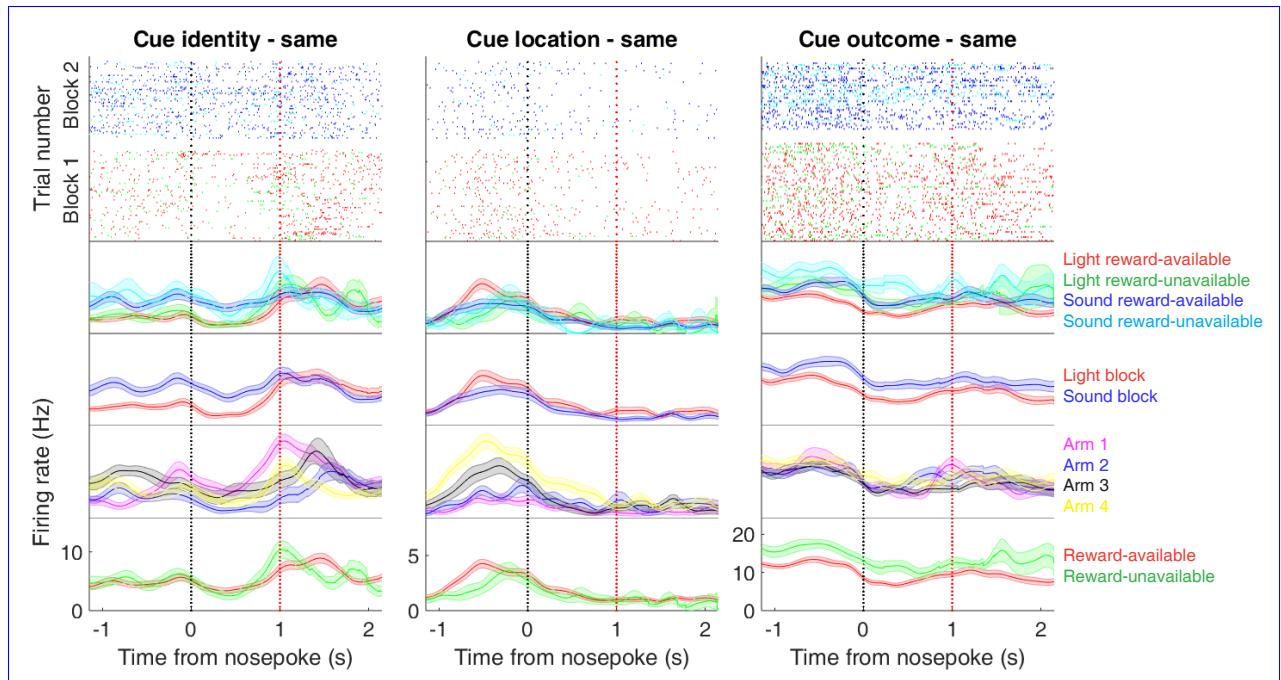


Figure supplement 1 Expanded neural examples.

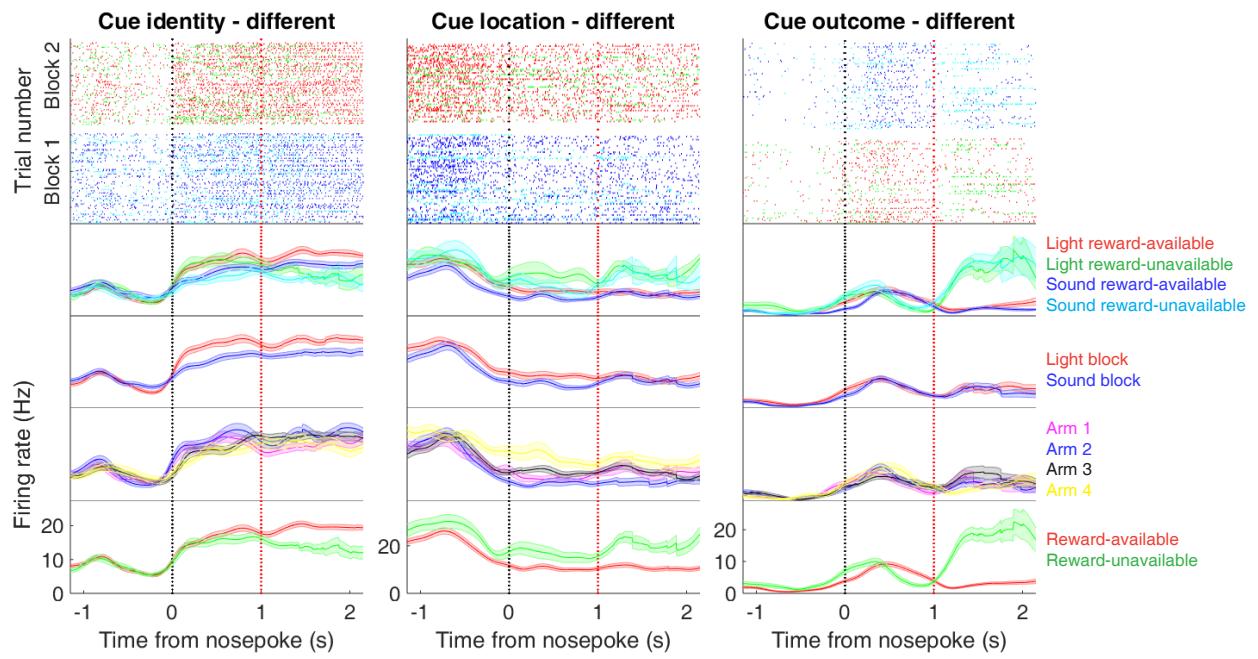


Figure supplement 2 Expanded neural examples.

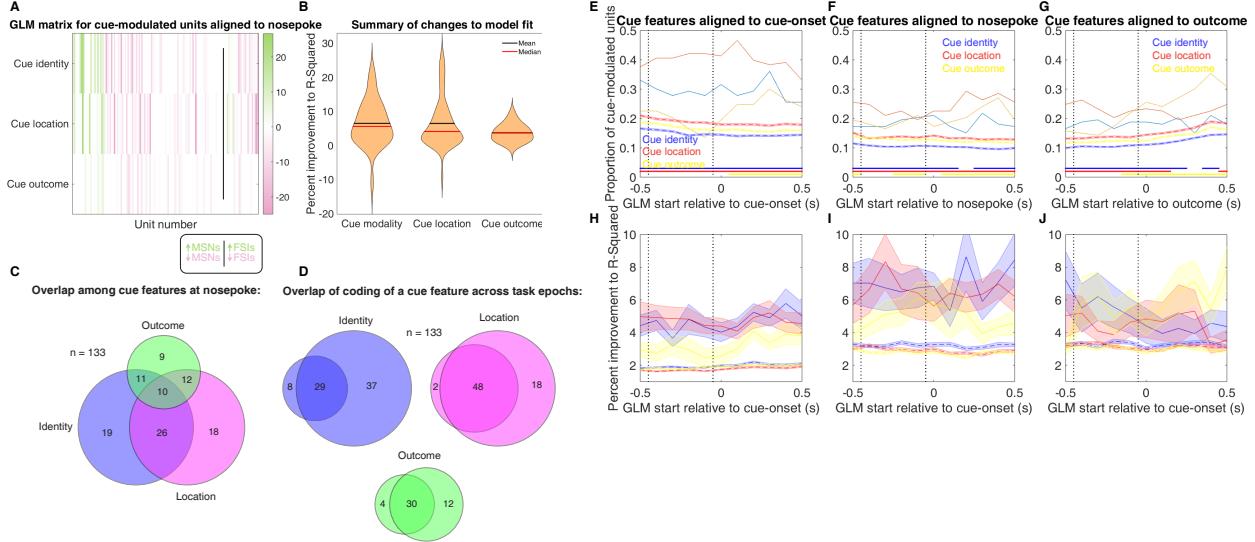


Figure 9: Summary of influence of various task parameters on cue-modulated NAc units during nosepoke. **A:** GLM matrix illustrating the contribution of various task parameters to NAc unit firing rates. A stepwise GLM was fit to each unit that showed evidence of cue modulation by a Wilcoxon signed-rank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

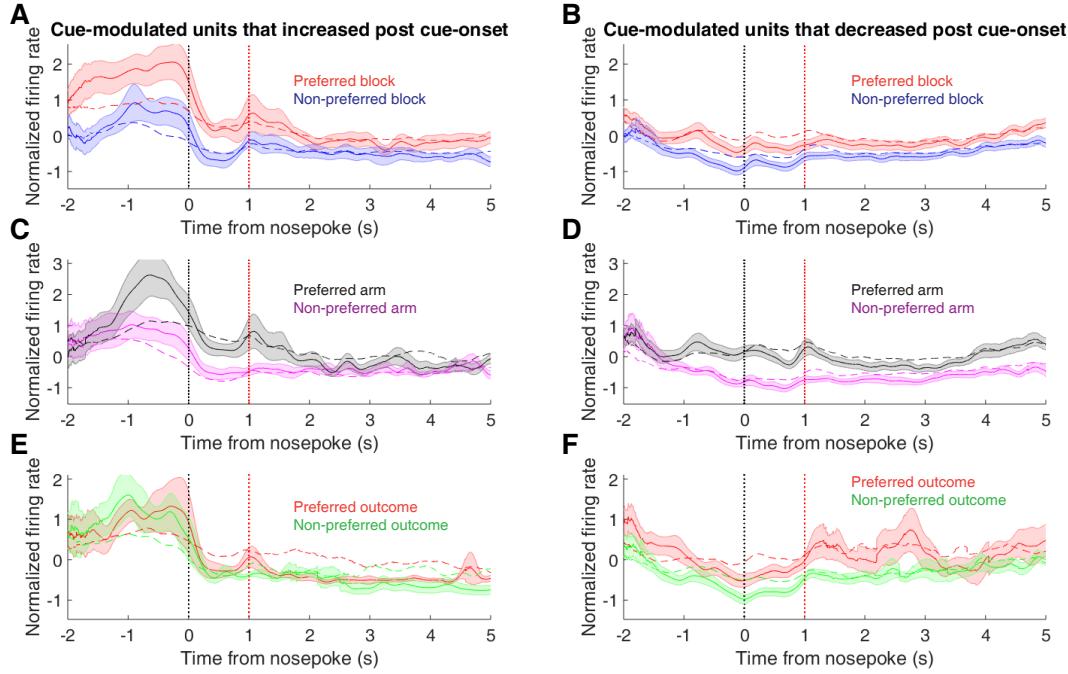


Figure supplement 1: Population-level averages of cue feature sensitive NAc units during a nosepoke. **A:** Average smoothed normalized (z-score) activity for cue-modulated units where cue identity was a significant predictor in the GLM, aligned to nosepoke with reward delivery occurring 1 s after nosepoke.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

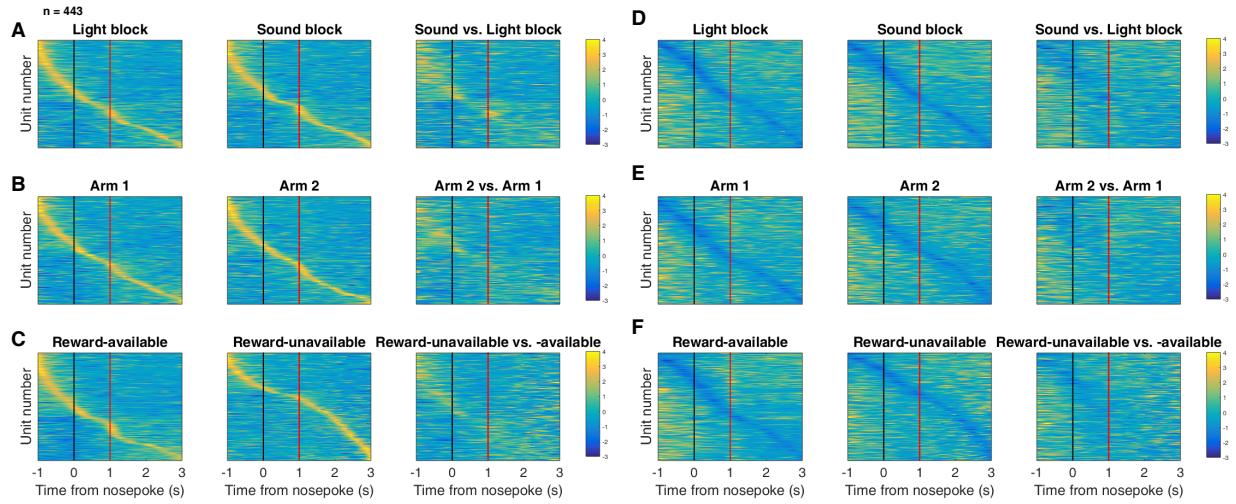


Figure supplement 2: Distribution of NAc firing rates across time surrounding nosepoke for approach trials. Each panel shows normalized (z-score) firing rates for all recorded NAc units (each row corresponds to one unit) as a function of time (time 0 indicates nosepoke), averaged across all approach 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.

240 **Discussion**

241 The main result of the present study is that NAc units encode not only the expected outcome of outcome-
242 predictive cues, but also the identity of such cues (Figure 1A). Importantly, this identity coding was main-
243 tained on approach trials both during a delay period where the rat held a nosepoke until the outcome was
244 received, ~~after and~~ immediately after outcome receipt (H2 and H3 in Figure 1B). ~~Units coding~~ The population
245 ~~of units that coded~~ for cue identity ~~showed partial overlap with those~~ was statistically independent from the
246 population coding for expected outcome (H3 in Figure 1A), i.e. overlap as expected from chance, Figure 1C).
247 This information was also present at the population level, with a higher than chance classification accuracy
248 for predicting the right cue condition using a pseudoensemble. Units that coded different cue features (iden-
249 tity, outcome, location) exhibited different temporal profiles as a whole, although across all recorded units a
250 tiling of task structure was observed such that all points within our analyzed task space was accounted for by
251 the ordered peak firing rates of all units. Furthermore, this tiling differed between various conditions with a
252 cue feature, such as light versus sound blocks. We discuss these observations and their implications below.

253 **Cue identity:**

254 Our finding that NAc units can discriminate between different outcome-predictive stimuli with similar moti-
255 vational significance (i.e. encodes cue identity) expands upon an extensive rodent literature examining NAc
256 correlates of conditioned stimuli *Setlow et al.* [2003]; *Nicola* [2004]; *Yun et al.* [2004]; *Roitman et al.*
257 [2005]; *Day et al.* [2006]; *Ambroggi et al.* [2008]; *Ishikawa et al.* [2008]; *Roesch et al.* [2009]; *Sad-*
258 *doris et al.* [2011]; *Goldstein et al.* [2012]; *Lansink et al.* [2012]; *Bissonnette et al.* [2013]; *McGinty et al.*
259 [2013]; *Atallah et al.* [2014]; *Sugam et al.* [2014]; *Cooch et al.* [2015]; *West and Carelli* [2016]; *Dejean*
260 *et al.* [2017]. Perhaps the most comparable work in rodents comes from a study that found distinct cod-
261 ing for an odor when it predicted separate but equally valued rewards *Cooch et al.* [2015]. The present
262 work is complementary to such *outcome identity* coding as it shows that NAc units encode *cue identity*, ~~both~~
263 ~~separately and integrated with~~ ~~in addition to~~ the reward it predicts (H2 and H3 in Figure 1A). Similarly,

264 Setlow, Schoenbaum, & Gallagher (2003) paired distinct odor cues with appetitive or aversive outcomes,
265 and found separate populations of units that encoded each cue. Furthermore, during a reversal they found
266 that the majority of units switched their selectivity, arguing that the NAc units were tracking the motivational
267 significance of these stimuli. Once again, our study was different in asking how distinct cues encoding the
268 same anticipated outcome are encoded. Such cue identity encoding suggests that even when the biological
269 relevance motivational significance of these stimuli is similaridentical, NAc dissociates their representations
270 at the level of the single-units. A possible interpretation of this coding of cue features alongside expected
271 outcome is that these representations are used to associate reward with relevant features of the environment,
272 so-called credit assignment in the reinforcement learning literature *Sutton and Barto* [1998]. A burgeoning
273 body of human and non-human primate work has started to elucidated neural correlates of credit assignment
274 in the PFC, particularly in the lateral orbitofrontal cortex *Chau and Sallet* [2015]; *Akaishi et al.* [2016];
275 *Asaad et al.* [2017]; *Noonan et al.* [2017]. Given the importance of cortical inputs in NAc associative
276 representations, it is possible that information related to credit assignment is relayed from the cortex to NAc
277 *Ishikawa et al.* [2008]; *Cooch et al.* [2015].

278 Viewed within the neuroeconomic framework of decision making, functional magnetic resonance imaging
279 (fMRI) studies have found support for NAc representations of *offer value*, a domain-general common currency
280 signal that enables comparison of different attributes such as reward identity, effort, and temporal proximity
281 *Peters and Bü* [2009]; *Levy and Glimcher* [2012]; *Bartra et al.* [2013]; *Sescousse et al.* [2015]. Our study
282 adds to a growing body of electrophysiological research that suggests the view of the NAc as a value center,
283 while informative and capturing major aspects of NAc processing, neglects additional contributions of NAc
284 to learning and decision making such as the offer (cue) identity signal reported here.

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

289 cue identities. Indeed, Sleezer, Castagno, & Hayden (2016) recorded NAc responses during the Wisconsin
290 Card Sorting Task, a common set-shifting task used in both the laboratory and clinic, and found units that
291 preferred firing to stimuli when a certain rule, or rule category was currently active. Further support for a
292 modulation of NAc responses by strategy comes from an fMRI study that examined BOLD levels during
293 a set-shifting task *FitzGerald et al.* [2014]. In this task, participants learned two sets of stimulus-outcome
294 contingencies, a visual set and an auditory set. During testing they were presented with both simultaneously,
295 and the stimulus dimension that was relevant was periodically shifted between the two. Here, they found
296 that bilateral NAc activity reflected value representations for the currently relevant stimulus dimension, and
297 not the irrelevant stimulus. The current finding of separate, but overlapping, populations of units encoding
298 independent coding of cue identity and expected outcome, suggests that the fMRI finding is generated by the
299 combined activity of several different functional cell types.

300 Our analyses were designed to eliminate several potential alternative interpretations to cue identity coding.
301 Because the different cues were separated into different blocks, units that discriminated between cue identi-
302 ties could instead be encoding time or other slowly-changing quantities. We excluded this possible confound
303 by excluding units that showed a drift in firing between the first and second half within a block. However, the
304 possibility remains that instead of or in addition to stimulus identity, these units encode a preferred context,
305 or even a macroscale representation of progress through the session. Indeed, encoding of the current strategy
306 could be an explanation for the sustained difference in population averaged firing across stimulus blocks
307 (Figure ??), as well as a potential explanation for the differentially tiling of task structure across blocks in
308 the current study (Figure 7).

309 A different potential confound is that between outcome and action value coding We discriminated between
310 these possibilities by analyzing error trials, where the rat approached reward as the rat was only rewarded for
311 left turns. Our GLM analysis dealt with this by excluding firing rate variance accounted for by a predictor
312 that represented whether the animal approached (left turn) after presentation of the reward-unavailable cue.
313 Units or skipped (right turn) the reward port at the choice point, thus we were able to identify units that

314 were modulated by the expected outcome of the cue ~~maintained their specific firing patterns even during~~
315 ~~error trials, as expected from outcome value coding but not after removing variance for potential~~ action
316 value coding. Additionally, NAc signals have been shown to be modulated by response vigor *McGinty et al.*
317 [2013]; to detangle this from our results we included trial length (i.e. latency to arrival at the reward site) as
318 a predictor in our GLMs, and found units with cue feature correlates independent of trial length.

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

330 **Encoding of position:**

331 Our finding that cue-evoked activity was modulated by cue location is in alignment with several previous
332 reports *Lavoie and Mizumori* [1994]; *Wiener et al.* [2003]; *Mulder et al.* [2005]; *Strait et al.* [2016]. The
333 NAc receives inputs from the hippocampus, and the communication of place-reward information across the
334 two structures suggests that the NAc tracks locations associated with reward *Tabuchi et al.* [2000]; *Pennartz*
335 [2004]; *Lansink et al.* [2008, 2009]; *van der Meer and Redish* [2011]; *Lansink et al.* [2016]; *Sjulson*
336 *et al.* [2017]. However, it is notable that in our task, location is explicitly uninformative about reward, yet
337 coding of this uninformative variable persists. The tiling of task space is in alignment with previous studies

338 showing that NAc units can also signal progress through a sequence of cues and/or actions *Shidara et al.*
339 [1998]; *Mulder et al.* [2004]; *Khamassi et al.* [2008]; *Berke et al.* [2009]; *Lansink et al.* [2012]; *Atallah*
340 *et al.* [2014], is similar to observations in the basal forebrain *Tingley and Buzsá* [2018], and may represent a
341 temporally evolved state value signal *Pennartz et al.* [2011]; *Hamid et al.* [2015]. Given that the current task
342 was pseudo-random, it is possible that the rats learned the structure of sequential cue presentation, and the
343 neural activity could reflect this. However, this is unlikely as including a previous trial variable in the analysis
344 did not explain a significant amount of firing rate variance in response to the cue for the vast majority of units.
345 In any case, NAc units on the present task continued to distinguish between different locations, even though
346 location, and progress through a sequence, were explicitly irrelevant in predicting reward. We speculate that
347 this persistent coding of location in NAc may represent a bias in credit assignment, and associated tendency
348 for rodents to associate motivationally relevant events with the locations where they occur.

349 **Implications:**

350 Maladaptive decision making, as occurs in schizophrenia, addiction, Parkinson's, among others, can result
351 from dysfunctional RPE and value signals *Frank et al.* [2004]; *Gradin et al.* [2011]; *Maia and Frank* [2011].
352 This view has been successful in explaining both positive and negative symptoms in schizophrenia, and
353 deficits in learning from feedback in Parkinson's *Frank et al.* [2004]; *Gradin et al.* [2011]. However, the
354 effects of RPE and value updating are contingent upon encoding of preceding action and cue features, the
355 eligibility trace *Sutton and Barto* [1998]; *Lee et al.* [2012]. Value updates can only be performed on these
356 aspects of preceding experience that are encoded when the update occurs. Therefore, maladaptive learning
357 and decision making can result from not only aberrant RPEs but also from altered cue feature encoding. For
358 instance, on this task the environmental stimulus that signaled the availability of reward was conveyed by
359 two distinct cues that were presented in four locations. While in our current study, the location and identity
360 of the cue did not require any adjustments in the animals behavior, we found coding of these features along-
361 side the expected outcome of the cue that could be the outcome of credit assignment computations computed
362 upstream. Identifying neural coding related to an aspect of credit assignment is important as inappropriate

363 credit assignment could be a contributor to conditioned fear overgeneralization seen in disorders with patho-
364 logical anxiety such as generalized anxiety disorder, post traumatic stress disorder, and obsessive-compulsive
365 disorder *Kaczkurkin and Lissek* [2013]; *Lissek et al.* [2014]; *Kaczkurkin et al.* [2017], and delusions ob-
366 served in disorders such as schizophrenia, Alzheimer's and Parkinson's *Kapur* [2003]; *Corlett et al.* [2010].
367 Thus, our results provide a neural window into the process of credit assignment, such that the extent and spe-
368 cific manner in which this process fails in syndromes such as schizophrenia, obsessive-compulsive disorder,
369 etc. can be experimentally accessed.

370 Methods

371 Subjects:

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

379 Overall timeline:

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

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

387 **Behavioral task and training:**

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

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

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

420 **Surgery:**

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

429 **Data acquisition and preprocessing:**

430 After recovery, rats were placed back on the task for recording. NAc signals were acquired at 20 kHz with a
431 RHA2132 v0810 preamplifier (Intan) and a KJE-1001/KJD-1000 data acquisition system (Amplipex). Sig-
432 nals were referenced against a tetrode placed in the corpus callosum above the NAc. Candidate spikes for
433 sorting into putative single units were obtained by band-pass filtering the data between 600-9000 Hz, thresh-
434 olding and aligning the peaks UltraMegaSort2k, *Hill et al.* [2011]. Spike waveforms were then clustered
435 with KlustaKwik using energy and the first derivative of energy as features, and manually sorted into units
436 (MClust 3.5, A.D. Redish et al., <http://redishlab.neuroscience.umn.edu/MClust/MClust.html>). Isolated units
437 containing a minimum of 200 spikes within a session were included for subsequent analysis. Units were clas-
438 sified as fast spiking interneurons (FSIs) by an absence of interspike intervals (ISIs) > 2 s, while medium
439 spiny neurons (MSNs) had a combination of ISIs > 2 s and phasic activity with shorter ISIs *Barnes et al.*
440 [2005]; *Atallah et al.* [2014].

441 **Data analysis:**

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

449 *Neural data.* ~~To Given that some of our analyses compare firing rates across time, particularly comparisons~~
450 ~~across blocks, we sought to exclude units with unstable firing rates that would generate spurious results~~
451 ~~reflecting a drift in firing rate over time unrelated to our task. To do this we ran a Mann-Whitney U test~~
452 ~~comparing the cue-evoked firing rates for the first and second half of trials within a block, and excluded~~
453 ~~99 of 443 units from analysis that showed a significant change for either block, leaving 344 units for~~

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

461 To determine the relative contribution of different task parameters to firing rate variance (as in Figures 4-5), a
462 forward selection stepwise general linear model (GLM) GLM using a Poisson distribution for the response
463 variable was fit to each cue-modulated unit, using data from every trial in a session. Cue identity (light
464 block, sound block), cue location (arm 1, arm 2, arm 3, arm 4), cue outcome (reward-available, reward-
465 unavailable), behavior (approach, skip), trial length, trial number, and trial history (reward availability on the
466 previous 2 trials) were used as predictors, and the 1 s post-cue firing rate as the response variable. Units were
467 classified as being modulated by a given task parameter if addition of the parameter significantly improved
468 model fit using deviance as the criterion ($p < .01$). A comparison of the R-squared value between the final
469 model and the final model minus the predictor of interest was used to determine the amount of firing rate
470 variance explained by the addition of that predictor for a given unit. To determine the degree to which
471 coding of cue identity, cue location, and cue outcome overlapped within units we computed pearsons r on
472 recoded beta coefficients from the GLM for the cue features. Specifically, for all cue-modulated units, we
473 coded a 1 if the cue feature was a significant predictor in the final model, and 0 if it was not, and calculated
474 pearsons r comparing an array of the coded 0s and 1s for cue-modulated units for one cue feature with a
475 similar array for another cue feature. The NAc was determined as coding a pair of cue features in a) a
476 separate populations of units if there was a significant negative correlation ($r < 0$), b) an independently
477 coded overlapping population of units if there was no significant correlation ($r = 0$), and c) a jointly coded
478 overlapping population of units if there was a significant positive correlation ($r > 0$). To investigate more
479 finely the temporal dynamics of the influence of task parameters to unit activity, we then fit a sliding window

480 GLM with the same task parameters using 500 ms bins and 100 ms steps, starting 500 ms before cue-onset,
481 up to 500 ms after cue-onset, and measured the proportion of units and average R-squared value for a given
482 time bin where a particular predictor contributed significantly to the final model. To control for the amount of
483 units that would be affected by a predictor by chance, we shuffled the trial order of firing rates for a particular
484 unit within a time bin, and took the average of this value over 100 shuffles. We then calculated how many
485 z-scores the observed proportion was from the mean of the shuffles samples, using a z-score of greater than
486 1.96 as a marker of significance.

487 To get a sense for the predictive power of these cue feature representations we trained a classifier using firing
488 rates from a pseudoensemble of our 133 cue-modulated units (Figure 6). We created a matrix of firing rates
489 for each time epoch surrounding cue-onset where each row was an observation representing the firing rate
490 for a trial, and each column was a variable representing the firing rate for a given unit. Trial labels, or classes,
491 were each condition for a cue feature (e.g. light and sound for cue identity) when trying to predict each cue
492 feature, and combinations of conditions across cue features when trying to make more specific predictions
493 (e.g. light reward-available, light reward-unavailable, sound reward-available, sound reward-unavailable),
494 making sure to align trial labels across units. We then ran LDA on these matrices, using 10-fold cross
495 validation to train the classifier on 90% of the trials and testing its predictions on the held out 10% of trials,
496 and repeated this approach to get the classification accuracy for 100 iterations. To test if the classification
497 accuracy was greater by chance, we shuffled the order of firing rates for each unit before we trained the
498 classifier. We repeated this for 25 shuffled matrices for each time point, and calculated how many z-scores
499 the mean classification rate of the observed data was from the mean of the shuffled samples, using a z-score
500 of 1.96 as the cut off for a significant deviation from the permuted distribution.

501 To better visualize responses to cues and enable subsequent population level analyses (as in Figures 4, ??,
502 and 7), spike trains were convolved with a Gaussian kernel ($\sigma = 100$ ms), and peri-event time histograms
503 (PETHs) were generated by taking the average of the convolved spike trains across all trials for a given task
504 condition. For analysis of population-level responses for cue features (Figure ??), convolved spike trains

505 for all units where cue identity, cue location, or cue outcome explained a significant portion of firing rate
506 variance were z-scored. Within a given cue feature, normalized spike trains were then separated according
507 to the preferred and non-preferred cue condition (e.g. light vs. sound block), and averaged across units to
508 generate population-level averages. To account for separation that would result from any random selection
509 of units, unit identity was shuffled and the shuffled average for preferred and non-preferred cue conditions
510 was generated for 1000 shuffles.

511 To visualize NAc representations of task space within cue conditions, normalized spike trains for all units
512 were ordered by the location of their maximum or minimum firing rate for a specified cue condition (Figure
513 7). To compare representations of task space across cue conditions for a cue feature, the ordering of units
514 derived for one condition (e.g. light block) was then applied to the normalized spike trains for the other
515 condition (e.g. sound block). For control comparisons within cue conditions, half of the trials for a condition
516 were compared against the other half. To look at the correlation of firing rates of all units within and across
517 various cue conditions, trials for each cue condition for a unit were shuffled and divided into two averages,
518 and averages within and across cue conditions were correlated. A linear mixed effects model was run for
519 each cue condition to determine if correlations of firing rates within cue conditions were more similar than
520 correlations across cue conditions.

521 To identify the responsivity of units to different cue features at the time of a nosepoke into a reward re-
522 ceptacle, and subsequent reward delivery, the same cue-responsive units from the cue-onset analyses were
523 analyzed at the time of nosepoke and outcome receipt using identical analysis techniques for all approach
524 trials (Figures 8, 9, ??, and ??). To compare whether coding of a given cue feature was accomplished by the
525 same or distinct population of units across time epochs, we ran the recoded coefficient correlation that was
526 used to assess the degree of overlap among cue features within a time epoch.

527 Given that some of our analyses compare firing rates across time, particularly comparisons across blocks,
528 we sought to exclude units with unstable firing rates that would generate spurious results reflecting a drift

529 ~~in firing rate over time unrelated to our task. To do this we ran a Mann-Whitney U test comparing the~~
530 ~~cue-evoked firing rates for the first and second half of trials within a block, and excluded 99 of 443 units~~
531 ~~from analysis that showed a significant change for either block, leaving 344 units for further analyses.~~
532 All analyses were completed in MATLAB R2015a, the code is available on our public GitHub repository
533 (<http://github.com/vandermeerlab/papers>), and the data can be accessed through DataLad.

534 **Histology:**

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

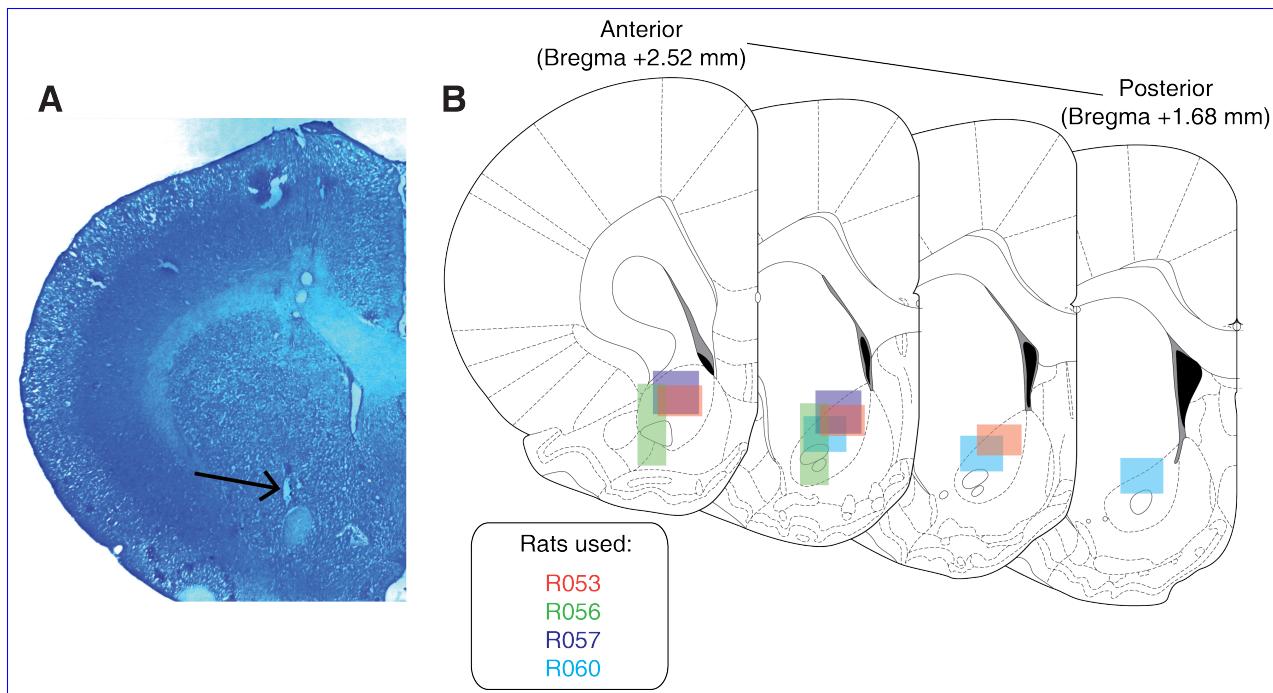


Figure 10: Histological verification of recording sites. Upon completion of experiments, brains were sectioned and tetrode placement was confirmed. **A:** Example section from R060 showing a recording site in the NAc core just dorsal to the anterior commissure (arrow). **B:** Schematic showing recording areas for all subjects.

541 **References**

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

- 572 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).
573 Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron*, 87(5),
574 1106–1118. doi: 10.1016/j.neuron.2015.08.018
- 575 Cheer, J. F., Aragona, B. J., Heien, M. L. A. V., Seipel, A. T., Carelli, R. M., & Wightman, R. M. (2007, apr). Coordinated
576 Accumbal Dopamine Release and Neural Activity Drive Goal-Directed Behavior. *Neuron*, 54(2), 237–244. doi:
577 10.1016/j.neuron.2007.03.021
- 578 Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in
579 schizophrenia. *Psychological Review*, 99(1), 45.doi: 10.1016/j.neuron.2007.03.021
- 580 Cooch, N. K., Stalnaker, T. A., Wied, H. M., Bali-Chaudhary, S., McDannald, M. A., Liu, T. L., & Schoenbaum, G. (2015,
581 jun). Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. *Nature
582 Communications*, 6, 7195. doi: 10.1038/ncomms8195
- 583 Corbit, L. H., & Balleine, B. W. (2011, aug). The General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer Are
584 Differentially Mediated by the Nucleus Accumbens Core and Shell. *Journal of Neuroscience*, 31(33), 11786–11794. doi:
585 10.1523/JNEUROSCI.2711-11.2011
- 586 Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in
587 Neurobiology*, 92, 345–369. doi: 10.1016/j.pneurobio.2010.06.007
- 588 Cromwell, H. C., & Schultz, W. (2003, may). Effects of Expectations for Different Reward Magnitudes on Neuronal Activity in
589 Primate Striatum. *Journal of Neurophysiology*, 89(5), 2823–2838. doi: 10.1152/jn.01014.2002
- 590 Day, J. J., Roitman, M. F., Wightman, R. M., & Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine
591 signaling in the nucleus accumbens. *Nature Neuroscience*, 10(8), 1020–1028. doi: 10.1038/nn1923
- 592 Day, J. J., Wheeler, R. A., Roitman, M. F., & Carelli, R. M. (2006, mar). Nucleus accumbens neurons encode Pavlovian ap-
593 proach behaviors: Evidence from an autoshaping paradigm. *European Journal of Neuroscience*, 23(5), 1341–1351. doi:
594 10.1111/j.1460-9568.2006.04654.x
- 595 Dejean, C., Sitko, M., Girardeau, P., Bennabi, A., Caillé, S., Cador, M., ... Le Moine, C. (2017, apr). Memories of Opiate
596 Withdrawal Emotional States Correlate with Specific Gamma Oscillations in the Nucleus Accumbens. *Neuropsychophar-
597 macology*, 42(5), 1157–1168. doi: 10.1038/npp.2016.272
- 598 du Hoffmann, J., & Nicola, S. M. (2014, oct). Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the
599 Nucleus Accumbens. *Journal of Neuroscience*, 34(43), 14349–14364. doi: 10.1523/JNEUROSCI.3492-14.2014
- 600 Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental
601 Psychology*, 32(2), 150–155. doi: 10.1037/h0058316
- 602 Fitzgerald, T. H. B., Schwartenbeck, P., & Dolan, R. J. (2014). Reward-Related Activity in Ventral Striatum Is Action Contingent
603 and Modulated by Behavioral Relevance. *Journal of Neuroscience*, 34(4), 1271–1279. doi: 10.1523/JNEUROSCI.4389-
604 13.2014

- 605 Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., ... Akil, H. (2011, jan). A selective role for dopamine
606 in stimulus-reward learning. *Nature*, 469(7328), 53–59. doi: 10.1038/nature09588
- 607 Floresco, S. B. (2015, jan). The Nucleus Accumbens: An Interface Between Cognition, Emotion, and Action. *Annual Review of
608 Psychology*, 66(1), 25–52. doi: 10.1146/annurev-psych-010213-115159
- 609 Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., & Magyar, O. (2006, mar). Dissociable roles for the nucleus accumbens core and
610 shell in regulating set shifting. *Journal of Neuroscience*, 26(9), 2449–2457. doi: 10.1523/JNEUROSCI.4431-05.2006
- 611 Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004, dec). By carrot or by stick: Cognitive reinforcement learning in Parkinson-
612 ism. *Science*, 306(5703), 1940–1943. doi: 10.1126/science.1102941
- 613 Giertler, C., Bohn, I., & Hauber, W. (2004, feb). Transient inactivation of the rat nucleus accumbens does not impair guidance of
614 instrumental behaviour by stimuli predicting reward magnitude. *Behavioural Pharmacology*, 15(1), 55–63. doi: 10.1126/sci-
615 ence.1102941
- 616 Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., ... Roesch, M. R. (2012, feb). Ventral
617 Striatum Encodes Past and Predicted Value Independent of Motor Contingencies. *Journal of Neuroscience*, 32(6), 2027–
618 2036. doi: 10.1523/JNEUROSCI.5349-11.2012
- 619 Goto, Y., & Grace, A. A. (2008, nov). Limbic and cortical information processing in the nucleus accumbens. *Trends in Neuro-
620 sciences*, 31(11), 552–558. doi: 10.1016/j.tins.2008.08.002
- 621 Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., ... Steele, J. D. (2011, jun). Expected value and prediction
622 error abnormalities in depression and schizophrenia. *Brain*, 134(6), 1751–1764. doi: 10.1093/brain/awr059
- 623 Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a
624 Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. doi: 10.1037/h0059831
- 625 Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., ... Berke, J. D. (2015, jan).
626 Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1), 117–126. doi: 10.1038/nn.4173
- 627 Hart, A. S., Rutledge, R. B., Glimcher, P. W., & Phillips, P. E. M. (2014, jan). Phasic Dopamine Release in the Rat Nucleus
628 Accumbens Symmetrically Encodes a Reward Prediction Error Term. *The Journal of Neuroscience*, 34(3), 698–704. doi:
629 10.1523/JNEUROSCI.2489-13.2014
- 630 Hassani, O. K., Cromwell, H. C., & Schultz, W. (2001, jun). Influence of Expectation of Different Rewards on Behavior-Related
631 Neuronal Activity in the Striatum. *Journal of Neurophysiology*, 85(6), 2477–2489. doi: 10.1152/jn.2001.85.6.2477
- 632 Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: the stimulus-reinforcer relation and directed action*. Psychonomic Society. doi:
633 10.1152/jn.2001.85.6.2477
- 634 Hill, D. N., Mehta, S. B., & Kleinfeld, D. (2011, jun). Quality Metrics to Accompany Spike Sorting of Extracellular Signals.
635 *Journal of Neuroscience*, 31(24), 8699–8705. doi: 10.1523/JNEUROSCI.0971-11.2011
- 636 Holland, P. C. (1992, jan). Occasion setting in pavlovian conditioning. *Psychology of Learning and Motivation*, 28(C), 69–125.
637 doi: 10.1016/S0079-7421(08)60488-0

- 638 Hollerman, J. R., Tremblay, L., & Schultz, W. (1998, aug). Influence of Reward Expectation on Behavior-Related Neuronal Activity
639 in Primate Striatum. *Journal of Neurophysiology*, 80(2), 947–963. doi: 10.1152/jn.1998.80.2.947
- 640 Honey, R. C., Iordanova, M. D., & Good, M. (2014, feb). Associative structures in animal learning: Dissociating elemental and
641 configural processes. *Neurobiology of Learning and Memory*, 108, 96–103. doi: 10.1016/j.nlm.2013.06.002
- 642 Hyman, S. E., Malenka, R. C., & Nestler, E. J. (2006, jul). NEURAL MECHANISMS OF ADDICTION: The Role
643 of Reward-Related Learning and Memory. *Annual Review of Neuroscience*, 29(1), 565–598. doi: 10.1146/an-
644 nurev.neuro.29.051605.113009
- 645 Ikemoto, S. (2007, nov). Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-
646 olfactory tubercle complex. *Brain Research Reviews*, 56(1), 27–78. doi: 10.1016/j.brainresrev.2007.05.004
- 647 Ishikawa, A., Ambroggi, F., Nicola, S. M., & Fields, H. L. (2008, may). Dorsomedial Prefrontal Cortex Contribution to Behav-
648 ioral and Nucleus Accumbens Neuronal Responses to Incentive Cues. *Journal of Neuroscience*, 28(19), 5088–5098. doi:
649 10.1523/JNEUROSCI.0253-08.2008
- 650 Joel, D., Niv, Y., & Ruppin, E. (2002, jun). Actor-critic models of the basal ganglia: new anatomical and computational perspectives.
651 *Neural Networks*, 15(4-6), 535–547. doi: 10.1016/S0893-6080(02)00047-3
- 652 Kaczkurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., ... Lissek, S. (2017, feb).
653 Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174(2), 125–134. doi:
654 10.1176/appi.ajp.2016.15121549
- 655 Kaczkurkin, A. N., & Lissek, S. (2013). Generalization of Conditioned Fear and Obsessive-Compulsive Traits. *Journal of Psychol-
656 ogy & Psychotherapy*, 7, 3. doi: 10.4172/2161-0487.S7-003
- 657 Kalivas, P. W., & Volkow, N. D. (2005, aug). The neural basis of addiction: A pathology of motivation and choice. *American
658 Journal of Psychiatry*, 162(8), 1403–1413. doi: 10.1176/appi.ajp.162.8.1403
- 659 Kapur, S. (2003, jan). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology
660 in schizophrenia. *American Journal of Psychiatry*, 160(1), 13–23. doi: 10.1176/appi.ajp.160.1.13
- 661 Khamassi, M., & Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and
662 model-free navigation strategies. *Frontiers in Behavioral Neuroscience*, 6, 79. doi: 10.3389/fnbeh.2012.00079
- 663 Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., & Wiener, S. I. (2008, nov). Anticipatory reward signals in ventral striatal
664 neurons of behaving rats. *European Journal of Neuroscience*, 28(9), 1849–1866. doi: 10.1111/j.1460-9568.2008.06480.x
- 665 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., & Pennartz, C. M. A. (2008, jun).
666 Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, 28(25),
667 6372–6382. doi: 10.1523/JNEUROSCI.1054-08.2008
- 668 Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., & Pennartz, C. M. (2009, aug). Hippocampus leads ventral
669 striatum in replay of place-reward information. *PLoS Biology*, 7(8), e1000173. doi: 10.1371/journal.pbio.1000173
- 670 Lansink, C. S., Jackson, J. C., Lankelma, J. V., Ito, R., Robbins, T. W., Everitt, B. J., & Pennartz, C. M. A. (2012, sep). Reward

- 671 Cues in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles. *Journal*
672 *of Neuroscience*, 32(36), 12444–12459. doi: 10.1523/JNEUROSCI.0593-12.2012
- 673 Lansink, C. S., Meijer, G. T., Lankelma, J. V., Vinck, M. A., Jackson, J. C., & Pennartz, C. M. A. (2016, oct). Reward Ex-
674 pectancy Strengthens CA1 Theta and Beta Band Synchronization and Hippocampal-Ventral Striatal Coupling. *Journal of*
675 *Neuroscience*, 36(41), 10598–10610. doi: 10.1523/JNEUROSCI.0682-16.2016
- 676 Lavoie, A. M., & Mizumori, S. J. (1994, feb). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons
677 of rats. *Brain Research*, 638(1-2), 157–168. doi: 10.1016/0006-8993(94)90645-9
- 678 Lee, D., Seo, H., & Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of*
679 *Neuroscience*, 35(1), 287–308. doi: 10.1146/annurev-neuro-062111-150512
- 680 Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*,
681 22, 1027–1038. doi: 10.1016/j.conb.2012.06.001
- 682 Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014, jun). Generalized anxiety disor-
683 der is associated with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75(11), 909–915. doi:
684 10.1016/j.biopsych.2013.07.025
- 685 Maia, T. V. (2009, dec). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and*
686 *Behavioral Neuroscience*, 9(4), 343–364. doi: 10.3758/CABN.9.4.343
- 687 Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature*
688 *Neuroscience*, 14(2), 154–162. doi: 10.1038/nn.2723
- 689 Malhotra, S., Cross, R. W., Zhang, A., & Van Der Meer, M. A. A. (2015). Ventral striatal gamma oscillations are highly variable
690 from trial to trial, and are dominated by behavioural state, and only weakly influenced by outcome value. *European Journal*
691 *of Neuroscience*, 42(10), 2818–2832. doi: 10.1038/nn.2723
- 692 McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011, feb). Ventral Striatum and Orbitofrontal Cortex
693 Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, 31(7), 2700–
694 2705. doi: 10.1523/JNEUROSCI.5499-10.2011
- 695 McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., & Nicola, S. M. (2013, jun). Invigoration of reward seeking by cue and
696 proximity encoding in the nucleus accumbens. *Neuron*, 78(5), 910–922. doi: 10.1016/j.neuron.2013.04.010
- 697 Mulder, A. B., Shibata, R., Trullier, O., & Wiener, S. I. (2005, may). Spatially selective reward site responses in tonically active
698 neurons of the nucleus accumbens in behaving rats. *Experimental Brain Research*, 163(1), 32–43. doi: 10.1007/s00221-
699 004-2135-3
- 700 Mulder, A. B., Tabuchi, E., & Wiener, S. I. (2004, apr). Neurons in hippocampal afferent zones of rat striatum parse routes into
701 multi-pace segments during maze navigation. *European Journal of Neuroscience*, 19(7), 1923–1932. doi: 10.1111/j.1460-
702 9568.2004.03301.x
- 703 Nicola, S. M. (2004, apr). Cue-Evoked Firing of Nucleus Accumbens Neurons Encodes Motivational Significance During a

- 704 Discriminative Stimulus Task. *Journal of Neurophysiology*, 91(4), 1840–1865. doi: 10.1152/jn.00657.2003
- 705 Nicola, S. M. (2010). The Flexible Approach Hypothesis: Unification of Effort and Cue-Responding Hypotheses for the Role of
706 Nucleus Accumbens Dopamine in the Activation of Reward-Seeking Behavior. *Journal of Neuroscience*, 30(49), 16585–
707 16600. doi: 10.1523/JNEUROSCI.3958-10.2010
- 708 Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007, mar). Tonic dopamine: Opportunity costs and the control of response vigor.
709 *Psychopharmacology*, 191(3), 507–520. doi: 10.1007/s00213-006-0502-4
- 710 Noonan, M. P., Chau, B. K., Rushworth, M. F., & Fellows, L. K. (2017, jul). Contrasting Effects of Medial and Lateral Orbitofrontal
711 Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience*, 37(29), 7023–7035.
712 doi: 10.1523/JNEUROSCI.0692-17.2017
- 713 Paxinos, G., & Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates* (4th ed.). San Diego: Academic Press. doi:
714 10.1523/JNEUROSCI.0692-17.2017
- 715 Pennartz, C. M. A. (2004, jul). The Ventral Striatum in Off-Line Processing: Ensemble Reactivation during Sleep and Modulation
716 by Hippocampal Ripples. *Journal of Neuroscience*, 24(29), 6446–6456. doi: 10.1523/JNEUROSCI.0575-04.2004
- 717 Pennartz, C. M. A., Ito, R., Verschure, P., Battaglia, F., & Robbins, T. (2011, oct). The hippocampalstriatal axis in learning,
718 prediction and goal-directed behavior. *Trends in Neurosciences*, 34(10), 548–559. doi: 10.1016/j.tins.2011.08.001
- 719 Peters, J., & Büchel, C. (2009, dec). Overlapping and distinct neural systems code for subjective value during intertemporal and
720 risky decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(50), 15727–34.
721 doi: 10.1523/JNEUROSCI.3489-09.2009
- 722 Rescorla, R. A., & Solomon, R. L. (1967). Two-Process Learning Theory: Relationships Between Pavlovian Conditioning and
723 Instrumental Learning. *Psychological Review*, 74(3), 151–182. doi: 10.1037/h0024475
- 724 Robinson, T. E., & Flagel, S. B. (2009, may). Dissociating the Predictive and Incentive Motivational Properties of
725 Reward-Related Cues Through the Study of Individual Differences. *Biological Psychiatry*, 65(10), 869–873. doi:
726 10.1016/j.biopsych.2008.09.006
- 727 Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009, oct). Ventral Striatal Neurons Encode the Value
728 of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience*, 29(42),
729 13365–13376. doi: 10.1523/JNEUROSCI.2572-09.2009
- 730 Roitman, M. F., Wheeler, R. A., & Carelli, R. M. (2005, feb). Nucleus accumbens neurons are innately tuned for reward-
731 ing and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, 45(4), 587–597. doi:
732 10.1016/j.neuron.2004.12.055
- 733 Saddoris, M. P., Stamatakis, A., & Carelli, R. M. (2011, jun). Neural correlates of Pavlovian-to-instrumental transfer in the nu-
734 cleus accumbens shell are selectively potentiated following cocaine self-administration. *European Journal of Neuroscience*,
735 33(12), 2274–2287. doi: 10.1111/j.1460-9568.2011.07683.x
- 736 Salamone, J. D., & Correa, M. (2012, nov). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron*, 76(3),

- 737 470–485. doi: 10.1016/j.neuron.2012.10.021
- 738 Samuelson, P. A. (1947). *Foundations of Economic Analysis*. Cambridge MA: Harvard University Press.doi:
739 10.1016/j.neuron.2012.10.021
- 740 Schultz, W. (2016, feb). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32. doi:
741 10.1038/nrn.2015.26
- 742 Schultz, W., Apicella, P., Scarnati, E., & Ljungberg, T. (1992, dec). Neuronal activity in monkey ventral striatum related to the
743 expectation of reward. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 12(12), 4595–610.
744 doi: 10.1523/JNEUROSCI.12-12-04595.1992
- 745 Schultz, W., Dayan, P., & Montague, P. R. (1997, mar). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.
746 doi: 10.1126/science.275.5306.1593
- 747 Sescousse, G., Li, Y., & Dreher, J.-C. (2015, apr). A common currency for the computation of motivational values in the human
748 striatum. *Social Cognitive and Affective Neuroscience*, 10(4), 467–473. doi: 10.1093/scan/nsu074
- 749 Setlow, B., Schoenbaum, G., & Gallagher, M. (2003, may). Neural encoding in ventral striatum during olfactory discrimination
750 learning. *Neuron*, 38(4), 625–636. doi: 10.1016/S0896-6273(03)00264-2
- 751 Shidara, M., Aigner, T. G., & Richmond, B. J. (1998, apr). Neuronal signals in the monkey ventral striatum related to progress
752 through a predictable series of trials. *Journal of Neuroscience*, 18(7), 2613–25.doi: 10.1016/S0896-6273(03)00264-2
- 753 Sjulson, L., Peyrache, A., Cumpelik, A., Cassataro, D., & Buzsáki, G. (2017, may). Cocaine place conditioning strengthens
754 location-specific hippocampal inputs to the nucleus accumbens. *bioRxiv*, 1–10. doi: 10.1101/105890
- 755 Sleezer, B. J., Castagno, M. D., & Hayden, B. Y. (2016, nov). Rule Encoding in Orbitofrontal Cortex and Striatum Guides Selection.
756 *Journal of Neuroscience*, 36(44), 11223–11237. doi: 10.1523/JNEUROSCI.1766-16.2016
- 757 Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., & Hayden, B. Y. (2016, mar). Neuronal selectivity
758 for spatial positions of offers and choices in five reward regions. *Journal of Neurophysiology*, 115(3), 1098–1111. doi:
759 10.1152/jn.00325.2015
- 760 Sugam, J. A., Saddoris, M. P., & Carelli, R. M. (2014, may). Nucleus accumbens neurons track behavioral preferences and reward
761 outcomes during risky decision making. *Biological Psychiatry*, 75(10), 807–816. doi: 10.1016/j.biopsych.2013.09.010
- 762 Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction* (Vol. 9) (No. 5). MIT Press, Cambridge, MA. doi:
763 10.1109/TNN.1998.712192
- 764 Tabuchi, E. T., Mulder, A. B., & Wiener, S. I. (2000, jan). Position and behavioral modulation of synchronization of hip-
765 pocampal and accumbens neuronal discharges in freely moving rats. *Hippocampus*, 10(6), 717–728. doi: 10.1002/1098-
766 1063(2000)10:6;717::AID-HIPO1009;3.0.CO;2-3
- 767 Takahashi, Y. K., Langdon, A. J., Niv, Y., & Schoenbaum, G. (2016, jul). Temporal Specificity of Reward Prediction Er-
768 rrors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. *Neuron*, 91(1), 182–193. doi:
769 10.1016/j.neuron.2016.05.015

- 770 Tingley, D., & Buzsáki, G. (2018, jun). Transformation of a Spatial Map across the Hippocampal-Lateral Septal Circuit. *Neuron*,
771 98(6), 1229–1242.e5. doi: 10.1016/J.NEURON.2018.04.028
- 772 van der Meer, M. A. A., & Redish, A. D. (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward
773 Information. *Journal of Neuroscience*, 31(8), 2843–2854. doi: 10.1523/JNEUROSCI.4869-10.2011
- 774 West, E. A., & Carelli, R. M. (2016, jan). Nucleus Accumbens Core and Shell Differentially Encode Reward-Associated Cues after
775 Reinforcer Devaluation. *Journal of Neuroscience*, 36(4), 1128–1139. doi: 10.1523/JNEUROSCI.2976-15.2016
- 776 Wiener, S. I., Shibata, R., Tabuchi, E., Trullier, O., Albertin, S. V., & Mulder, A. B. (2003, oct). Spatial and behavioral correlates
777 in nucleus accumbens neurons in zones receiving hippocampal or prefrontal cortical inputs. *International Congress Series*,
778 1250(C), 275–292. doi: 10.1016/S0531-5131(03)00978-6
- 779 Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral
780 and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, 24(12), 2923–2933. doi:
781 10.1523/JNEUROSCI.5282-03.2004

782 –