

# Neurons in the Nucleus Accumbens Promote Selection Bias for Nearer Objects

Sara E. Morrison<sup>1</sup> and Saleem M. Nicola<sup>1,2</sup>

<sup>1</sup>Department of Psychiatry and Behavioral Science and <sup>2</sup>Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, New York 10461

Both animals and humans often prefer rewarding options that are nearby over those that are distant, but the neural mechanisms underlying this bias are unclear. Here we present evidence that a proximity signal encoded by neurons in the nucleus accumbens drives proximate reward bias by promoting impulsive approach to nearby reward-associated objects. On a novel decision-making task, rats chose the nearer option even when it resulted in greater effort expenditure and delay to reward; therefore, proximate reward bias was unlikely to be caused by effort or delay discounting. The activity of individual neurons in the nucleus accumbens did not consistently encode the reward or effort associated with specific alternatives, suggesting that it does not participate in weighing the values of options. In contrast, proximity encoding was consistent and did not depend on the subsequent choice, implying that accumbens activity drives approach to the nearest rewarding option regardless of its specific associated reward size or effort level.

**Key words:** decision-making; effort; impulsivity; nucleus accumbens; proximity; reward

## Introduction

People and animals show a pronounced preference for rewarding options that are nearby as opposed to farther away (Perrings and Hannon, 2001; Stevens et al., 2005; Kralik and Sampson, 2012; O'Connor et al., 2014), a phenomenon we call proximate reward bias. Neural activity in the hippocampus, prefrontal cortex, and nucleus accumbens (NAc) reflects the proximity of rewards (Hok et al., 2005; Lee et al., 2006; van der Meer and Redish, 2009; Young and Shapiro, 2011; Lansink et al., 2012; Pfeiffer and Foster, 2013); however, the mechanisms by which proximity biases decision-making are not well understood. Recently, dopamine in the striatum (Howe et al., 2013) and activity in the ventral tegmental area (VTA) (Gomperts et al., 2013) were found to increase as animals approached a reward, suggesting that dopamine promotes movement toward proximate rewards. Consistent with these findings, we have shown that NAc neurons encode proximity to an approach target (McGinty et al., 2013), raising the possibility that this signal biases animals to choose nearer objects. This could occur in two ways: via cognitive evaluation of the options or a noncognitive process of conditioned approach.

Some investigators have proposed that NAc activity might participate in decision-making either directly (Setlow et al., 2003; Salamone et al., 2007; Day et al., 2011) or as the “critic” in actor–critic models of reinforcement learning (O'Doherty et al., 2004; Khamassi et al., 2008; Roesch et al., 2009; van der Meer and Redish, 2011). Indeed, NAc activity can be modulated by reward availability, size, delay, and effort requirement (Setlow et al., 2003; Nicola et al., 2004; Roitman et al., 2005; Roesch et al., 2009; Day et al., 2011). Such signals might participate in calculating the values of different choice targets (e.g., levers to press) by integrating costs and benefits and weighing the values to choose the optimal target. Thus, NAc neurons that encode proximity could contribute to decision-making by enhancing the value of nearby choice targets or discounting the value of more distal targets.

On the other hand, choosing a nearby over a distal rewarding stimulus might simply reflect conditioned approach to a nearby rewarding object. Conditioned approach can be evoked by a reward-associated stimulus even under circumstances in which the approach behavior is itself nonadaptive (Dayan et al., 2006; Clark et al., 2012). We hypothesized that, rather than contributing to cognitive evaluation of the options, the NAc proximity signal might drive impulsive approach behavior that is preferentially directed toward proximate stimuli. We tested this hypothesis with a decision-making task in which proximity to the choice targets was highly variable, and in which the reward size and effort associated with each target were systematically varied. Although subjects' choices were mostly adaptive, rats were much more likely to choose a lever that was in close proximity whether or not it was the optimal choice. Here we present evidence that NAc proximity signaling contributes to impulsive approach driven by the close proximity of a reward-associated target. Notably, we found separate encoding of proximity, reward size, and

Received May 30, 2014; revised Sept. 9, 2014; accepted Sept. 11, 2014.

Author contributions: S.E.M. and S.M.N. designed research; S.E.M. performed research; S.E.M. analyzed data; S.E.M. and S.M.N. wrote the paper.

This work was supported by National Institutes of Health Grants DA019473, DA038412, and MH092757 to S.M.N. and Grant DA034465 to S.E.M., National Alliance for Research on Schizophrenia and Depression, the Klarman Family Foundation, and the Peter F. McManus Charitable Trust to S.M.N., and the Charles H. Revson Foundation to S.E.M. We thank S. Lardeux, V. McGinty, and P. Rudebeck for comments on the manuscript; members of the S.M.N. laboratory for helpful discussions; A. Geevarghese and D. Johnson for help with rat training; and J. Kim for invaluable technical assistance.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Saleem M. Nicola, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461. E-mail: saleem.nicola@einstein.yu.edu.

DOI:10.1523/JNEUROSCI.2197-14.2014

Copyright © 2014 the authors 0270-6474/14/3414147-16\$15.00/0

effort in the NAC, arguing against the idea that this brain area represents the values of options in a common currency.

## Materials and Methods

All procedures were performed in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and were approved by the Institutional Animal Care and Use Committee of the Albert Einstein College of Medicine.

**Subjects.** Subjects were 9 male Long–Evans rats obtained from Charles River Laboratories. Rats weighed 275–300 g upon arrival. They were singly housed and placed on a 12 h light/dark cycle; all experiments were conducted during the light phase. After arrival, rats were allowed to acclimate to the housing colony for at least 3 d. They were then habituated to human contact and handling over 2 or 3 sessions before the start of training. Behavioral training took place over 10–24 weeks, followed by surgery, 5–7 d of recovery, and collection of neural and behavioral data over a period of 2–12 weeks.

Subjects were provided with water *ad libitum* throughout, and food (standard chow) was provided *ad libitum* until the start of training. After acclimation and habituation, rats were placed on a restricted diet of 14 g of chow per day (BioServ F-137 dustless pellets), which continued throughout behavioral training and experiments (except for a 5–7 d recovery period after surgery). If necessary, extra food was provided to maintain a minimum of 90% of prerestriction body weight.

**Behavioral apparatus and task.** Training and electrophysiological recording took place in a Plexiglas chamber measuring 40 cm × 40 cm (height 60 cm). The chamber was outfitted with a reward receptacle flanked by two retractable levers (MED Associates), a speaker for the auditory stimulus, and a separate speaker providing constant white noise, as well as an overhead video camera (Plexon) used for motion tracking (see Figure 1C). The speaker for the auditory stimulus was located 55 cm above floor level on the same wall as the receptacle, and the intensity of the stimulus (75 dB) was virtually constant throughout the chamber when measured at rats' head level. The behavioral task was controlled by MED-PC software (MED Associates), and behavioral events were collected at 1 ms resolution and recorded concurrently with neural data and video data using Plexon hardware and software.

During the task (see Fig. 1A, B), freely moving rats chose between two levers (left or right) to obtain a liquid sucrose reward (10%). During reward size blocks, one lever was associated with a large reward (75  $\mu$ l) and the other with a small reward (45  $\mu$ l), and both levers had a low effort requirement (one press). During effort blocks, both levers were associated with the large reward, but one lever had a high effort requirement (16 presses), whereas the other had a low effort requirement (one press). Each block was composed of 20 forced choice trials, during which only one randomly selected lever was available, followed by 40 free choice trials, during which both levers were available simultaneously. Two reward size blocks and two effort blocks were presented during each session. The order of blocks and starting lever assignments were varied among sessions such that all possible block sequences were represented equally, with the exception that reward blocks were always presented back-to-back, as were effort blocks. Reward/effort contingencies were reversed between blocks; the reversal was signaled only by a return to forced choice trials.

Trials were presented after a variable intertrial interval (ITI) that was exponentially distributed (mean = 10 s), approximating a constant cue onset probability. Each trial was signaled by the presentation of an auditory cue that played continuously until a lever was selected (up to 15 s); the extension of one lever (forced choice trials) or both levers (free choice trials); and the illumination of a light above each active lever. The auditory cue was an intermittent pure tone of 5350 kHz played at 80 dB. If no lever press was detected within 15 s, the cue was terminated, both levers retracted, both cue lights extinguished, and a new ITI was begun. After the rat made one press on a lever, the other lever was immediately retracted and its light extinguished; the selected lever was retracted and its light extinguished after the required number of lever presses was completed. Reward was then delivered upon nosepoke into the receptacle.

**Training procedure.** Training proceeded at an individualized pace for each subject but generally followed the same stages:

- (1) **Receptacle training.** The auditory cue was presented at random intervals drawn from an exponential distribution with mean = 10 s. Rats learned to enter the receptacle during the cue to earn a bolus of sucrose.
- (2) **Lever training.** The auditory cue was presented at intervals of 8 s; simultaneously, one lever was randomly extended and the cue light above the lever illuminated. Rats learned to press the lever once and then enter the receptacle to earn a bolus of sucrose. Upon lever press, the auditory cue ceased, the lever was retracted, and the cue light flashed until receptacle entry was detected. (If there was a lever press but no receptacle entry by 15 s after the start of the auditory cue, the cue light would stop flashing and a new ITI would begin.) After achieving an acceptable number of lever presses, this stage was repeated using intervals of 12 and 20 s and then using a random ITI as in the final task. All subsequent stages used a random ITI.
- (3) **Effort reversal training.** Subjects were simultaneously trained to perform multiple lever presses and to choose between two levers. As in the final task, rats were initially given 20 forced choice trials, in which the auditory cue was presented simultaneously with the extension of one lever and illumination of its cue light. The left or right lever alternately was assigned to the "high effort" condition. In the first phase, rats were required to press the high-effort lever twice, or the low-effort lever once, and then enter the receptacle to receive a "large" reward (same reward for both effort levels). After the forced choice trials, rats were given 40 free choice trials, which were identical to forced choice trials except both levers were extended and rats could choose either. After completion of the free choice block, the effort contingencies were reversed, and rats experienced another forced choice block followed by another free choice block. Rats experienced up to 8 reversals over the course of the 2 h session. After rats learned to perform the task at a given effort level, they advanced to the next effort level, accomplishing the following numbers of lever presses: 2, 4, 6, 8, 12, and 16. No specific choice behavior was required to advance to the next level, but a preference for the low-effort lever gradually emerged over the course of training. Rats were held at the 16 lever press condition until choice percentages were determined to be relatively stable from day to day (typically 8 sessions). Except for the multiple reversals, this phase was identical to the "effort blocks" used in the final task.
- (4) **Reward size reversal training.** Rats were initially given 20 forced choice trials, in which the auditory cue was presented simultaneously with the extension of one lever and illumination of its cue light. The left or right lever alternately was assigned to the "large reward" condition. The rat was required to press the lever once to receive either a large reward or a small reward. After the forced choice trials, rats were given 40 free choice trials, which were identical to forced choice trials except both levers were extended and rats could choose either. After completion of the free choice block, the reward size contingencies were reversed, and rats experienced another forced choice block followed by another free choice block. Rats experienced up to 8 reversals over the course of the 2 h session. Except for the multiple reversals, this phase was identical to the "reward size blocks" used in the final task. Rats gradually developed a preference for the lever associated with large reward. Subjects were held at this phase until choice percentages were determined to be relatively stable from day to day (typically 8–12 sessions). Occasionally during this phase (typically once a week), effort reversal blocks were given to ensure behavioral maintenance.
- (5) **Mixed reversal training.** Rats were presented with either two effort blocks (one reversal) followed by several (2–6) reward size blocks or two reward size blocks (one reversal) followed by several (2–6) effort blocks. Effort blocks and reward size blocks were presented first on alternating days. This phase was identical to the final task, except the final task was terminated after one reversal of each type, rather than after a fixed amount of time (typically 2 h). Subjects were held at this phase until choice percentages were determined

to be relatively stable from day to day on both block types (minimum 8 sessions, 4 of each type). After completing this phase, rats were considered ready for surgery and data collection.

**Implantation of electrode arrays.** Following initial training, we implanted custom-built drivable microelectrode arrays bilaterally targeted at the dorsal border of the NAc core (coordinates in mm relative to bregma: anteroposterior 1.4, mediolateral  $\pm 1.5$ , dorsoventral 6.25) using standard aseptic procedures. Recording arrays comprised 8 Teflon-insulated tungsten wires (AM Systems) cut by hand to achieve an impedance of 95–115  $M\Omega$  and mounted in a circular pattern (diameter  $\sim 1$  mm) in a custom-designed microdrive system (du Hoffmann et al., 2011). Rats were anesthetized with isoflurane (5% in oxygen for induction of anesthesia; 1%–2% for maintenance) and treated with an antibiotic (Baytril; 10 mg/kg) and analgesic (ketoprofen; 10 mg/kg) after surgery.

**Histology.** After completion of data collection, animals were deeply anesthetized with pentobarbital (100 mg/kg) and direct current (50  $\mu A$ ) passed through each of the electrodes in the array for 10 s. Next, animals were transcardially perfused with 0.9% saline, followed by 8% PFA; brains were removed and placed in PFA. Later, brains were sunk in 30% sucrose for at least 3 d before sectioning on a cryostat (50  $\mu m$  slices). Slices were mounted on glass slides and stained with cresyl violet using standard histological procedures. Final locations of arrays were determined by mapping electrolytic lesions and electrode tracks (where visible), and positions of recorded cells reconstructed using experimental notes.

**Neural data acquisition.** After at least 1 week of recovery from surgery, we recorded the activity of individual neurons during task performance using Plexon hardware and software. Voltages were amplified by 5000–18,000 (typically 10,000) and bandpass filtered at 250 and 8800 Hz. Putative spikes were time-stamped and stored in segments of 1.4 ms. Spikes were sorted offline (Offline Sorter, Plexon) using principal component analysis and visual inspection of waveform clusters in 3D feature-based space. Only putative units with peak amplitude  $>75 \mu V$ , a signal-to-noise ratio exceeding 2:1, and fewer than 0.1% of interspike intervals  $<2$  ms were analyzed. We verified isolation of single units by inspecting autocorrelograms for each recording, as well as cross-correlograms for those neurons recorded on the same electrode.

After collecting data at a particular depth (typically for 2–4 sessions), the electrodes were advanced as a group in increments of  $\sim 150 \mu m$ , which was typically adequate to eliminate previously observed neural activity and introduce new units. To ensure that no neurons were included twice, only one session per channel per dorsoventral location was used as part of the dataset.

**Video recording and analysis of locomotion.** During each session, a video recording was obtained synchronously with neural and behavioral data using Plexon hardware and software (Cineplex V3.0). The software automatically tracked the positions of two LEDs (red and green) mounted on the recording headstage. The entire video image (1.5 mm resolution) was saved to disk at a rate of 30 frames/s.

The video tracking data was preprocessed as follows. First, we corrected for visual distortion as well as skew introduced by the angle of the camera, as described previously (McGinty et al., 2013). Next, we used linear interpolation to fill in any missing data points (e.g., because of occlusion of the LEDs by the recording cable). Then we excluded from analysis stretches of missing data  $>0.5$  s, as well as any stretch of data in which  $>70\%$  of points were missing over the course of 1 s. We also excluded data points in which the distance between the two LEDs fell outside of 2 SDs of the mean distance for that session (which usually indicated that the software was picking up reflections of the LEDs in the chamber walls, rather than the LEDs themselves).

For the purposes of this study, we used the video tracking data to obtain the following measures: distance from the left lever, right lever, and receptacle at cue onset; latency to reach maximum velocity after cue onset; and latency to begin motion after cue onset (“motion onset latency”). To identify the time of motion onset, we calculated a locomotor index (LI) as described previously (Drai et al., 2000; McGinty et al., 2013). Briefly, for each time point  $t$ , the LI is calculated by determining the SD of the distances between the rat’s position at each successive time

point over the 9 video frames surrounding  $t$ . Thus, the LI is a temporally and spatially smoothed representation of speed.

Using the LI, we identified trials in which the rat was relatively motionless at cue onset. For each of these, we calculated the time of motion onset by determining the first video frame when the LI exceeded a threshold value. This threshold was determined individually for each behavioral session based on the specific distribution of LIs for that session. For each session, we fit the log-transformed LI with a Gaussian mixture model that was the sum of three Gaussian distributions. The threshold for motion onset was designated as the average of the two highest Gaussian means. We confirmed that motion onset times calculated using this method corresponded well to motion onset times derived from manual scoring of the video recording.

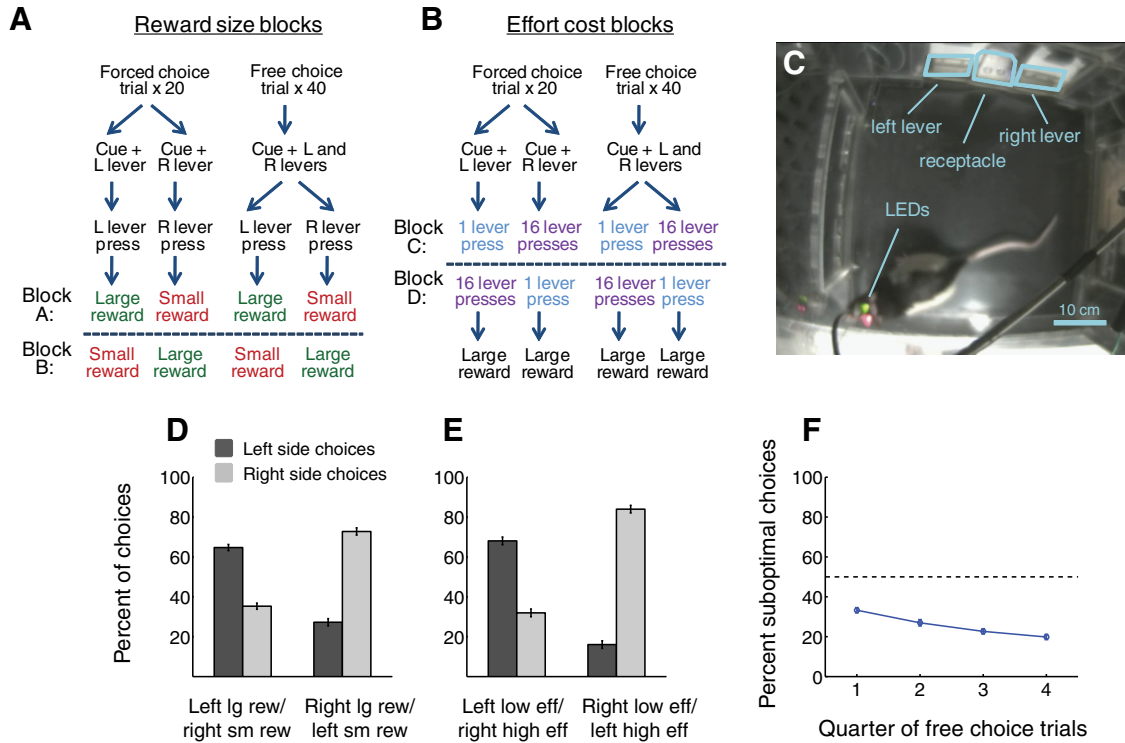
For the purposes of both behavioral and neural analysis, we identified trials when the subject was “near” each lever at cue onset (see Fig. 2A). “Near left lever” was operationally defined as the rat’s head being within a  $\sim 12.5$  cm radius of the left lever, and closer to the left lever than the right lever; “near right lever” was defined similarly. Thus, in “near” trials, at cue onset the animal was within the approximate one-third of the behavioral chamber most proximate to the levers. Trials in which the rat was not determined to be “near” either lever at cue onset were categorized as “far.”

**Analysis of neural data.** To identify cue-evoked excitations, we first defined a Poisson distribution that approximated the baseline firing rate during the 1 s before cue onset. Cue-excited neurons were classified as such by the presence of three or more consecutive 10 ms bins in the 200 ms time window following cue onset in which firing rate exceeded the 99.9% confidence interval of the baseline distribution. We also evaluated whether the cue response was primarily excitatory or inhibitory by assessing the average  $Z$ -score  $>200$  ms after cue onset, calculated in 10 ms bins in relation to 1 s of pre-cue baseline; we excluded from analysis neurons with a negative average  $Z$ -score. We also excluded from analysis 10 neurons with a baseline firing rate  $>18$  Hz because they were unlikely to be medium spiny neurons (Gage et al., 2010).

Peristimulus time histograms for individual neurons (see Figs. 4A–C and 6) were calculated in 10 ms bins and are shown unsmoothed. Population peristimulus time histograms (see Figs. 4D–G and 5E) were calculated in 10 ms bins, normalized to a 1 s pre-cue baseline before averaging across neurons, and smoothed with a 5-bin moving average. Population activity on different trial types was compared using firing rates from the window 50–150 ms after cue onset (capturing the peak of the population cue-evoked response).  $p$  values were Bonferroni corrected where appropriate.

Except where otherwise stated, all other analyses were performed using activity from the 200 ms following cue onset. We used two modes of analysis for characterizing the coding properties of neurons: receiver operating characteristic (ROC) analysis and application of a generalized linear model (GLM). We used ROC analysis to calculate a reward preference index, which represents a comparison between activity before choices of the large reward and small reward on the same lever; and an effort preference index, representing a comparison between activity before expending high effort and low effort on the same lever (see Fig. 7A–C). These indices were calculated separately for each lever. Preference index values of 0.5 indicate no difference;  $>0.5$  indicates greater firing before large reward or high effort;  $<0.5$  indicates greater firing before small reward or low effort.

We also used ROC analysis to calculate a chosen target preference index (see Fig. 8) comparing activity before the two choices available within each block (e.g., left/large reward vs right/small reward). We compared the magnitude of this index to the magnitude of the reward or effort preference index by normalizing both indices to 0.5 (i.e., subtracting 0.5 and taking the absolute value) and computing the difference between the normalized indices (“difference index”). Strongly negative values of the difference index are consistent with target value encoding; strongly positive values are consistent with target identity encoding; and values near zero are consistent with choice value encoding. Finally, we used ROC analysis to calculate a lever proximity index (see Fig. 9A,B), which represents a comparison between cue-evoked activity on trials when the subject was near either lever or far from both levers (see Video recording and analysis of locomotion, above). Values of the lever prox-



**Figure 1.** Task and behavior. **A, B**, Sequence of events during a decision-making task in which reward size (**A**) and effort cost (**B**) were systematically varied. In Block A, a left lever press resulted in the availability of a large reward and a right lever press resulted in the availability of a small reward. In Block B, these contingencies were reversed. In Blocks C and D, the effort level required to obtain reward was varied in a similar manner, but reward size was held constant. The four blocks were presented in pseudo-random order during different sessions, but Blocks A and B were always presented back-to-back, as were Blocks C and D. **C**, Rat in behavioral chamber. The rat’s movements were recorded using video tracking software and two head-mounted LEDs. **D, E**, Average behavioral performance during Blocks A and B (**D**) and Blocks C and D (**E**) over 45 sessions. Error bars indicate SEM. **F**, Probability of suboptimal choices over the course of 40 free choice trials, averaged over 45 sessions. Error bars indicate SEM.

imity index >0.5 indicate higher firing on “near” trials; values <0.5 indicate higher firing on “far” trials.

To investigate the relationship between neural firing and lever proximity, we used two versions of a GLM. The first was a “locomotor factor only” GLM (see Figs. 4H, I and 5C, D) as follows:

$$\ln(Y) = \beta_0 + \beta_{dist}x_{dist} + \beta_{lat}x_{lat} + \varepsilon \quad (1)$$

where  $x_{dist}$  is lever distance,  $x_{lat}$  is latency to maximum velocity,  $\beta_0$ ,  $\beta_{lat}$ , and  $\beta_{dist}$  are the regression coefficients resulting from the model fit,  $\varepsilon$  is the residual error, and  $Y$  is the response variable (i.e., the spike count in the 200 ms window following cue onset). This form of GLM assumes that the response variable follows a Poisson distribution.

We also applied a “locomotor + decision variable” GLM (see Figs. 9C, D and 10) that incorporated both the continuous locomotor factors and the binary factors that describe choice type:

$$\ln(Y) = \beta_0 + \beta_{side}x_{side} + \beta_Lx_L + \beta_Rx_R + \beta_{dist}x_{dist} + \beta_{lat}x_{lat} + \varepsilon \quad (2)$$

where  $x_{side}$  is a dummy variable representing the identity of the chosen target (1 for left lever; 0 for right lever),  $x_L$  and  $x_R$  are dummy variables representing the reward value or effort level associated with the left lever and right lever, respectively (1 for large reward or low effort; -1 for small reward or high effort; 0 if the other lever was chosen), and  $\beta_{side}$ ,  $\beta_L$ , and  $\beta_R$  are the corresponding regression coefficients resulting from the model fit. (All other terms are the same as in Eq. 1.)

To facilitate comparison between the regression estimates, we converted them to an expected percentage change in firing rate resulting from an interdecile shift of the regressor (for continuous factors; Eq. 3) or a shift of the regressor from one state to another (for binary factors; Eq. 4):

$$\text{Percent change} = (e^{\beta_A \times IDR_A} - 1) \times 100\% \quad (3)$$

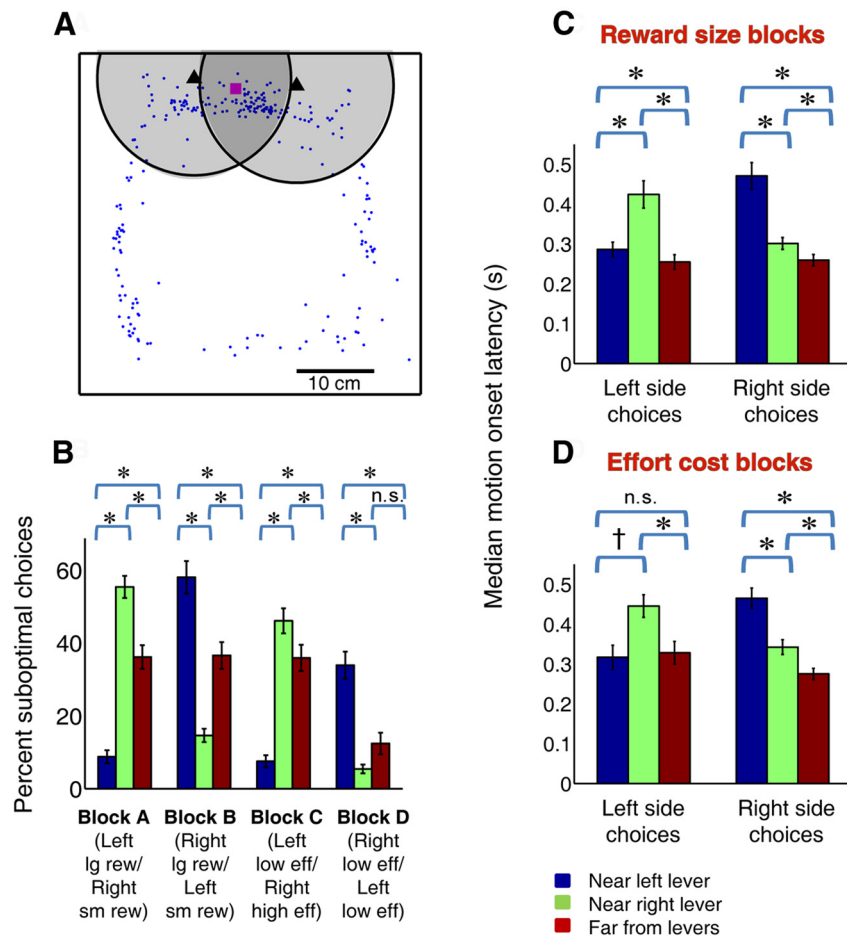
$$\text{Percent change} = (e^{2\beta_B} - 1) \times 100\% \quad (4)$$

where  $\beta_A$  is the regression estimate for the locomotor variable  $A$  and  $IDR_A$  is the interdecile range of that regressor (i.e., the difference between the 10th and 90th percentile); or  $\beta_B$  is the regression estimate for the binary variable  $B$ .

## Results

We trained rats ( $n = 9$ ) to perform a decision-making task (Fig. 1A, B) in which they chose one of two targets (left or right lever) based on the reward size and effort cost (i.e., number of lever presses) associated with each target. The task consisted of four trial blocks: two “reward size blocks” in which either large or small sucrose reward was available while effort remained constant (Blocks A and B; Fig. 1A) and two “effort cost blocks” in which only large reward was available and either low or high effort was required (Blocks C and D; Fig. 1B). The order of blocks was varied from session to session. Each block of 40 free choice trials was preceded by 20 “forced choice” trials in which subjects were exposed to the current target-reward and target-effort contingencies.

Each trial began with an auditory cue and the extension of one lever (forced choice) or both levers (free choice; Fig. 1C). After the rat pressed either lever, the other lever was immediately retracted; the chosen lever was retracted after the required number of presses was completed. Reward was then delivered upon entry into the receptacle. Trials were presented after a variable ITI averaging 10 s. A video tracking system and head-mounted LEDs were used to monitor rats’ position within the chamber.



**Figure 2.** Rats exhibit impulsive responding at whichever lever is in close proximity. **A**, Distribution of a subject's head position at cue onset (blue dots) within the chamber during a single example session. Representation of the chamber is to scale. Scale bar, 10 cm. Black triangles represent positions of left and right lever. Magenta square represents location of reward receptacle. Shaded circles (radius 12.5 cm) represent the regions in which the subject was considered to be "near" one of the levers. In the overlapping shaded region, subjects were considered to be near the left lever when they were closer to the left lever than the right lever, and vice versa. **B**, Average percentage suboptimal choices (small reward or high effort) during each of the four blocks when rats were near the left lever (blue), near the right lever (green), or far from both levers (red) at cue onset.  $p \leq 0.001$ ,  $\dagger p \leq 0.01$  (corrected; Bonferroni *post hoc* test). n.s., Not significant. Suboptimal choices are most frequent when the rat is in close proximity to the suboptimal lever and least frequent when the rat is in close proximity to the optimal lever (associated with large reward or low effort). **C, D**, Average of median motion onset latencies in the three proximity conditions during reward size blocks (**C**) and effort blocks (**D**) for trials in which the rat was stationary at cue onset. Color conventions and symbols same as in **B**. **A–D**, Error bars indicate SEM. Trials in which the rat starts near the nonchosen lever (i.e., must overcome the propensity to choose a lever in close proximity) have the longest motion onset latencies.

### Proximate reward bias is a product of impulsive approach

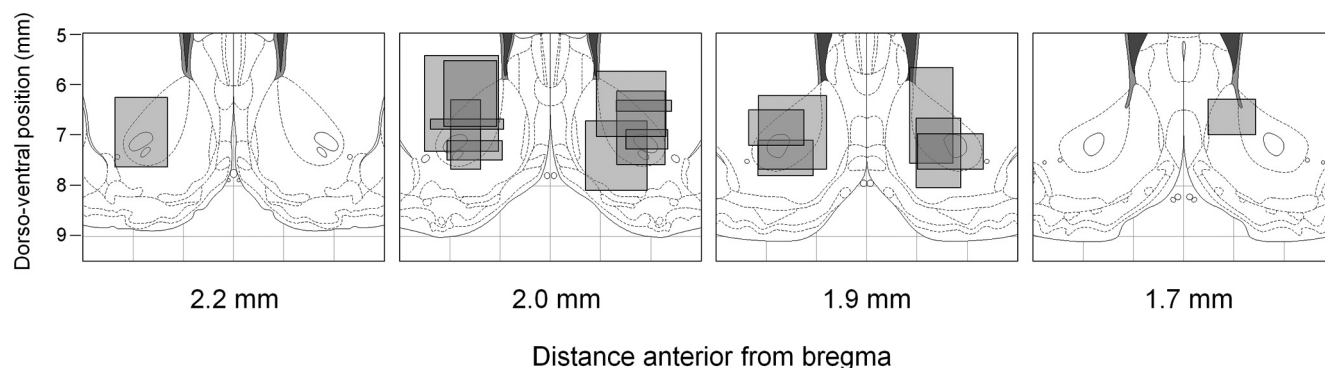
In each block, one lever yielded a more optimal outcome than the other (i.e., larger reward or lower effort). Rats made significantly more optimal than suboptimal choices (paired *t* test,  $p < 0.001$  in all blocks), even at the beginning of free choice blocks, demonstrating that their decisions were largely based on learning the reward size and effort cost associated with each lever (Fig. 1*D–F*). Surprisingly, however, a substantial minority of choices were suboptimal (mean small reward choices, 31%; mean high effort choices, 24%). Suboptimal choices occurred throughout the free choice block, even after performance reached a plateau, so they were unlikely to reflect incomplete learning of the contingencies (Fig. 1*F*). The task design suggests a different potential explanation for this effect: because the levers were extended noncontingently and at unpredictable times, the distance between the animal and each lever at trial onset varied across trials. We hypothesized that subop-

timal choices were due to proximate reward bias: a propensity to choose nearer reward-associated objects.

To test this hypothesis, we examined the likelihood of suboptimal choice as a function of distance from the left and right levers (Fig. 2*A, B*). As illustrated in Figure 2*A*, we operationally defined being "near" a particular lever as a head position within 12.5 cm radius of that lever (and closer to that lever than the other lever); all other trials (on average,  $25.6 \pm 1.8\%$ ) were categorized as "far." Rats chose the lever associated with small reward or high effort significantly more often when they were near that lever at the time of cue onset than when they were near the other lever (Fig. 2*B*;  $p < 0.001$  in all conditions, one-way ANOVA with Bonferroni *post hoc* test) or far from both levers ( $p < 0.001$  in all conditions). This suggests that many suboptimal choices were the result of impulsive action: rats simply approached the nearer lever, despite having learned the location of a better option not much farther away, as evidenced by their overall performance (Fig. 1*D–F*).

We considered the possibility that rats might have chosen the nearer lever because it took extra time and/or energy to respond on the farther lever. If this were the case, then proximate reward bias might have produced choices that appeared impulsive but actually resulted from a cognitive evaluation process involving the application of a steep discounting function. Arguing against this notion, rats made many choices of the high-effort lever even though it took more time (and likely energy expenditure) to perform 16 lever presses (average  $\pm$  SEM of session medians:  $3.1 \pm 0.13$  s) than to approach and press the low-effort lever ( $1.1 \pm 0.05$  s) when they were located near the high-effort lever at cue onset. Although we cannot exclude the possibility that suboptimal choices were the result of

a more complex cognitive process (e.g., perhaps rats apply different discounting functions to different forms of effort) the parsimonious interpretation is that suboptimal choices were truly impulsive. Moreover, rats made many more optimal choices (choices of the lever associated with large reward and/or low effort) when they were in close proximity to that lever, compared with when they were far from both levers ( $p < 0.001$  in all conditions except Block D). Rats' position at cue onset did not differ when the left versus right lever was associated with the optimal outcome (percentage of "near left" trials not different,  $p = 0.08$ , Wilcoxon; percentage of "near right" trials not different,  $p = 0.92$ ), implying that subjects did not predominantly use a strategy of standing near the optimal lever during the intertrial interval. These results suggest that a substantial fraction of both optimal and suboptimal choices are the result of impulsive approach to the nearer lever.



**Figure 3.** Location of recording arrays. Panels are coronal atlas sections (Paxinos and Watson, 1998) showing the approximate locations of recording arrays. Each gray box represents one array and was plotted on the atlas section nearest to the center of the array in the anteroposterior direction. The width of each box indicates the mediolateral extent over which the electrolytic lesions and/or electrode tracks were identifiable for each array. The height of each box indicates the dorsoventral distance over which the array was driven during the course of the experiment. Although some arrays are shown to project outside the NAc, we took care to ensure that only NAc neurons were included in the study.

Further supporting this idea, subjects were significantly slower to initiate locomotor approach to the chosen lever when they were located near the nonchosen lever at cue onset (Fig. 2C,D;  $p < 0.01$  for left side choices during effort cost blocks;  $p < 0.001$  in all other conditions, ANOVA with Bonferroni *post hoc* test). This suggests that rats must override a prepotent inclination to choose the nearer lever, a process that takes extra time, in order to approach the optimal lever (or, rarely, the suboptimal lever) when it is farther away. Meanwhile, compared with trials in which rats were near the subsequently chosen lever at cue onset, the latency to initiate movement when the subject was far from both levers was statistically identical (left side choices during effort cost blocks) or even slightly faster (all other conditions; Fig. 2C,D); this observation indicates that animals were no less motivated for reward when they were farther from the levers and that they remained actively engaged in the task regardless of lever proximity. Together, these data show that proximate reward bias arises from a mechanism separate from, and competitive with, that used to weigh the costs and benefits associated with different choice targets.

### Proximate reward bias may be driven by proximity encoding in the NAc

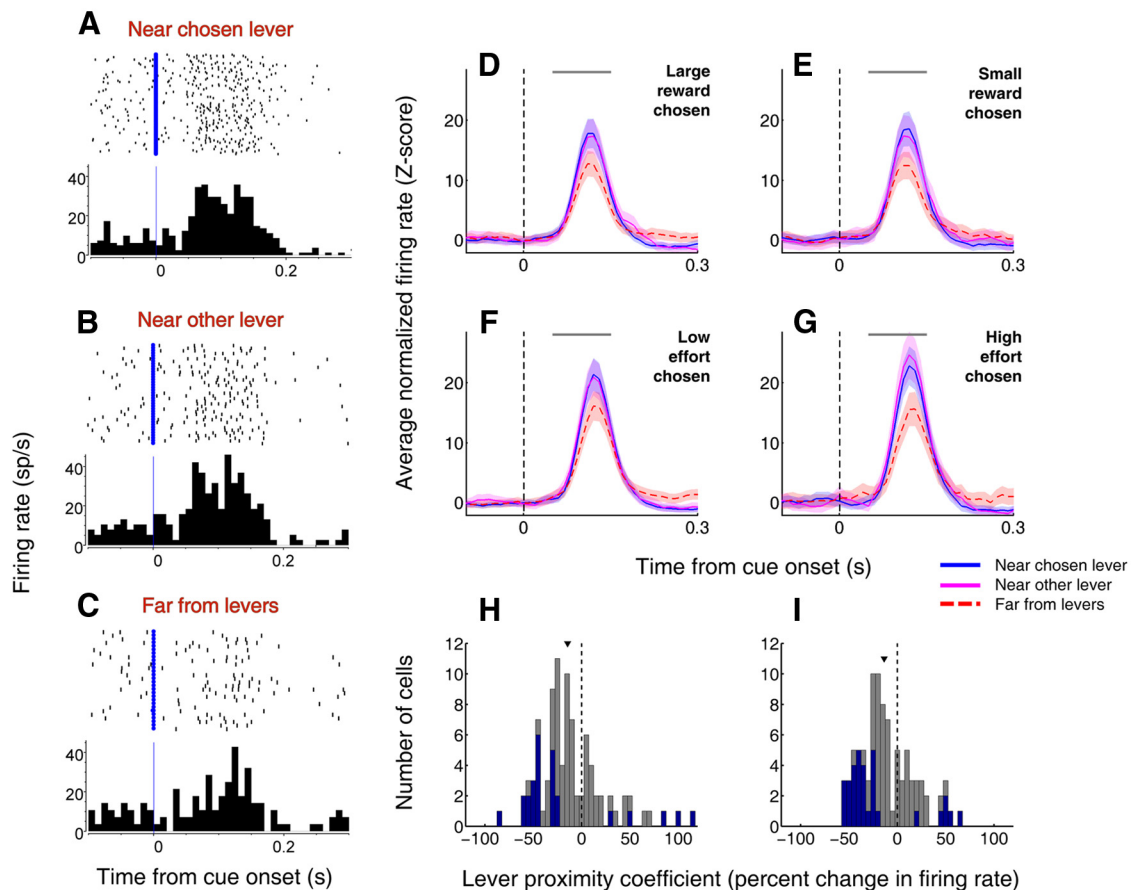
We next asked how proximity might be reflected in the activity of neurons in the NAc. We recorded from 159 neurons in the NAc during 49 behavioral sessions. Recording locations based on histological reconstruction are shown in Figure 3. Most neurons included in the dataset were located in the NAc core, but some were likely to have been located in one of the shell subregions; however, the firing characteristics of these cells were very similar to those putatively located in the NAc core.

Consistent with our previous observations (Nicola et al., 2004; Ambroggi et al., 2011), a large proportion of the neurons ( $n = 92$ , 58%) displayed fast phasic responses to the cues signaling the start of each trial. To characterize this activity, we focused our analyses on cue-evoked firing in a “decision window” from 0 to 200 ms after cue onset. We found that cue-evoked excitations during this time window strongly encoded proximity to the levers. An example of this effect is shown in Figure 4A–C. As was typical, this neuron exhibited stronger excitation in response to the cue when the animal was near either of the levers, including the nonchosen lever (Fig. 4B), at cue onset: activity in the decision window was significantly greater on “near chosen” and “near nonchosen” trials than on “far” trials (both comparisons,  $p <$

0.05, Wilcoxon) but did not differ between the two “near” conditions ( $p > 0.1$ , Wilcoxon).

Proximity was encoded in a similar manner by the population overall (Fig. 4D–G): regardless of which choice was ultimately made, we observed greater cue-evoked excitation on trials in which the rat started near a lever than those in which the rat was far from the levers at cue onset. We found that peak population activity on “near chosen” and “near nonchosen” trials was significantly greater than activity on “far” trials (all comparisons,  $p \leq 0.05$ , Wilcoxon). Meanwhile, population activity in the two “near” conditions was not significantly different ( $p > 0.05$ , Wilcoxon). Thus, population average activity in the NAc robustly reflected proximity to a reward-associated choice target whether or not that target was subsequently chosen.

Using a simpler cued approach task, we previously showed that two of the strongest contributors to cue-evoked activity are distance from the approach target at cue onset and latency to movement onset (McGinty et al., 2013). Therefore, to quantify the contribution of proximity to NAc cue-evoked signaling after accounting for effects of latency, we fit a GLM with regressors “distance from chosen lever” and “latency to maximum speed” to activity during the decision window. We used latency to maximum speed as a regressor because it correlates strongly with movement onset latency ( $r^2 = 0.40$ ,  $p < 0.001$ ) but can be calculated for all trials, including those in which the animal is already moving at cue onset; substitution of the direct regressor “latency to motion onset” produced similar results (data not shown). We then calculated a normalized “lever proximity coefficient” by expressing the regression estimate for distance as the expected change in firing rate resulting from an interdecile shift of the regressor (i.e., a change from the 10th to the 90th percentile; see Materials and Methods). We found that a large proportion of neurons ( $n = 46$ ; 50%) had a significant contribution of lever proximity to firing rate on one or both tasks. The proportion of cue-excited neurons with a significant contribution of lever proximity was at least 36% in each of the 9 subjects, indicating that the effect was not limited to specific individuals. Moreover, the resulting distributions were significantly shifted in the negative direction (reward size and effort blocks, medians  $< 0$ ;  $p < 0.01$ , sign rank test), indicating greater firing when the rat was nearer the chosen lever at cue onset (Fig. 4H,I). Using a regressor more representative of distance from both levers (distance from the receptacle) yielded nearly identical results (data not shown).



**Figure 4.** Cue-evoked excitations in the NAc are stronger when rats are located in close proximity to either lever upon cue onset. **A–C**, Example neuron that exhibits stronger cue-evoked firing when the rat is located near one of the levers (**A, B**) than when it is far from both levers (**C**) at cue onset. All trials from both reward size and effort blocks are shown in chronological order with the first trial on top. Blue lines indicate cue onset. **D–F**, Population peristimulus time histograms aligned on cue onset for trials in which the rat chooses the lever associated with large reward (**D**), small reward (**E**), low effort (**F**), or high effort (**G**). Blue solid line indicates rat located near chosen lever; magenta solid line indicates near nonchosen lever; red dashed line indicates far from both levers. Shading represents SEM. Gray bar represents window of analysis for comparisons; median activity in this window differs significantly between “near chosen” and “far” ( $p < 0.05$  for all conditions, Wilcoxon), and between “near nonchosen” and “far” ( $p < 0.005$  for all conditions) but does not differ between “near chosen” and “near nonchosen” ( $p > 0.05$  for all conditions). **H, I**, Distributions of lever proximity coefficient during reward size blocks (**H**) and effort cost blocks (**I**) derived from a GLM with factors of distance from chosen lever and latency to maximum velocity. The coefficient is the normalized  $\beta$  value for distance from the chosen lever. Blue indicates neurons with significant  $\beta$  values ( $p < 0.05$ ). Arrowhead indicates distribution median. In both plots, one outlier with normalized  $\beta$  value  $> 150$  is not shown. Both medians are significantly shifted to the negative ( $p \leq 0.01$ , sign rank test), indicating that cue-evoked activity tends to increase with decreasing distance from the chosen lever.

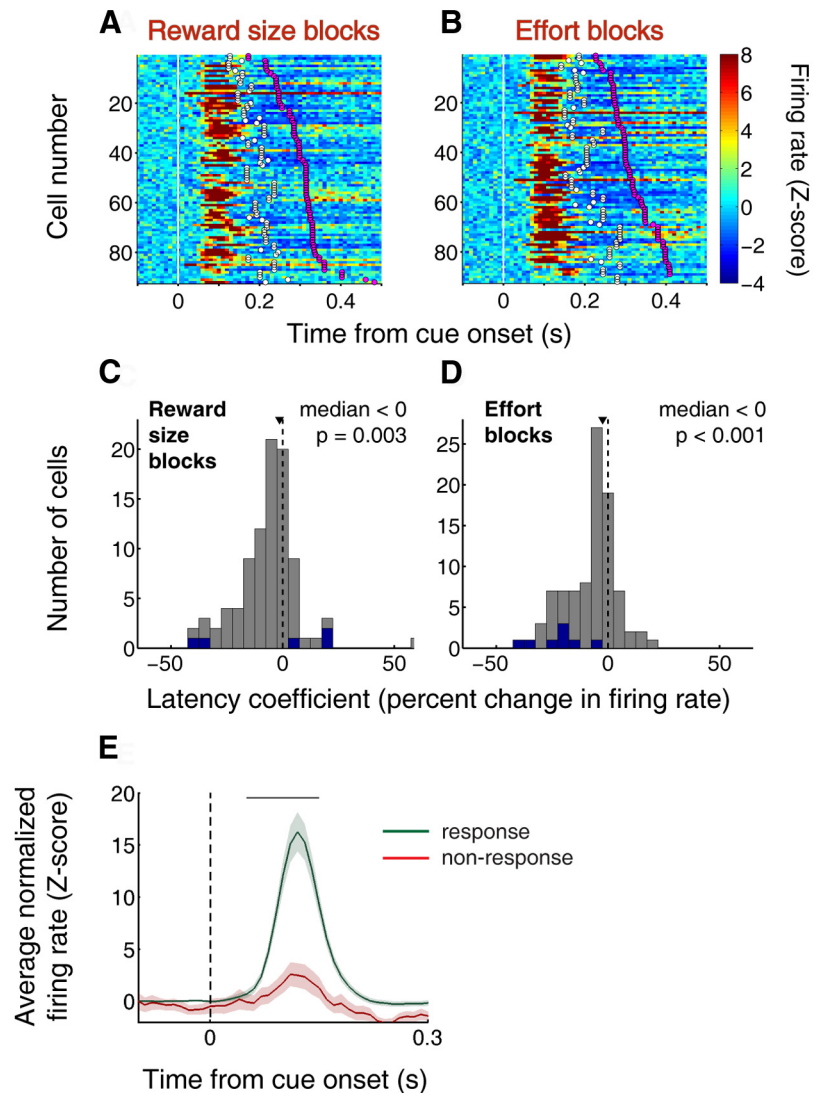
Importantly, both the onset and peak of cue-evoked responses occurred before the subject’s motion onset in the great majority of trials (Fig. 5*A, B*). The median latency for cue-evoked excitations (see Materials and Methods) was 60 ms; in contrast, on trials in which the subject was motionless at cue onset, the median motion onset latency averaged 313 ms, with an average lower quartile value of 204 ms. Neural activity following the rat’s onset of locomotion often featured an inhibition in firing below baseline, which occurred after the cessation of the excitatory response in many neurons (Fig. 5*A, B*, deep blue). We examined whether proximity encoding was also present in this time window by applying a similar GLM to the one used in Figure 4*H, I* to the 200 ms following motion onset. We found that many neurons had a positive proximity coefficient during this time period, in contrast to the negative coefficients predominating during the postcue time window, indicating a higher firing rate when the animal was located farther from the levers at cue onset. Indeed, the distribution of proximity coefficients (data not shown) was significantly shifted in the positive direction (reward size and effort blocks, medians  $> 0$ ;  $p < 0.01$ , sign rank test). Thus, stronger inhibitions, much like stronger excitations, were positively correlated with proximity to the levers.

Because they occur before motion onset, cue-evoked excitations, which robustly encode proximity to a lever, could contribute to driving subsequent approach behavior, perhaps including the choice of which lever to approach. In support of this notion, the magnitude of cue-evoked excitation was related to the prompt initiation of approach: based on the GLM described above, we calculated a “latency coefficient” (Fig. 5*C, D*) in a similar manner to the proximity coefficient (Fig. 4*H, I*). We found that 13% of cells ( $n = 12$ ) had a significant contribution of latency to firing rate on one or both tasks. Although fewer individual neurons exhibited a significant contribution of latency than of proximity, the resulting distributions were significantly shifted in the negative direction, indicating greater firing on trials with shorter latency to initiate approach to the lever. Furthermore, population cue-evoked activity was dramatically lower (indeed, excitation was virtually absent) on the small subset of trials (1.8%; mean = 4.8 per session) in which the animal did not press a lever in response to the cue (Fig. 5*E*). Overall, NAc cue-evoked excitations encode both proximity and latency/probability of approach to the levers, suggesting that greater excitation could promote impulsive approach to a nearby lever, resulting in proximate reward bias.

### NAc reward size and effort signals are unlikely to participate in value-based decision-making

NAc signals that drive impulsive approach might compete or cooperate with signals from other brain areas that promote “rational” choice behavior, such as the decision to approach the optimal lever even when it is farther away. Alternatively, in addition to encoding proximity, NAc cue-evoked excitations could contribute to nonimpulsive choice based on expected reward and/or effort. Therefore, we investigated whether these excitations encode reward size, effort cost, and/or target identity (i.e., left or right lever). Figure 6A–D shows an example of cue-evoked activity that was modulated by both reward size and target identity during the decision window. Before choices of the left lever, this neuron exhibited significantly greater cue-evoked firing when the lever was associated with large reward than with small reward (Fig. 6A,C;  $p < 0.05$ , Wilcoxon). No such differential firing was observed on trials in which the right lever was chosen (Fig. 6B,D;  $p > 0.1$ , Wilcoxon). Other neurons encoded effort cost, along with target identity, in a similar fashion. The neuron shown in Figure 6E–H exhibited significantly greater cue-evoked firing before choices of the right lever when it was associated with high effort than with low effort (Fig. 6F,H;  $p < 0.05$ , Wilcoxon), whereas no such differential firing was observed on trials in which the left lever was chosen (Fig. 6E,G;  $p > 0.1$ , Wilcoxon).

We noted that some neurons (e.g., Fig. 6A–D) displayed greater firing when the higher value option (large reward or low effort) was subsequently chosen, whereas others (e.g., Fig. 6E–H) responded more strongly before choices associated with small reward or high effort. To determine which of these response profiles predominated, we used ROC analysis (see Materials and Methods) to calculate a “reward preference index” and “effort preference index” for each neuron (separately for each lever). Index values  $>0.5$  indicate higher firing before large reward or high effort choices, whereas index values  $<0.5$  indicate higher firing before small reward or low effort choices. Many individual neurons had preference indices that were significantly different from 0.5 (permutation test,  $p < 0.05$ ): 36 neurons (39.1%) had at least one significant reward preference index, and 20 neurons (21.7%) had at least one significant effort preference index. In both cases, this is far more than would be expected by chance ( $p < 0.0001$ , binomial test). Surprisingly, however, we found that, over the population, approximately equal numbers of NAc neurons responded more strongly before choices of large reward and small reward, as well as low effort and high effort; for

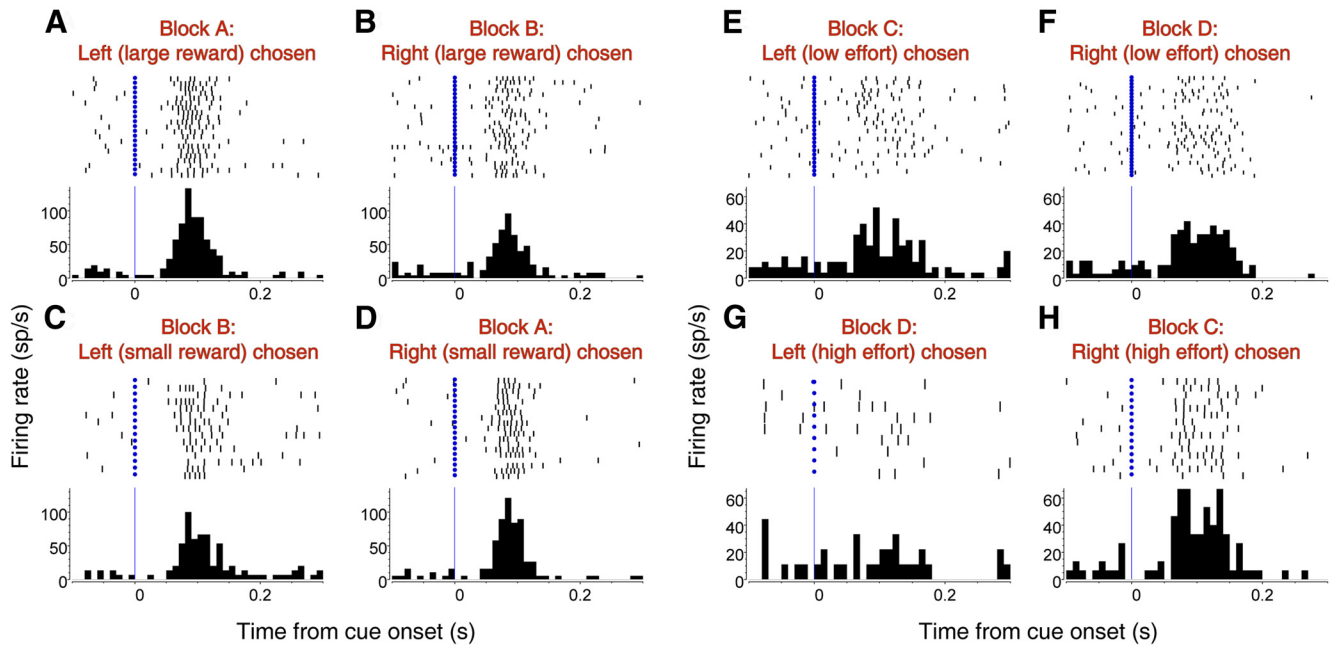


**Figure 5.** Cue-evoked excitations occur before motion onset and are related to response initiation. **A, B**, Each row represents a single neuron's average firing rate, aligned to cue onset, calculated in 10 ms bins and normalized to a 1 s precue baseline. All 92 cue-excited neurons are shown during reward size blocks (**A**) and effort cost blocks (**B**). Neurons are sorted by median time of rat's motion onset (magenta circles); white circles represent bottom quartile. Several consecutively displayed neurons have the same median and lower quartile motion onset values because they were recorded during the same behavioral session. Excitations almost always precede motion onset and are time-locked to the cue, rather than to the subject's movement. **C, D**, Results of the GLM shown in Figure 4H, I with respect to the factor “latency to reach maximum velocity.” The latency index is derived from the  $\beta$  value associated with the latency factor and normalized as the expected percentage change in firing rate over an interdecile shift. Blue indicates significant  $\beta$  values ( $p < 0.05$ ). Arrowhead indicates distribution median. Both medians are significantly shifted to the negative ( $p < 0.01$ , sign rank test), indicating that cue-evoked activity tends to increase with decreasing motion latency. Similar results were obtained by substituting the factor “latency to motion onset.” **E**, Neural activity for 40 cells recorded during sessions in which at least one nonresponse occurred. Activity is aligned on cue onset for all trials in which the subject made a response (pressed a lever at least once; green) or did not make a response (red). Shading represents SEM. Gray bar represents window of analysis for comparison; median activity in this window differs significantly between response and nonresponse trials ( $p < 0.001$ , Wilcoxon).

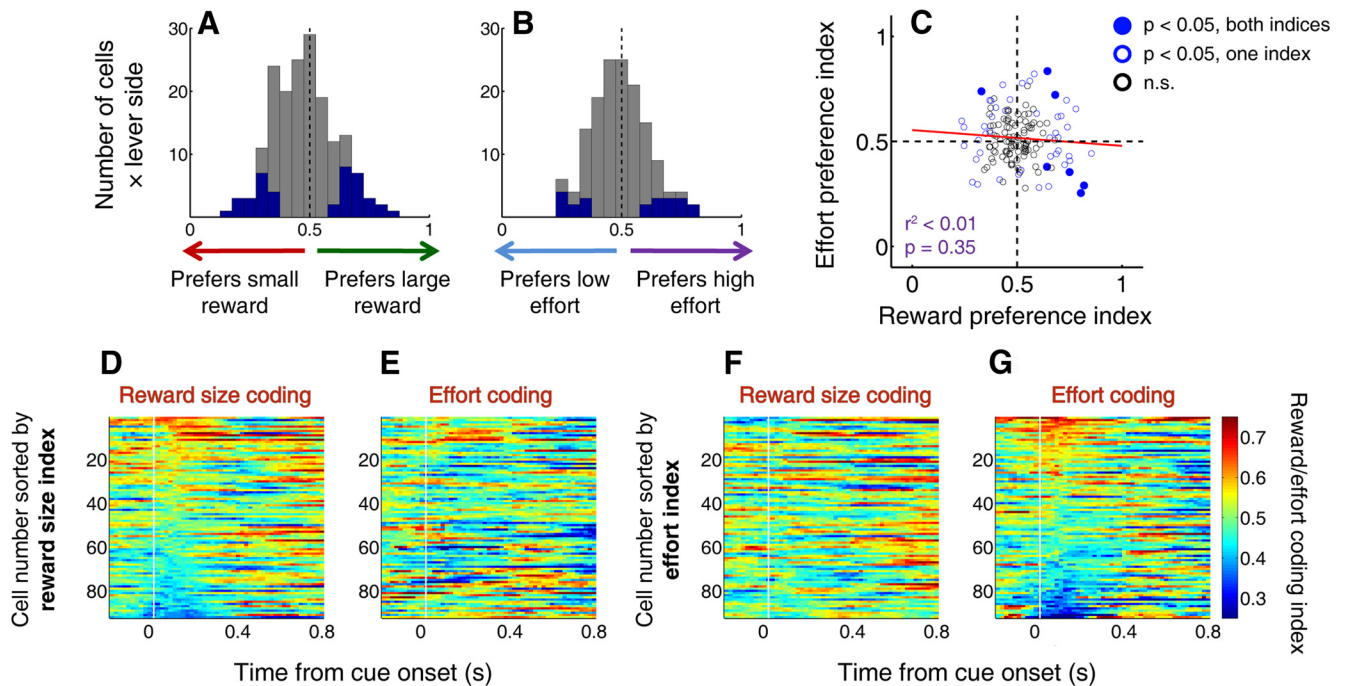
both reward and effort, the median preference index was not different from 0.5, indicating no overall preference (Fig. 7A,B;  $p > 0.1$ , sign rank test).

Next, we used the preference indices to determine whether neurons that fired more strongly before large reward choices were the same neurons that fired more strongly before low effort choices. Counter to our expectations, there was virtually no correlation between reward preference index and effort preference index (Fig. 7C): the reward value/target identity combination “preferred” by an individual neuron did not predict the preferred

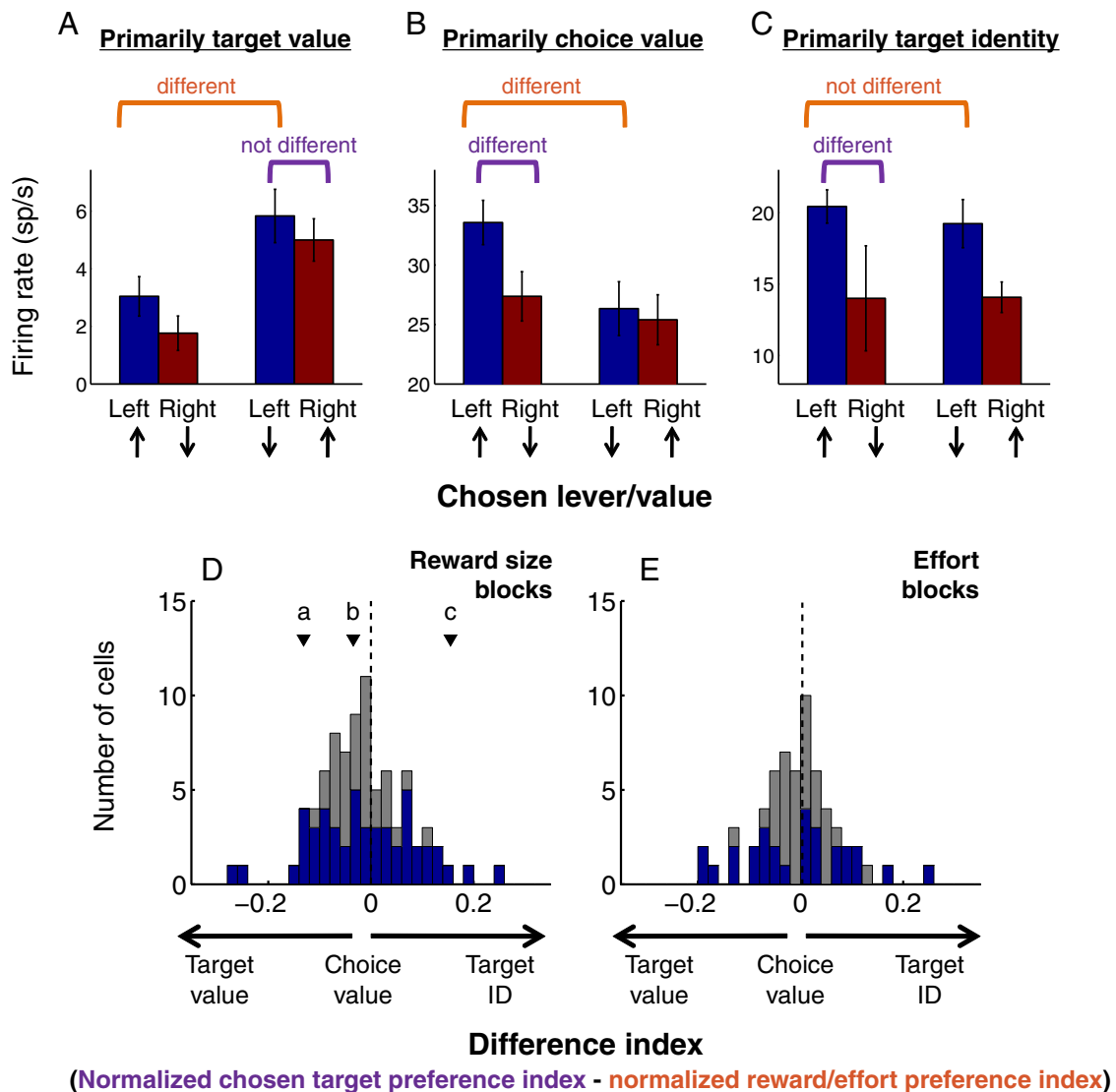




**Figure 6.** Cue-evoked activity of single neurons in the NAC may be modulated by expected reward size or effort cost. **A–D**, Example neuron that exhibits greater cue-evoked activity when the subject chooses to obtain a large reward by pressing the left lever (**A**) rather than a small reward by pressing the same lever (**C**). No significant difference in activity is observed before choices of the right lever (**B, D**). **E–H**, Example neuron that exhibits greater cue-evoked activity when the subject chooses to press the right lever 16 times (**H**) rather than one time (**F**) to obtain the same reward. No significant difference in activity is observed for choices of the left lever (**E, G**).



**Figure 7.** Reward size coding bears no clear relationship to effort coding among individual neurons. **A, B**, Distribution of reward preference index (**A**) and effort preference index (**B**) calculated using ROC analysis for each target (left or right lever) for 92 cue-excited cells. Preference index > 0.5 indicates greater activity preceding large reward choices (**A**) or high effort choices (**B**). Blue represents preference index significantly different from 0.5 ( $p < 0.05$ , permutation test). Overall, neither median is significantly shifted from 0.5 ( $p > 0.05$ , sign rank test), indicating a lack of preference for large or small reward, or for high or low effort, on the population level. **C**, Reward size preference index versus effort preference index for all cue-excited cells with sufficient data (at least 3 trials of each type). Solid blue dots indicate both indices significant ( $p < 0.05$ , permutation test); open blue dots indicate one index significant; open black dots indicate neither index significant. Red line indicates results of a linear regression. **D–G**, Time course of reward size preference index (**D, F**) or effort preference index (**E, G**), combined across targets, for each cue-excited neuron. Preference indices are calculated in 300 ms windows slid by 20 ms steps. Cells are sorted from top to bottom in decreasing order of average reward preference index (**D, E**) or effort preference index (**F, G**) over the first 200 ms (10 bins) after cue onset. Times are placed at the start of bins. Cells with strong reward size encoding (intense red or blue) do not systematically show strong effort encoding in the same direction (high or low value) and vice versa.



**Figure 8.** Population-wide, cue-evoked excitations do not track the value of specific action targets. **A–C**, Response profiles of individual neurons encoding primarily target value (**A**), choice value (**B**), or target identity (**C**) during reward size blocks. Up arrows indicate choice associated with large reward; down arrows indicate small reward. Average firing rates are obtained from the 200 ms post-cue onset time window preceding each choice type. Error bars indicate SEM. Orange bracket indicates reward preference index comparison; purple bracket indicates chosen target preference index comparison. As in Figure 7, preference indices are calculated using ROC analysis; “different” indicates  $p < 0.1$ , permutation test. **D, E**, Distribution of a “difference index” representing the relative encoding of reward size associated with a choice versus reward size associated with a specific target (**D**) or effort cost associated with a choice versus effort cost associated with a specific target (**E**). Preference indices are normalized by subtracting 0.5 and taking the absolute value; the difference index is then calculated by subtracting the normalized reward/effort preference index from the normalized chosen target preference index. For each cue-excited neuron, the difference index is calculated with respect to the preferred target only (e.g., for the neuron represented in **A**, the left lever). Thus, a neuron that encodes target value would have a strongly negative difference index; a neuron that encodes target identity would have a strongly positive difference index; and a neuron that encodes choice value would have a difference index near zero. Arrowheads labeled “a,” “b,” and “c” indicate the difference index associated with the neurons shown in **A–C**, respectively. Blue represents the subset of neurons with at least one significant preference index ( $p < 0.1$ , permutation test). For both this subset and the population as a whole, the median difference index is not different from zero ( $p > 0.1$ , sign rank test), indicating a lack of consistent encoding of target value, choice value, or target identity on a population level.

effort level/target identity combination, nor vice versa. Indeed, most neurons that strongly encoded reward size during the decision window did not significantly differentiate among effort levels; similarly, most neurons that encoded effort level did not encode reward size (Fig. 7C, open blue circles). Finally, we used a sliding ROC analysis to compare the trajectory of signals carrying information about reward size and effort over the course of the trial (Fig. 7D–G). We found that the time course of reward size coding bore no clear relationship to the time course of effort coding in the same neuron. Thus, reward size and effort level were encoded by largely nonoverlapping populations of neurons in the NAc.

We reasoned that these reward and effort signals might participate in weighing the value of different options only if they

encode the costs and/or benefits associated with a specific target (i.e., right or left lever), regardless of which option is subsequently chosen. This “target value” signaling is typified by the neuron shown in Figure 8A, which fired more strongly whenever the right lever was associated with large reward (or, alternatively, when the left lever was associated with small reward) whether or not the subject went on to select that lever. In contrast, signals that encode the reward or effort associated with the subsequent choice necessarily reflect the outcome of a decision that has already been made by an “upstream” brain area. This “choice value” signaling is typified by the neuron shown in Figure 8B (same as Fig. 6A–D), which fired more strongly when the left lever was associated with large reward, but only when that lever was

subsequently chosen. A third possibility is that neural signals might reflect which target was chosen without regard to the value of that option. This “target identity” signaling is typified by the neuron shown in Figure 8C, which exhibited greater firing whenever the left lever was chosen, regardless of whether it was associated with large or small reward.

To determine which of these three types of encoding predominated, we compared the reward or effort preference index for each lever (Fig. 7A,B) with a “chosen target preference index.” For each lever, we determined the neuron’s preferred reward size or effort level; then, within the block that included that lever/reward or lever/effort combination, we calculated the index by using ROC analysis to compare cue-evoked activity before the preferred choice or the other available choice. Both indices were normalized by subtracting 0.5 and taking the absolute value, resulting in an unsigned measure of difference. Thus, for a neuron that encodes target value (Fig. 8A), the reward preference index will be high (orange bracket), whereas the chosen target preference index will be low (purple bracket): activity is modulated only by the current value of one of the targets, not by the subsequent choice. For a neuron that encodes choice value (Fig. 8B), on the other hand, the chosen target preference index (purple bracket) will be high because activity is dependent on the subsequent choice.

We subtracted these two indices to obtain a “difference index” that describes the degree of choice versus target value encoding for each cue-excited neuron. Ideal encoding of target value yields a strongly negative difference index; ideal encoding of target identity, a strongly positive index; and ideal encoding of choice value, a difference index near zero. Across the population, choice and target value encoding varied along a spectrum, with a median difference index indistinguishable from zero (Fig. 8D,E; both blocks,  $p > 0.05$ , sign rank test). This remained true when we limited the analysis to cells with a strong reward/effort preference index or chosen target preference index ( $p \leq 0.1$ , permutation test; Fig. 8D,E, blue bars), indicating that the distribution is not skewed by non-task-modulated neurons. Overall, a plurality of cue-evoked responses were modulated by some aspect of the animal’s subsequent choice: reward size, effort level, and/or identity. Notably, only a small subset of neurons could be characterized as encoding target value. Therefore, the great majority of NAc cue-excited neurons did not represent reward size or effort in a way that could be used to weigh options against each other and thereby support adaptive decision-making.

### Proximity is encoded independently from reward size or effort

The infrequent encoding of target value (Fig. 8) contrasts with the robust encoding of proximity to choice targets (Fig. 4), suggesting that these cue-evoked excitations primarily promote approach toward nearby reward-associated targets. We reasoned that the NAc might support such approach by integrating the stimulus properties that influence the probability and vigor of approach, including proximity as well as associated reward size and/or effort cost. To test this hypothesis, we calculated a “lever proximity index” by using ROC analysis to compare cue-evoked firing on trials in which the rat was near a lever ( $< 12.5$  cm) or far from the levers ( $> 12.5$  cm) at cue onset (Fig. 2A). An index  $> 0.5$  indicates greater firing on “near” trials, whereas an index value  $< 0.5$  indicates greater firing on “far” trials. Consistent with the GLM results (Fig. 4H,I), the median lever proximity index was significantly  $> 0.5$  ( $p < 0.001$ , sign rank test) during both blocks (Fig. 9A,B, red arrowheads), and 43 neurons (46.7%) had a lever

proximity index that was significantly  $> 0.5$  ( $p < 0.05$ , permutation test) on one or both tasks. Using the lever proximity index, we then compared the direction and magnitude of proximity encoding with the direction and magnitude of reward size and effort encoding (Fig. 7A,B). We found that there was virtually no correlation between encoding of reward size or effort cost and lever proximity within the same neurons (Fig. 9A,B). Therefore, the activity of individual NAc neurons did not integrate proximity with the costs and benefits associated with a target.

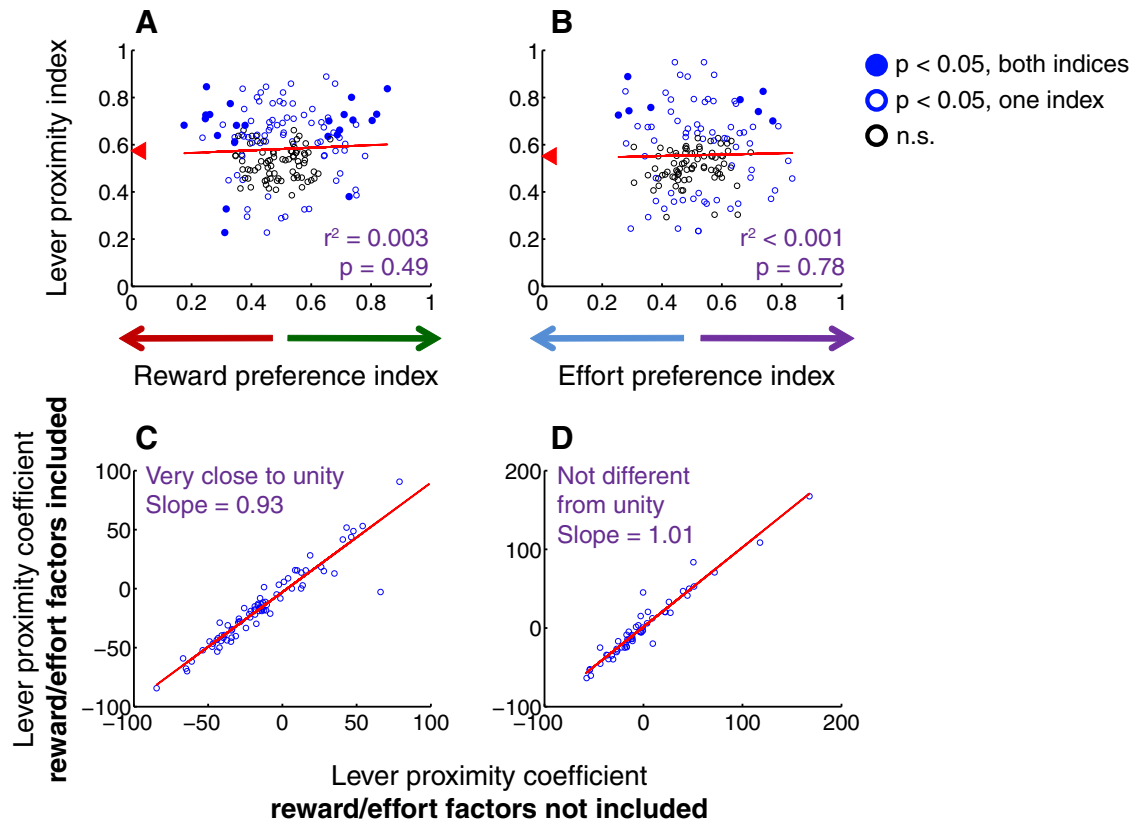
We also considered the notion that proximity-related variability in the cue-evoked response might be better explained by reward magnitude and/or effort cost. Therefore, we added factors representing reward size and effort level to the GLM described previously (Figs. 4H,I and 5C,D). The lever proximity coefficient remained virtually identical in strength and direction even when accounting for variability explained by reward size and effort (Fig. 9C,D). Thus, proximity encoding was not secondary to encoding of reward size or effort; rather, proximity and decision variables were represented simultaneously as independent parts of a multiplexed signal.

Next, the absence of a stable representation of expected value across contexts (Fig. 7C–G) led us to examine whether the representation of lever proximity was more stable than that of reward size and effort. Using the expanded GLM, we calculated a reward preference coefficient and effort preference coefficient for each neuron; these represent the expected change in firing rate resulting from a shift in chosen reward magnitude from small to large, or a shift in the chosen effort level from high to low. Consistent with the results shown in Figure 7C, there was no significant correlation between reward size preference coefficient and effort preference coefficient within the same neurons (Fig. 10A). In contrast, the lever distance coefficient was much more stable across reward size and effort blocks (Fig. 10B).

Finally, we asked whether the representations of reward size, effort, and lever proximity observed during free choice trials were present during forced choice trials, when only one option was available. Surprisingly, there was little correlation between reward size preference or effort preference during free choice trials and forced choice trials (Fig. 10C,E). Lever proximity, on the other hand, was encoded in a relatively consistent manner during free and forced choice trials (Fig. 10D,F). Overall, even though rats’ choices were clearly influenced by an understanding of reward and effort associated with each lever (Fig. 1D,E), representations of reward size and effort cost were highly variable depending on context, whereas proximity to choice targets was represented far more reliably. This is consistent with the idea that cue-evoked excitations in the NAc play a strong role in promoting impulsive approach to nearby targets, rather than participating in adaptive choice based on expected reward or effort cost.

## Discussion

Humans and animals tend to choose rewarding options that are nearer rather than more distant (Perrings and Hannon, 2001; Stevens et al., 2005; Kralik and Sampson, 2012; O’Connor et al., 2014). We investigated the neural basis of proximate reward bias using a decision-making task in which the subject’s distance from the choice targets (levers) was highly variable. Animals usually chose the optimal lever; however, when animals were near one of the levers, they tended to approach it regardless of costs and benefits. This cannot be explained by a straightforward discounting function because high effort choices required more time and, presumably, energy expenditure than low effort choices. Furthermore, the increased latency to approach a distant optimal lever



**Figure 9.** Proximity to lever is encoded independently from expected reward size or effort. **A, B**, Lever proximity index versus reward size preference index (**A**) or effort preference index (**B**) for each target among all cue-excited cells. Lever proximity index is calculated using ROC analysis to compare firing rates on trials in which the subject is located near either lever or far from both levers at cue onset. Index values  $>0.5$  indicate higher firing rate on “near” trials; values  $<0.5$  indicate higher firing rate on “far” trials. Red arrowheads indicate median lever proximity index. Reward and effort preference index are calculated as in Figure 7. Solid blue dots indicate both indices significant ( $p < 0.05$ , permutation test); open blue dots indicate one index significant; open black dots indicate neither index significant. Red represents regression lines. **C, D**, Results of a GLM that includes factors of “distance from chosen lever” with or without target identity and interaction between target identity and reward size (**C**) or effort level (**D**). Latency to maximum speed is also included as a factor in the model. One outlier (lever proximity coefficient =  $\sim 400$ ) is excluded from the regression in **C**. Note that, for nearly all cells, the lever proximity coefficient changes very little with inclusion of the other factors; regression line (red) is very close to unity.

(Fig. 2) suggests a cognitive evaluation process distinct from the faster noncognitive process that produces impulsive approach to proximate objects.

The activity of individual neurons in the NAc suggests a possible neural mechanism for proximate reward bias. Many neurons were excited by the cue signaling choice availability, and the excitations were consistently greater when animals were closer to the levers (Fig. 4). It is possible that these signals might actually encode related variables, such as the estimated time or effort requirement; however, because these variables are inextricably related to spatial proximity, “proximity” is a parsimonious descriptor of this encoding. This proximity signal was widespread and embedded within cue-evoked excitations that preceded movement onset and predicted movement onset latency (Fig. 5), suggesting that these excitations could drive impulsive approach to a nearby reward-associated choice target. Intriguingly, the proximity signal was equivalent regardless of which lever (optimal or suboptimal) was chosen (Fig. 4). An alternative task design (e.g., with levers that are farther apart) might have revealed differential encoding of proximity to the two options; however, this finding would not have fundamentally altered the potential role of the NAc in driving impulsive approach to the nearer lever.

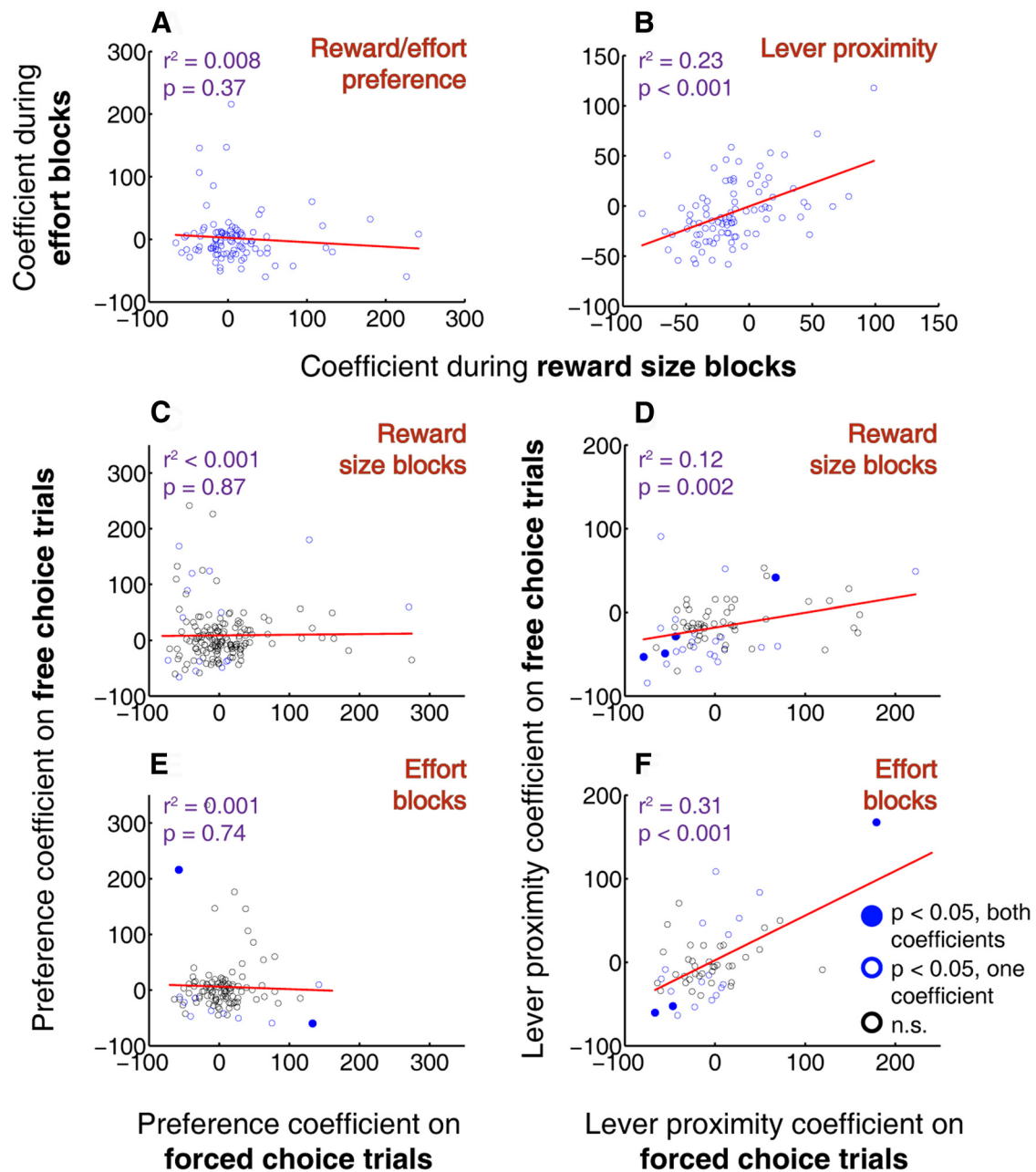
In contrast to their robust proximity encoding, we found little evidence that NAc neurons contribute to decision-making based on reward size or effort cost. Because population activity showed no preference for the higher value option (Fig. 7), downstream

brain areas cannot simply integrate this activity to guide animals toward optimal choices. Although it is possible that specific subsets of NAc neurons could contribute to this function, most cue-evoked activity did not track the reward or effort associated with specific choice targets, nor did it integrate the costs and benefits associated with either option (Figs. 7, 8, 9, and 10). Therefore, these signals lack the information needed for weighing options against each other.

Together, these results indicate that cue-evoked activity in the NAc is well suited for driving impulsive approach to proximate reward-associated objects but ill suited for supporting choice based on anticipated reward or effort. However, animals often did choose the optimal lever, particularly when they were far from both levers and NAc cue-evoked firing was at its lowest. Therefore, we propose that a separate, NAc-independent neural circuit weighs costs and benefits, promoting optimal target selection, and competes with the NAc proximity signal for control over the animal’s behavior.

#### Proximate reward bias as a form of impulsive approach

Dopamine in the NAc core is required for Pavlovian conditioned approach (PCA) to objects with incentive salience (Parkinson et al., 1999, 2002; Di Ciano et al., 2001; Saunders and Robinson, 2012). PCA does not depend on a neural representation of the anticipated outcome, but on the incentive value of the conditioned stimulus itself (Clark et al., 2012). This can lead to “para-



**Figure 10.** Proximity to lever, unlike reward size or effort, is encoded consistently across different task conditions. **A, B**, Reward and effort preference coefficients (**A**) and lever proximity coefficients (**B**) compared during reward size blocks versus effort blocks. Note the lack of correlation between the reward preference coefficient and effort preference coefficient, consistent with Figure 7C. **C–F**, Reward and effort preference coefficients (**C, E**) and lever proximity coefficients (**D, F**) compared during forced choice versus free choice trials within reward size blocks (**C, D**) or effort blocks (**E, F**). One outlier (lever proximity coefficient =  $\sim 400$  under all conditions) is excluded from the regression in **D** and **F**. Lever proximity coefficient is calculated as described previously; reward size and effort preference coefficients are based on the  $\beta$  values from each of the two interaction factors (e.g., reward size/left and reward size/right) normalized as the percentage change in firing rate associated with a shift in that binary factor (see Materials and Methods). Solid blue dots, both coefficients significant ( $p < 0.05$ ); open blue dots, one coefficient significant; open black dots, neither coefficient significant. Latency to maximum speed is also included as a factor in all models. Regression lines in red.

doxical” behavior (Dayan et al., 2006); for example, approach to a reward-associated object may persist even when the outcome is contingent on not approaching (Williams and Williams, 1969), or PCA can be directed toward objects associated with reward (“sign tracking”) rather than the location where reward is expected (“goal tracking”) (Flagel et al., 2010). Similarly, in the current task, subjects tended to approach a nearby rewarding object even if it was associated with a suboptimal outcome. These phenomena might share a neural mechanism: cue-evoked NAc activity that increases with greater proximity to the reward-associated object and drives approach behavior. Supporting this

notion, both sign tracking (but not goal tracking) (Flagel et al., 2011; Saunders and Robinson, 2012) and NAc cue-evoked excitations (Yun et al., 2004; du Hoffmann and Nicola, 2014) are dependent on the NAc’s dopaminergic input from the VTA.

Cue-evoked excitations in the NAc are ideally suited to drive approach to objects based on incentive salience: they occur at short latency after the object appears, precede the onset of approach and predict its latency, and do not reliably encode the expected value of the outcome. Moreover, cue-related firing is strong when the cue results in approach, but weak when the cue lacks sufficient incentive salience to do so (Nicola et al., 2004;

Ambroggi et al., 2011; McGinty et al., 2013). Robust encoding of proximity is consistent with this role because proximity enhances the salience of stimuli. Such encoding might be facilitated by a dopamine signal that increases with diminishing target distance (Gomperts et al., 2013; Howe et al., 2013).

Proximate reward bias could be considered a form of impulsive behavior. Researchers commonly divide impulsive behavior into two categories: impulsive action and impulsive choice (Winstanley et al., 2006; Bari and Robbins, 2013). In impulsive choice, which is insensitive to manipulations of NAc dopamine (Winstanley et al., 2005), subjects choose the less rewarding option when a more valuable option would require enduring a greater delay. Impulsive action, in contrast, comprises premature responding when the subject must withhold a response until reward is available (Bari and Robbins, 2013). Impulsive actions, which depend on NAc dopamine (Cole and Robbins, 1989; Economidou et al., 2012; Moreno et al., 2013), likely result from failure to suppress PCA to a reward-associated stimulus, even when the outcome is suboptimal; consistent with this idea, rats that are prone to sign tracking also exhibit increased impulsive action in these tasks (Flagel et al., 2010). Even though the current task involved decision-making, the subjects' impulsive behavior is better described by the classical definition of impulsive action: suboptimal choices likely resulted from inability to suppress PCA rather than from temporal discounting. Overall, our results suggest that impulsive action could generally be driven by proximity-modulated signals in the NAc that promote conditioned approach.

### Does the NAc participate in decision-making based on economic value?

Many studies have shown that NAc single-unit and BOLD responses are modulated by reward size, delay, or effort (Hollerman et al., 1998; Cromwell and Schultz, 2003; Setlow et al., 2003; Roitman et al., 2005; Kable and Glimcher, 2007; Gregorios-Pippas et al., 2009; Roesch et al., 2009; Day et al., 2011; Lim et al., 2011). Accordingly, some authors speculate that the NAc plays a role as a goal selector that encodes the value of conditioned stimuli in a common currency (Walton et al., 2006; Peters and Büchel, 2009; Mannella et al., 2013) or as the “critic” in an actor–critic model of decision-making (Khamassi et al., 2008; Roesch et al., 2009; van der Meer and Redish, 2011). Inconsistent with these roles, we found that NAc activity rarely encoded the value of specific targets, nor did it encode a consistent, integrated representation of the value of expected outcomes. Instead, we found that NAc neurons encoded in parallel aspects of the upcoming choice (location, reward size, and effort), apparently reflecting a decision made upstream. These results are consistent with observations that NAc activity was modulated by the value of upcoming choices rather than specific options (Ito and Doya, 2009; Kim et al., 2009; Roesch et al., 2009; Day et al., 2011) and with recent models emphasizing that NAc signals are context-dependent and represent many aspects of goals and actions (McDannald et al., 2011; Pennartz et al., 2011).

From these findings we conclude that, in this context, NAc cue-evoked activity does not directly contribute to a cognitive form of decision-making. This is supported by the finding that NAc dopamine depletion does not influence effort-based decision-making (Walton et al., 2009), although other studies report that manipulations of the NAc bias animals to choose lower effort options (Mingote et al., 2005; Salamone et al., 2007; Ghods-Sharifi and Floresco, 2010). However, in these studies, the position of the high-effort/high-reward target was invariant

across training, imbuing it with high incentive salience and likely leading to PCA. Locomotion toward freely available food, the usual low-effort alternative, does not rely on NAc dopamine (Baldo et al., 2002); but because PCA is highly dependent on NAc dopamine, manipulations of the NAc result in reduced approach to the high-effort/high-reward option, perhaps via suppression of the approach-promoting signals we have identified.

Prior studies of decision-making may not have identified proximity encoding in the NAc because, typically, subjects chose among a limited set of actions (“inflexible approach”) (Nicola, 2010). In contrast, our task required “flexible approach” (i.e., different trajectories to the chosen lever) which, unlike inflexible approach, is dependent on NAc dopamine (Nicola, 2010). This difference may explain reported preferences for larger reward or shorter delay in NAc activity (Cromwell and Schultz, 2003; Roesch et al., 2009; Goldstein et al., 2012): it is possible that the NAc participates in value-based decision-making only when selecting among a constrained set of actions. In contrast, our results indicate that, when animals must choose among targets to flexibly approach (as in the natural environment), NAc value encoding is unreliable. Likewise, proximate reward bias, along with NAc proximity encoding, might become apparent only when subjects must decide among flexible approach targets. Future research should consider the possibility that proximity encoding might play a strong, specific role in the selection of targets of conditioned approach.

### References

- Ambroggi F, Ghazizadeh A, Nicola SM, Fields HL (2011) Roles of nucleus accumbens core and shell in incentive-cue responding and behavioral inhibition. *J Neurosci* 31:6820–6830. [CrossRef Medline](#)
- Baldo BA, Sadeghian K, Basso AM, Kelley AE (2002) Effects of selective dopamine D1 or D2 receptor blockade within nucleus accumbens subregions on ingestive behavior and associated motor activity. *Behav Brain Res* 137:165–177. [CrossRef Medline](#)
- Bari A, Robbins TW (2013) Inhibition and impulsivity: behavioral and neural basis of response control. *Prog Neurobiol* 108:44–79. [CrossRef Medline](#)
- Clark JJ, Hollon NG, Phillips PE (2012) Pavlovian valuation systems in learning and decision making. *Curr Opin Neurobiol* 22:1054–1061. [CrossRef Medline](#)
- Cole BJ, Robbins TW (1989) Effects of 6-hydroxydopamine lesions of the nucleus accumbens septi on performance of a 5-choice serial reaction time task in rats: implications for theories of selective attention and arousal. *Behav Brain Res* 33:165–179. [CrossRef Medline](#)
- Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol* 89:2823–2838. [CrossRef Medline](#)
- Day JJ, Jones JL, Carelli RM (2011) Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *Eur J Neurosci* 33:308–321. [CrossRef Medline](#)
- Dayan P, Niv Y, Seymour B, Daw ND (2006) The misbehavior of value and the discipline of the will. *Neural Netw* 19:1153–1160. [CrossRef Medline](#)
- Di Ciano P, Cardinal RN, Cowell RA, Little SJ, Everitt BJ (2001) Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of pavlovian approach behavior. *J Neurosci* 21:9471–9477. [Medline](#)
- Drai D, Benjamini Y, Golani I (2000) Statistical discrimination of natural modes of motion in rat exploratory behavior. *J Neurosci Methods* 96: 119–131. [CrossRef Medline](#)
- du Hoffmann J, Kim JJ, Nicola SM (2011) An inexpensive drivable cannulated microelectrode array for simultaneous unit recording and drug infusion in the same brain nucleus of behaving rats. *J Neurophysiol* 106: 1054–1064. [CrossRef Medline](#)
- du Hoffmann J, Nicola SM (2014) Dopamine invigorates reward seeking by promoting cue-evoked excitation in the nucleus accumbens. *J Neurosci*, in press.
- Economidou D, Theobald DE, Robbins TW, Everitt BJ, Dalley JW (2012)

- Norepinephrine and dopamine modulate impulsivity on the five-choice serial reaction time task through opponent actions in the shell and core sub-regions of the nucleus accumbens. *Neuropsychopharmacology* 37: 2057–2066. [CrossRef Medline](#)
- Flagel SB, Robinson TE, Clark JJ, Clinton SM, Watson SJ, Seeman P, Phillips PE, Akil H (2010) An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: implications for addiction. *Neuropsychopharmacology* 35:388–400. [CrossRef Medline](#)
- Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H (2011) A selective role for dopamine in stimulus–reward learning. *Nature* 469:53–57. [CrossRef Medline](#)
- Gage GJ, Stoetzer CR, Wiltschko AB, Berke JD (2010) Selective activation of striatal fast-spiking interneurons during choice execution. *Neuron* 67: 466–479. [CrossRef Medline](#)
- Ghods-Sharifi S, Floresco SB (2010) Differential effects on effort discounting induced by inactivations of the nucleus accumbens core or shell. *Behav Neurosci* 124:179–191. [CrossRef Medline](#)
- Goldstein BL, Barnett BR, Vasquez G, Tobia SC, Kashtelyan V, Burton AC, Bryden DW, Roesch MR (2012) Ventral striatum encodes past and predicted value independent of motor contingencies. *J Neurosci* 32:2027–2036. [CrossRef Medline](#)
- Gomperts SN, Chen Z, Wilson MA (2013) VTA neuronal activity during learning in a spatial appetitive choice task. *Soc Neurosci Abstr* 39:771.09.
- Gregorios-Pippas L, Tobler PN, Schultz W (2009) Short-term temporal discounting of reward value in human ventral striatum. *J Neurophysiol* 101: 1507–1523. [CrossRef Medline](#)
- Hok V, Save E, Lenck-Santini PP, Poucet B (2005) Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. *Proc Natl Acad Sci U S A* 102:4602–4607. [CrossRef Medline](#)
- Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Neurophysiol* 80:947–963. [Medline](#)
- Howe MW, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM (2013) Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* 500:575–579. [CrossRef Medline](#)
- Ito M, Doya K (2009) Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J Neurosci* 29:9861–9874. [CrossRef Medline](#)
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633. [CrossRef Medline](#)
- Khamassi M, Mulder AB, Tabuchi E, Douchamps V, Wiener SI (2008) Anticipatory reward signals in ventral striatal neurons of behaving rats. *Eur J Neurosci* 28:1849–1866. [CrossRef Medline](#)
- Kim H, Suh JH, Huh N, Lee D, Jung MW (2009) Role of striatum in updating values of chosen actions. *J Neurosci* 29:14701–14712. [CrossRef Medline](#)
- Kralik JD, Sampson WW (2012) A fruit in hand is worth many more in the bush: steep spatial discounting by free-ranging rhesus macaques (*Macaca mulatta*). *Behav Processes* 89:197–202. [CrossRef Medline](#)
- Lansink CS, Jackson JC, Lankelma JV, Ito R, Robbins TW, Everitt BJ, Pennartz CM (2012) Reward cues in space: commonalities and differences in neural coding by hippocampal and ventral striatal ensembles. *J Neurosci* 32:12444–12459. [CrossRef Medline](#)
- Lee I, Griffin AL, Zilli EA, Eichenbaum H, Hasselmo ME (2006) Gradual translocation of spatial correlates of neuronal firing in the hippocampus toward prospective reward locations. *Neuron* 51:639–650. [CrossRef Medline](#)
- Lim SL, O'Doherty JP, Rangel A (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J Neurosci* 31:13214–13223. [CrossRef Medline](#)
- Mannella F, Gurney K, Baldassarre G (2013) The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. *Front Behav Neurosci* 7:135. [CrossRef Medline](#)
- McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G (2011) Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J Neurosci* 31:2700–2705. [CrossRef Medline](#)
- McGinty VB, Lardeux S, Taha SA, Kim JJ, Nicola SM (2013) Invigoration of reward seeking by cue and proximity encoding in the nucleus accumbens. *Neuron* 78:910–922. [CrossRef Medline](#)
- Mingote S, Weber SM, Ishiwari K, Correa M, Salamone JD (2005) Ratio and time requirements on operant schedules: effort-related effects of nucleus accumbens dopamine depletions. *Eur J Neurosci* 21:1749–1757. [CrossRef Medline](#)
- Moreno M, Economidou D, Mar AC, López-Granero C, Caprioli D, Theobald DE, Fernando A, Newman AH, Robbins TW, Dalley JW (2013) Divergent effects of D(2)/(3) receptor activation in the nucleus accumbens core and shell on impulsivity and locomotor activity in high and low impulsive rats. *Psychopharmacology* 228:19–30. [CrossRef Medline](#)
- Nicola SM (2010) The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *J Neurosci* 30:16585–16600. [CrossRef Medline](#)
- Nicola SM, Yun IA, Wakabayashi KT, Fields HL (2004) Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task. *J Neurophysiol* 91:1840–1865. [CrossRef Medline](#)
- O'Connor DA, Meade B, Carter O, Rossiter S, Hester R (2014) Behavioral sensitivity to reward is reduced for far objects. *Psychol Sci* 25:271–277. [CrossRef Medline](#)
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454. [CrossRef Medline](#)
- Parkinson JA, Olmstead MC, Burns LH, Robbins TW, Everitt BJ (1999) Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *J Neurosci* 19:2401–2411. [Medline](#)
- Parkinson JA, Dalley JW, Cardinal RN, Bamford A, Fehner B, Lachenal G, Rudarakanchana N, Halkerston KM, Robbins TW, Everitt BJ (2002) Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behav Brain Res* 137:149–163. [CrossRef Medline](#)
- Paxinos G, Watson C (1998) The rat brain in stereotaxic coordinates, fourth edition. New York: Academic.
- Pennartz CM, Ito R, Verschure PF, Battaglia FP, Robbins TW (2011) The hippocampal–striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci* 34:548–559. [CrossRef Medline](#)
- Perrings C, Hannon B (2001) An introduction to spatial discounting. *J Regional Sci* 41:23–38. [CrossRef](#)
- Peters J, Büchel C (2009) Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *J Neurosci* 29:15727–15734. [CrossRef Medline](#)
- Pfeiffer BE, Foster DJ (2013) Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497:74–79. [CrossRef Medline](#)
- Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G (2009) Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J Neurosci* 29:13365–13376. [CrossRef Medline](#)
- Roitman MF, Wheeler RA, Carelli RM (2005) Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* 45:587–597. [CrossRef Medline](#)
- Salamone JD, Correa M, Farrar A, Mingote SM (2007) Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology* 191:461–482. [CrossRef Medline](#)
- Saunders BT, Robinson TE (2012) The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. *Eur J Neurosci* 36:2521–2532. [CrossRef Medline](#)
- Setlow B, Schoenbaum G, Gallagher M (2003) Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* 38:625–636. [CrossRef Medline](#)
- Stevens JR, Rosati AG, Ross KR, Hauser MD (2005) Will travel for food: spatial discounting in two new world monkeys. *Curr Biol* 15:1855–1860. [CrossRef Medline](#)
- van der Meer MA, Redish AD (2009) Covert expectation-of-reward in rat ventral striatum at decision points. *Front Integr Neurosci* 3:1. [CrossRef Medline](#)
- van der Meer MA, Redish AD (2011) Ventral striatum: a critical look at models of learning and evaluation. *Curr Opin Neurobiol* 21:387–392. [CrossRef Medline](#)

- Walton ME, Kennerley SW, Bannerman DM, Phillips PE, Rushworth MF (2006) Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw* 19:1302–1314. [CrossRef Medline](#)
- Walton ME, Groves J, Jennings KA, Croxson PL, Sharp T, Rushworth MF, Bannerman DM (2009) Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-based decision making. *Eur J Neurosci* 29:1678–1691. [CrossRef Medline](#)
- Williams DR, Williams H (1969) Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *J Exp Anal Behav* 12:511–520. [CrossRef Medline](#)
- Winstanley CA, Theobald DE, Dalley JW, Robbins TW (2005) Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. *Neuropsychopharmacology* 30:669–682. [CrossRef Medline](#)
- Winstanley CA, Eagle DM, Robbins TW (2006) Behavioral models of impulsivity in relation to ADHD: translation between clinical and preclinical studies. *Clin Psychol Rev* 26:379–395. [CrossRef Medline](#)
- Young JJ, Shapiro ML (2011) Dynamic coding of goal-directed paths by orbital prefrontal cortex. *J Neurosci* 31:5989–6000. [CrossRef Medline](#)
- Yun IA, Wakabayashi KT, Fields HL, Nicola SM (2004) The ventral tegmental area is required for the behavioral and nucleus accumbens neuronal firing responses to incentive cues. *J Neurosci* 24:2923–2933. [CrossRef Medline](#)