

Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards

Matthew R. Roesch,¹ Teghpal Singh,² P. Leon Brown,² Sylvina E. Mullins,² and Geoffrey Schoenbaum^{1,3}

¹Department of Anatomy and Neurobiology, ²Program in Neuroscience, and ³Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland 21201

The ventral striatum (VS) is thought to serve as a gateway whereby associative information from the amygdala and prefrontal regions can influence motor output to guide behavior. If VS mediates this “limbic–motor” interface, then one might expect neural correlates in VS to reflect this information. Specifically, neural activity should reflect the integration of motivational value with subsequent behavior. To test this prediction, we recorded from single units in VS while rats performed a choice task in which different odor cues indicated that reward was available on the left or on the right. The value of reward associated with a left or rightward movement was manipulated in separate blocks of trials by either varying the delay preceding reward delivery or by changing reward size. Rats’ behavior was influenced by the value of the expected reward and the response required to obtain it, and activity in the majority of cue-responsive VS neurons reflected the integration of these two variables. Unlike similar cue-evoked activity reported previously in dopamine neurons, these correlates were only observed if the directional response was subsequently executed. Furthermore, activity was correlated with the speed at which the rats’ executed the response. These results are consistent with the notion that VS serves to integrate information about the value of an expected reward with motor output during decision making.

Introduction

The ventral striatum (VS) is thought to serve as a “limbic–motor” interface (Mogenson et al., 1980). This hypothesis has been derived primarily from the connectivity of this area with decision/motor-related areas including the prefrontal cortex, limbic-related areas including the hippocampus, amygdala, orbitofrontal cortex, and midbrain dopamine neurons, along with its outputs to motor regions, such as ventral pallidum (Groenewegen and Russchen, 1984; Heimer et al., 1991; Brog et al., 1993; Wright and Groenewegen, 1995; Voorn et al., 2004; Gruber and O’Donnell, 2009). Through these connections, the ventral striatum is thought to integrate information about the value of expected outcomes with motor information to guide motivated behavior. Consistent with this proposal, manipulations of VS impair changes in response latencies associated with different quantities of reward (Hauber et al., 2000; Gierler et al., 2003) and impact other measures of vigor, salience, and arousal thought to reflect the value of expected rewards (Berridge and Robinson, 1998; Cardinal et al., 2002a,b; Di Chiara, 2002; Nicola, 2007).

From these and other studies (Wadenberg et al., 1990; Ikemoto and Panksepp, 1999; Di Ciano et al., 2001; Di Chiara, 2002; Salamone and Correa, 2002; Wakabayashi et al., 2004; Yun et al., 2004; Gruber et al., 2009), it has been suggested that VS is indeed critical for motivating behavior in response to reward-predicting cues. However, there is ample contradictory evidence (Amalric and Koob, 1987; Cole and Robbins, 1989; Robbins et al., 1990; Reading and Dunnett, 1991; Reading et al., 1991; Brown and Bowman, 1995; Gierler et al., 2004) and little direct single-unit recording data from VS in tasks designed to directly address this question (Hassani et al., 2001; Cromwell and Schultz, 2003). Specifically, most VS studies have not varied both expected reward and response direction. Furthermore, no studies have examined how VS neurons respond when animals are making decisions between differently valued rewards, to assess the relationship between the cue-evoked activity and the decision.

To address these issues, we recorded from single neurons in VS while rats performed a choice task for differently valued rewards (Roesch et al., 2006, 2007a,b). On every trial, rats were instructed or chose between two wells (left or right) to receive reward. In different trial blocks, we manipulated the value of the expected reward by increasing either the delay to or size of reward (10% sucrose solution). Here we report that cue-evoked activity in VS neurons integrated the value of the expected reward and the direction of the upcoming movement. Increased firing required that the response be executed and was not observed if the reward was available but the animal chose to execute a different response. Furthermore, increased firing was correlated with the speed at which the rats’ executed that response. These results are consis-

Received June 2, 2009; revised Aug. 22, 2009; accepted Aug. 24, 2009.

This work was supported by National Institute on Drug Abuse Grants R01-DA015718 (G.S.) and K01DA021609 (M.R.R.) and National Institute on Aging Grant R01-AG027097 (G.S.).

Correspondence should be addressed to either Matthew R. Roesch, Department of Psychology, Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD 20742, E-mail: mroesch@psyc.umd.edu; or Geoffrey Schoenbaum, Department of Anatomy and Neurobiology, University of Maryland School of Medicine, 20 Penn Street, HSF-25251, Baltimore, MD 21201, E-mail: schoenbg@schoenbaumlab.org.

DOI:10.1523/JNEUROSCI.2572-09.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/2913365-12\$15.00/0

tent with the notion that VS serves to integrate information about the value of an expected reward with motor output during decision making.

Materials and Methods

Subjects. Male Long–Evans rats were obtained at 175–200 g from Charles River Laboratories. Rats were tested at the University of Maryland School of Medicine in accordance with School of Medicine and National Institutes of Health guidelines.

Surgical procedures and histology. Surgical procedures followed guidelines for aseptic technique. Electrodes were manufactured and implanted as in previous recording experiments. Rats had a drivable bundle of 10 25- μ m diameter. Iron–nickel–chrome wires (Stablohm 675; California Fine Wire) chronically implanted in the left hemisphere dorsal to VS ($n = 6$; 1.6 mm anterior to bregma, 1.5 mm laterally, and 4.5 mm ventral to the brain surface). Immediately before implantation, these wires were freshly cut with surgical scissors to extend ~ 1 mm beyond the cannula and electroplated with platinum (H_2PtCl_6 ; Aldrich) to an impedance of ~ 300 k Ω . Cephalexin (15 mg/kg, p.o.) was administered twice daily for 2 weeks postoperatively to prevent infection. Rats were ~ 3 months old at the time of surgery and were individually housed on a 12 h light/dark cycle; experiments were conducted during the light phase.

Behavioral task. Recording was conducted in aluminum chambers ~ 18 inches on each side with sloping walls narrowing to an area of 12×12 inches at the bottom. A central odor port was located above and two adjacent fluid wells on a panel in the right wall of each chamber. Two lights were located above the panel. The odor port was connected to an air flow dilution olfactometer to allow the rapid delivery of olfactory cues. Task control was implemented via computer. Port entry and licking was monitored by disruption of photobeams.

The basic design of a trial is illustrated in Figure 1. Trials were signaled by illumination of the panel lights inside the box. When these lights were on, nose poke into the odor port resulted in delivery of the odor cue to a small hemicylinder located behind this opening. One of three different odors was delivered to the port on each trial, in a pseudorandom order. At odor offset, the rat had 3 s to make a response at one of the two fluid wells located below the port. One odor (Verbena Olifac) instructed the rat to go to the left to get reward, a second odor (Camkol DH) instructed the rat to go to the right to get reward, and a third odor (Cedryl Acet Trubek) indicated that the rat could obtain reward at either well. Odors were presented in a pseudorandom sequence such that the free-choice odor was presented on 7 of 20 trials and the left/right odors were presented in equal numbers (± 1 over 250 trials). In addition, the same odor could be presented on no more than three consecutive trials. Odor identity did not change over the course of the experiment.

Once the rats were shaped to perform this basic task, we introduced blocks in which we independently manipulated the size of the reward delivered at a given side and the length of the delay preceding reward delivery. Once the rats were able to maintain accurate responding through these manipulations, we began recording sessions. For recording, one well was randomly designated as short (500 ms) and the other

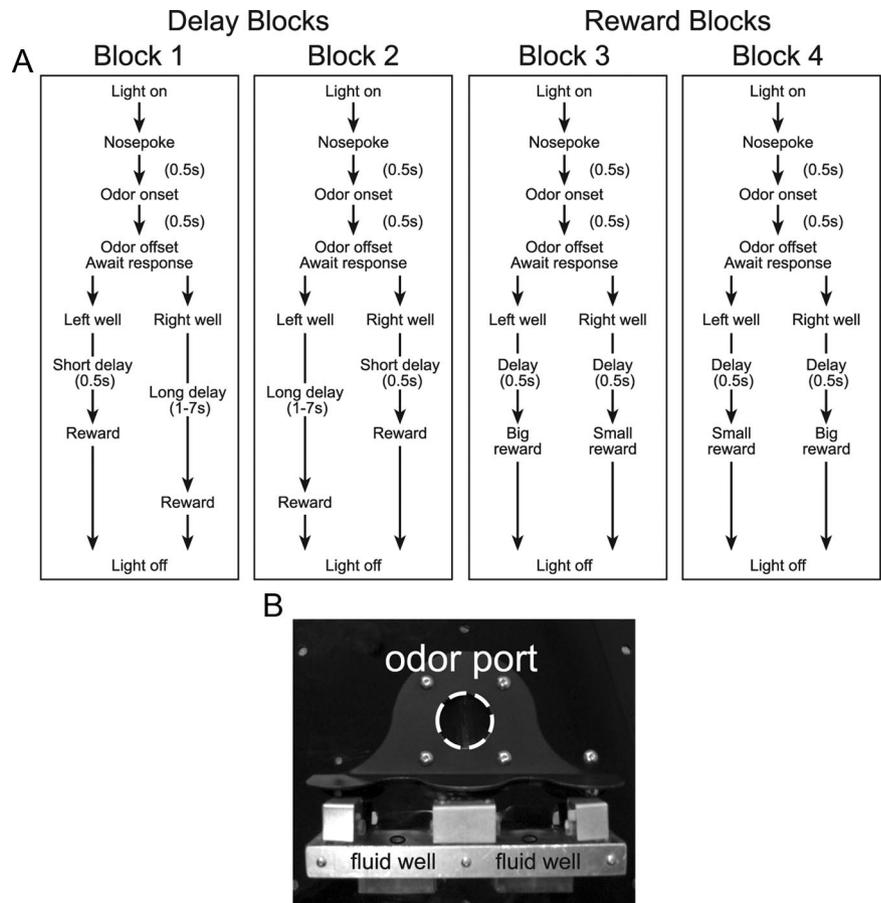


Figure 1. Task and behavior. **A**, Choice task during which we varied the delay preceding reward delivery and the size of reward. Figure shows sequence of events in each trial in four blocks in which we manipulated the time to reward or the size of reward. Trials were signaled by illumination of the panel lights inside the box. When these lights were on, nose poke into the odor port resulted in delivery of the odor cue to a small hemicylinder located behind this opening. One of three different odors was delivered to the port on each trial, in a pseudorandom order. At odor offset, the rat had 3 s to make a response at one of the two fluid wells located below the port. One odor instructed the rat to go to the left to get reward, a second odor instructed the rat to go to the right to get reward, and a third odor indicated that the rat could obtain reward at either well. One well was randomly designated as short and the other long at the start of the session (block 1). In the second block of trials, these contingencies were switched (Block 2). In later blocks (Blocks 3, 4), we held the delay preceding reward delivery constant while manipulating the size of the expected reward. **B**, Apparatus used to delivery odors and reward.

long (1–7 s) at the start of the session (Fig. 1A, Block 1). Rats were required to wait in the well to receive reward. In the second block of trials, these contingencies were switched (Fig. 1A, Block 2). The length of the delay under long conditions abided the following algorithm. The side designated as long started off as 1 s and increased by 1 s every time that side was chosen until it became 3 s. If the rat continued to choose that side, the length of the delay increased by 1 s up to a maximum of 7 s. If the rat chose the side designated as long as less than 8 of the last 10 choice trials, then the delay was reduced by 1 s to a minimum of 3 s. The reward delay for long forced-choice trials was yoked to the delay in free-choice trials during these blocks. In later blocks, we held the delay preceding reward delivery constant (500 ms) while manipulating the size of the expected reward (Fig. 1A). The reward was a 0.05 ml bolus of 10% sucrose solution. For big reward, an additional bolus was delivered after 500 ms. At least 60 trials per block were collected for each neuron. Rats were mildly water deprived (~ 30 min/d water *ad libitum*) with *ad libitum* access on weekends.

Single-unit recording. Procedures were the same as described previously (Roesch et al., 2006, 2007a). Wires were screened for activity daily; if no activity was detected, the rat was removed, and the electrode assembly was advanced 40 or 80 μ m. Otherwise, active wires were selected to be recorded, a session was conducted, and the electrode was advanced at the end of the session. Neural activity was recorded using two identical

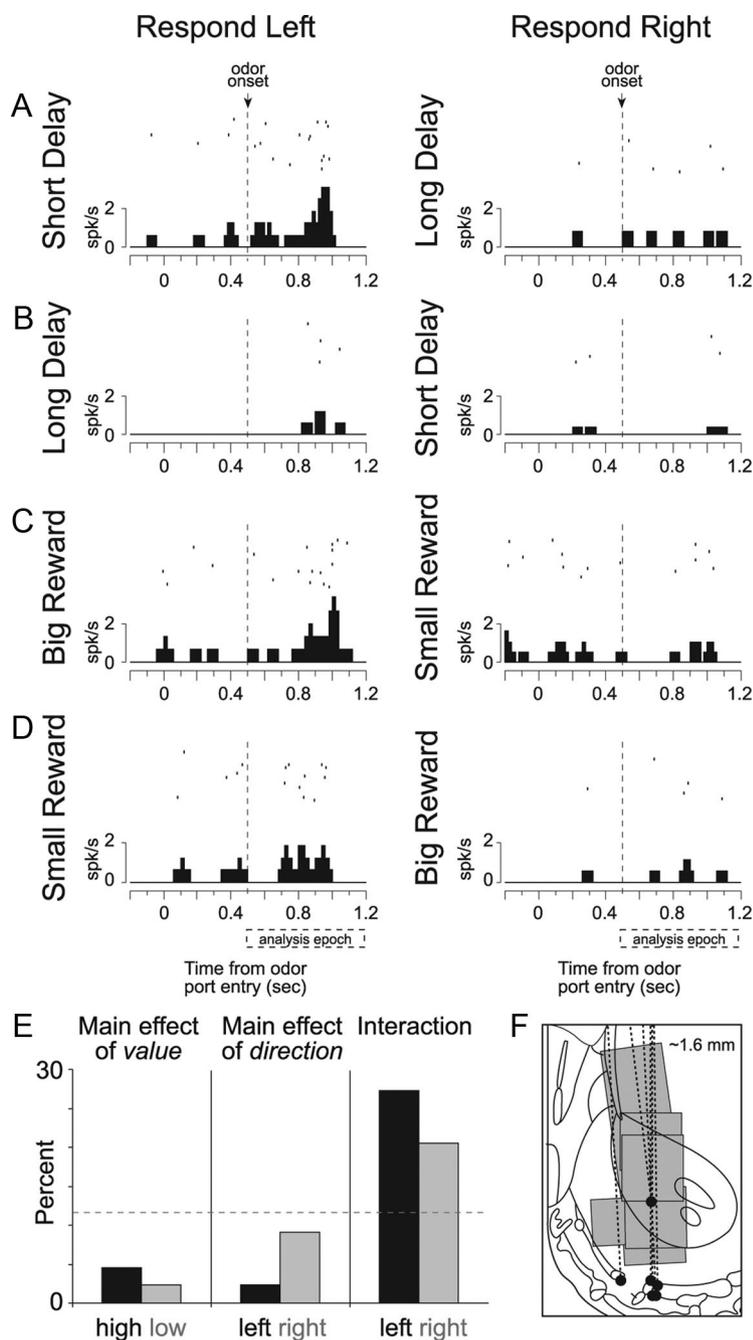


Figure 2. Activity of single neurons in VS reflect an interaction between expected value and direction. **A–D**, Activity of a single VS neuron averaged over all trials for each condition aligned on odor port entry during all eight conditions (4 rewards \times 2 directions). **E**, Results of a two-factor ANOVA with value and direction as factors ($p < 0.05$). Firing rate was taken from odor onset to port exit. The height of each bar indicates the percentage of odor-responsive neurons that exhibited a main effect of value, a main effect of direction, or an interaction between the value and direction. **F**, Boxes represent the extent of recording locations in six rats. Vertical bars on the drawing indicate the center of the electrode track in each rat; boxes indicate approximate extent of recording sessions vertically during transition through VS and give an estimate of lateral (and anteroposterior) spread of the wires (~ 1 mm).

Plexon Multichannel Acquisition Processor systems, interfaced with odor discrimination training chambers. Signals from the electrode wires were amplified 20 \times by an operational amplifier head stage (HST/8050-G20-GR; Plexon), located on the electrode array. Immediately outside the training chamber, the signals were passed through a differential preamplifier (PBX2/16sp-r-G50/16fp-G50; Plexon), in which the single-unit signals were amplified 50 \times and filtered at 150–9000 Hz. The single-unit signals were then sent to the Multichannel Acquisition Processor box, in which they were further filtered at 250–8000 Hz, digitized at 40 kHz, and amplified at 1–32 \times . Waveforms ($>2.5:1$ signal-to-noise) were extracted from active channels and recorded to disk by an asso-

ciated workstation with event timestamps from the behavior computer. Waveforms were not inverted before data analysis.

Data analysis. Units were sorted using Offline Sorter software from Plexon, using a template-matching algorithm. Sorted files were then processed in Neuroexplorer to extract unit timestamps and relevant event markers. These data were subsequently analyzed in Matlab (MathWorks). To examine activity related to the decision, we examined activity from odor onset to odor port exit. Wilcoxon’s tests were used to measure significant shifts from zero in distribution plots ($p < 0.05$). t tests or ANOVAs were used to measure within-cell differences in firing rate ($p < 0.05$). Pearson’s χ^2 tests ($p < 0.05$) were used to compare the proportions of neurons.

Results

Rats were trained on a choice task illustrated in Figure 1A (Roesch et al., 2006, 2007a). On each trial, rats responded to one of two adjacent wells after sampling an odor at a central port. Rats were trained to respond to three different odor cues: one odor that signaled reward in the right well (forced-choice), a second odor that signaled reward in the left well (forced-choice), and a third odor that signaled reward at either well (free-choice). Across blocks of trials, we manipulated value by increasing the length of the delay preceding reward delivery (Fig. 1A, Blocks 1, 2) or by increasing (Fig. 1A, Blocks 3, 4) the number of rewards delivered. Essentially, there were four types of rewards (short-delay, long-delay, big-reward, and small-reward) and two response directions (left and right), resulting in a total of eight conditions.

Rats’ behavior on both free- and forced-choice trials reflected manipulations of value. On free-choice trials, rats chose shorter delays and larger rewards over their respective counterparts (t test; $df = 119$; t values >16 ; p values <0.0001). Likewise, on forced-choice trials, rats were faster and more accurate when responding for a more immediate or larger reward (t test; $df = 119$; t values >9 ; p values <0.0001). Thus, rats perceived the differently delayed and sized rewards as having different values and were more motivated under short-delay and big-reward conditions than under long-delay and small-reward conditions, respectively.

We recorded 257 VS neurons across 75 sessions in six rats during performance of all four trial blocks. Recording locations are illustrated in Figure 2F. Because forced-choice trials present an evenly balanced neural dataset with equal numbers of responses to each well, we will first address our hypothesis by analyzing data from these trials. Thus, we will ask whether neural activity in VS neurons reflects value and direction of responding across blocks, particularly after learning (last 10 trials in each direction).

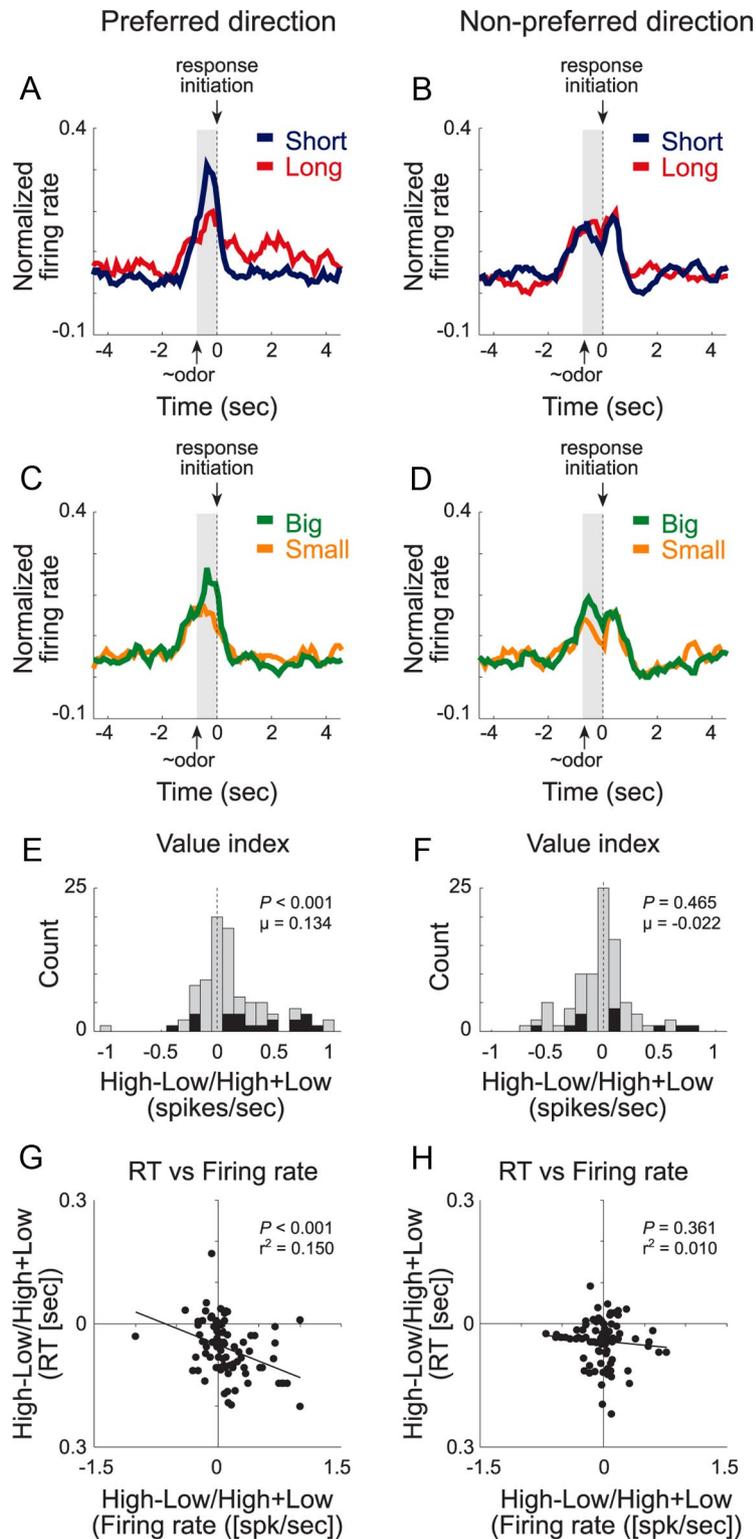


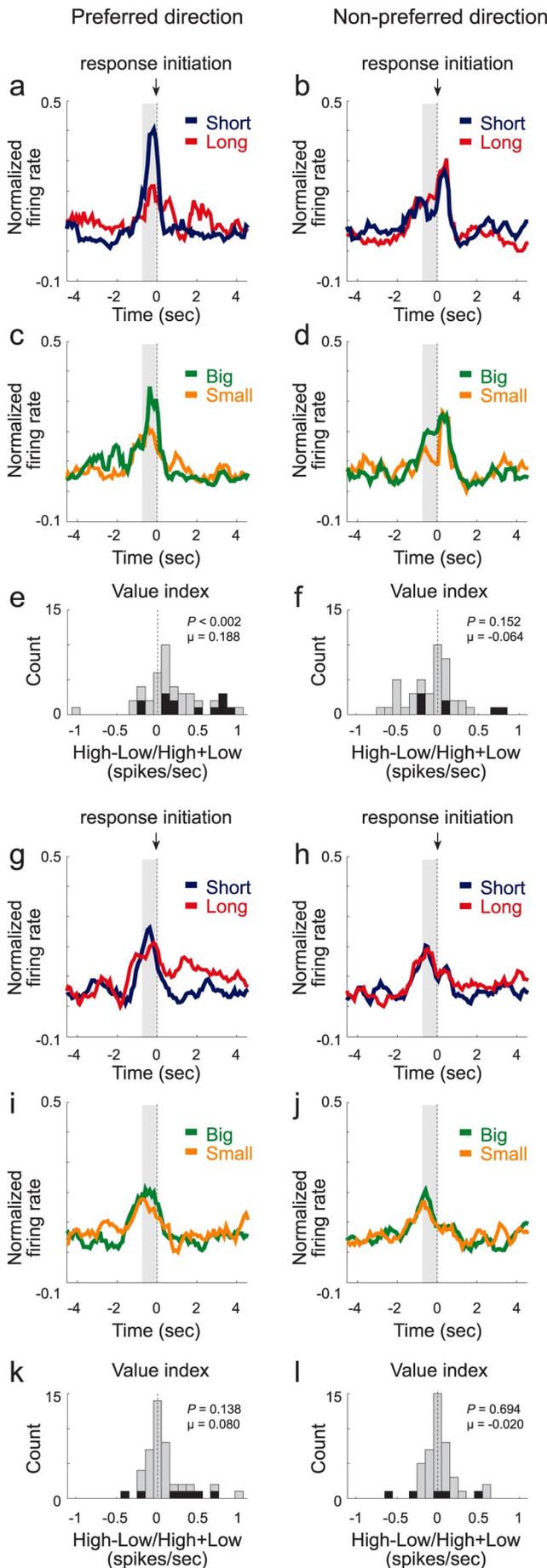
Figure 3. Population activity of odor-responsive neurons reflected motivational value and response direction on forced-choice trials. **A–D**, Curves representing normalized population firing rate during performance of forced-choice trials for the 44 odor-responsive neurons as a function of time under the eight task conditions (short, blue; long, red; big, green; small, orange). Data are aligned on odor port exit. Preferred and nonpreferred directions are represented in left and right columns, respectively. For each neuron, the direction that yielded the maximal response was designated as preferred. Population histograms were normalized by dividing all conditions by the maximum firing rate elicited among the eight conditions. To allow for qualitative comparisons across different sessions baseline firing rate (1 s before nose poke) was also subtracted out. All quantitative statistical analysis (i.e., distributions, counts of single cells, and correlations) were performed on raw firing rates. **E, F**, Distribution of value indices (delay index = short – long/short + long and reward index = big – small/big + small) in preferred (**E**) and nonpreferred (**F**) directions representing the difference in firing (odor on to port exit; gray bar) when high- and low-value outcomes were expected. For preferred and nonpreferred distributions, each cell contributed two counts: delay index and size index ($n = 88$). Significance

Activity in VS reflected the value and direction of the upcoming response

As has been reported previously (Carelli and Deadwyler, 1994; Nicola et al., 2004; Taha and Fields, 2006; Robinson and Carelli, 2008), many VS neurons were excited ($n = 44$; 17%) or inhibited ($n = 76$; 30%) during cue sampling (odor onset to port exit) versus baseline (1 s before nose poke; t test comparing baseline with cue sampling over all trials collapsed across condition; $p < 0.05$). An example of the former is illustrated in Figure 2A–D. Consistent with the hypothesis put forth in Introduction, activity of this neuron reflected the integration of associative information about the value of the reward predicted by the cue and the subsequent response. Thus, cue-evoked activity on forced-choice trials after learning was strongest for the cue that indicated reward in the left well, and this neural response was highest when value predicted for that well was high (on short and big trials). To quantify this effect, we performed a two-factor ANOVA with value and direction as factors during the last 10 forced-choice trials in each block ($p < 0.05$). Of the 44 cue-responsive neurons, 21 (47%) showed a similar significant interaction between direction and value. This count was significantly above chance given our threshold for statistical significance in our unit analysis (χ^2 test; $p < 0.0001$), and there was no directional bias to the left or right across the population (Fig. 2E) ($p = 0.98$). In contrast, and in keeping with the most rigorous account of the hypothesis that VS integrates value and direction information, only five (11%) showed a main effect of direction alone, and only three (7%) showed a main effect of value alone (Fig. 2E) (ANOVA; $p < 0.05$); these counts did not exceed chance (χ^2 test; p values > 0.05).

The overall effect is illustrated in Figure 3, which plots the average activity across all cue-responsive neurons on forced-choice trials during the last 10 trials for all eight conditions. For each cell, direction was referenced to its preferred response before averaging; thus, by definition, activity was higher in the preferred direction (left col-

←
measured by Wilcoxon's test. Black bars represent neurons that showed a significant difference (t test; $p < 0.05$) within single neurons. **G, H**, Correlations in the preferred (**G**) and nonpreferred (**H**) direction between value indices (short – long/short + long and big – small/big + small) computed for firing rate (during odor sampling) and reaction time (RT) speed at which rats exited the odor port after sampling the odor).



umn). Like the single-cell example, population activity during cue sampling was stronger in the preferred direction when value was high. That is, activity was stronger before a response in the preferred direction of the cell (left column) when the expected outcome was either a short delay (blue) or a large reward (green) compared with a long delay (red) or a small reward (orange), respectively. Notably, although activity in these populations did begin to increase during entry into the odor port, the difference in firing was only present during actual delivery of the odor (Fig. 3, gray shading).

Distributions of delay and size indices for each neuron, defined by the difference between high and low value divided by the sum of the two, are illustrated for each direction (preferred and nonpreferred) during the odor epoch (odor onset to port exit) in Figure 3, E and F. Only when value was manipulated in the preferred direction of the cell was the index significantly shifted above zero, indicating higher firing rates for more valued outcomes (Wilcoxon's test; $\mu = 0.134$; $z = 3.56$; $p < 0.001$). Cases in which neurons exhibiting stronger firing for high-value reward [$n = 16$ (18%)] outnumbered those showing the opposite effect [$n = 4$ (5%); χ^2 test; $p < 0.008$]. Neither the shift in the distribution nor the difference in number of cases in which activity was stronger for high or low value achieved significance in the non-preferred direction (Fig. 3F) (p values > 0.4).

Activity in VS was correlated with motivational level

VS is thought to motivate or invigorate behavior (Robbins and Everitt, 1996; Cardinal et al., 2002a). If the neural signal integrating value and directional response, identified above, relates to that function, then one should expect this activity to be correlated with the motivational differences between high- and low-value reward in our task. To address this question, we next examined the relationship between neural activity and reaction time (speed at which rat made the decision to move and exited the odor port). In previous sections, we showed that the reaction time was faster and activity was stronger (Fig. 3) when more valued reward (short delay and big reward) was at stake. To ask whether the two were correlated, we plotted neural activity (high - low/high + low) versus reaction time (high - low/high + low) independently for preferred and nonpreferred directions. We found that there was a significant negative correlation between the two in the preferred direction of the neuron (Fig. 3G) ($p < 0.001$; $r^2 = 0.150$). This relationship was not evident in the nonpreferred direction (Fig. 3H) ($p = 0.361$; $r^2 = 0.010$).

To examine this phenomenon more closely, we divided sessions into those with a strong versus a weak motivational difference between high- and low-value reward. According to the correlation described above, we would expect activity to be stronger for higher-value reward in sessions in which rats

Figure 4. Selectivity in VS reflected differences in motivation level. **A–D**, Curves representing normalized population firing rate ($n = 22$) during performance of forced-choice trials for those cells in which reaction time differences between high- and low-value trials were strong (short, blue; long, red; big, green; small, orange). **E, F**, Distribution of value indices (delay index = short - long/short + long and reward index = big - small/big + small) in preferred (**E**) and nonpreferred (**F**) directions representing the difference in firing (odor on to port exit; gray bar) when high- and low-value outcomes were expected in those cells illustrated in **A–D**. For preferred and nonpreferred distributions, each cell contributed two counts: delay index and size index. Black bars represent neurons that showed a significant difference (t test; $p < 0.05$) within single neurons (odor on to port exit; gray bar). **G–L**, Same as **A–F** but for those cells in which reaction time differences were weak ($n = 22$). For breakdown of sessions, see Results.

showed a strong difference between high- and low-value outcomes. To test this, we sorted sessions based on each rat's reaction time difference between high- and low-value trial types (small – big; long – short). In the top half of the distribution, the average reaction time on high- and low-value trials was 156 and 285 ms, respectively (t test; $df = 43$; $t = 17$; $p < 0.0001$), whereas in the lower half, reaction times on high- and low-value trials were 207 and 234 ms, respectively (t test; $df = 43$; $t = 5$; $p < 0.01$). Although both halves exhibited significant differences between high- and low-value outcomes, the differences were significantly larger in the top half (t test; $df = 43$; $t = 17$; $p < 0.0001$).

Remarkably, the neural signal identified above was only evident in sessions in which the rats were more strongly invigorated by high-value reward (Fig. 4A–D). This is illustrated in both delay and size blocks by higher firing rate during odor sampling for short-delay (blue) and big-reward (green) conditions over long-delay (red) and small-reward (orange) conditions, respectively. Value index distributions were significantly shifted above zero in the preferred direction in these sessions (Fig. 4E) (Wilcoxon's test; $\mu = 0.188$; $z = 3$; $p < 0.002$). In sessions in which rats were less concerned about the outcome (Fig. 4G–L), there was only a modest nonsignificant difference in activity in the preferred direction (Wilcoxon's test; $\mu = 0.080$; $z = 1$; $p = 0.138$).

Notably, the differences between sessions with strong and weak reaction time differences did not seem to reflect satiation, which has been shown to lead to slower overall reaction times (Holland and Straub, 1979; Sage and Knowlton, 2000). Overall speed of responding was not significantly different between sessions with strong and weak reaction time differences (220 vs 221 ms; t test; $df = 43$; $t = 0.02$; $p = 0.986$), and value correlates were no more likely to be observed early in a session versus late. The number of cells exhibiting value selectivity during the first two blocks of a session did not significantly differ from those observed during the last two blocks of a session (12 neurons or 27% vs 11 neurons or 25%; χ^2 ; $p = 0.85$).

Rats also appeared to learn the contingencies similarly in the two session types; rats chose the more valuable reward on 69% of trials (strong, 69.1%; weak, 69.4%; t test; $df = 43$; $t = 0.2$; $p = 0.852$). This indicates that latency differences did not reflect a learning effect. Together, these data suggest that differences in reaction time did not result from satiation or insufficient learning. Instead, when rats were goal oriented and strongly motivated by differences in expected value, activity in VS clearly reflected the animals' motivational output.

Activity in VS reflected the value of the decision

Up to this point, we have only analyzed forced-choice trials, in which odors instruct rats to respond to the left or the right well. We have assumed that this directional selectivity reflects the impending movement; however, directional selectivity might also

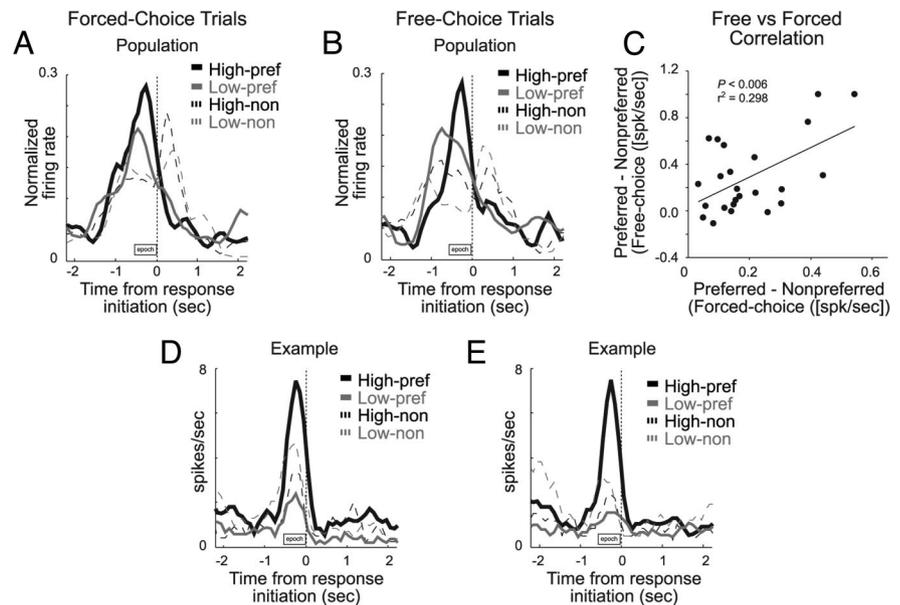


Figure 5. Activity on free- and forced-choice trials was positively correlated. **A, B,** Figures show average activity on forced- and free-choice trials, collapsed across value manipulation (high value, black; low value, gray), for preferred (solid) and nonpreferred (dashed) response directions. To control for learning, we only included trials after behavior reflected the contingencies in the current block (>50% choice of more valuable option). Furthermore, to control for the possibility that low-value choices might still be more frequent early during this block of trials, we paired each free-choice trial with the immediately preceding and following forced-choice trial of the same value. The line graphs show average activity from these trials on forced-choice (**A**) and free-choice (**B**) trials in each condition, aligned to odor port exit. **C,** Correlation between activity on free- and forced-choice trials. x -axis represents the difference between the trial type that elicited the strongest firing rate (preferred) and its counterpart (nonpreferred) in the same block of trials under forced-choice trials. y -axis represents same activity difference on free-choice trials. Thus, the response direction and the value of the expected reward are the same. The only difference between these two trial types was the odor that preceded the response. **D, E,** Single-cell example of firing on free- and forced-choice trials. Same conventions as in **A** and **B**.

represent the identity of the odor, regardless of whether or not that response is executed. This is because, on forced-choice trials, the odor and movement direction are confounded, because one odor means go right and the other means go left.

To address this issue, we compared activity on forced-choice trials with that on free-choice trials. This comparison can resolve this issue because, on free-choice trials, a different odor (rather than forced-choice) indicated the freedom to choose either direction (i.e., reward was available on each side). Moreover, rats chose the lower-value direction on a significant number of free-choice trials. Thus, by comparing firing on free- and forced-choice trials, we can disambiguate odor from movement selectivity. If the directional signal identified on forced-choice trials reflects only the impending movement, then it should be identical on free- and forced-choice trials, provided the rat makes the same response. Conversely, if the signal differs on free- and forced-choice trials when the rat makes the same response, then this would suggest that the proposed directional selectivity incorporates information about the sensory features of the odor.

For this analysis, we included all trials after learning (>50% choice performance) and collapsed across delay and size blocks. This procedure allowed us to increase our sample of low-value free-choice trials, which were sparse at the end of trial blocks. To further control for any differences that might arise during learning (rats typically chose low-value outcomes earlier on free-choice trials but were forced to choose low-value outcomes throughout the entire block on forced-choice trials), we paired each free-choice trial with the immediately preceding and following forced-choice trial of the same value.

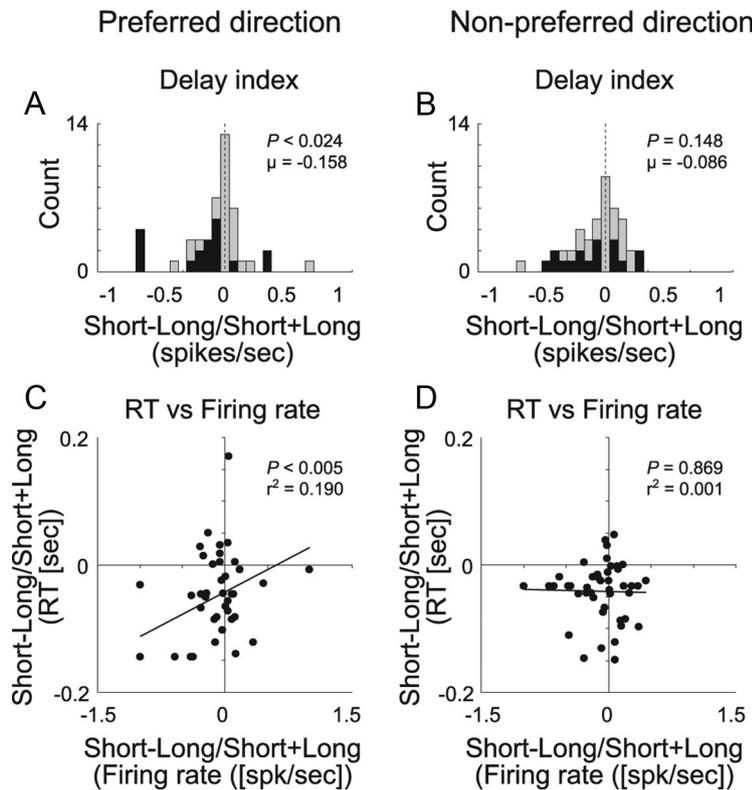


Figure 6. Activity after the behavioral response was stronger when rats expected the delay to be long. **A, B**, Distribution of delay indices (short – long/short + long) in preferred (**A**) and nonpreferred (**B**) directions representing the difference in firing (3 s after the response) when short- and long-value outcomes were expected. Significance measured by Wilcoxon's test. Black bars represent neurons that showed a significant difference (*t* test; $p < 0.05$) within single neurons. Data are aligned on reward delivery. **C, D**, Correlations between value indices (short – long/short + long) for firing rate (3 s after response) and reaction time (RT) (speed at which rats exited the odor port after sampling the odor). Preferred and nonpreferred directions are represented in left and right columns, respectively.

The results of this analysis are illustrated in Figure 5. Figure 5, **A** and **B**, represents the average activity over all neurons that showed a significant interaction between direction and value when rats responded in the preferred (solid) and nonpreferred (dashed) direction of the cell for high-value (black) and low-value (gray) outcomes during forced- and free-choice trials, respectively. As described previously, cue-evoked activity on forced-choice trials was stronger for high-value outcomes but only in one direction (Fig. 5A). Activity during free-choice trials showed exactly the same pattern. Thus, firing was higher on free-choice trials but only when the rat chose the high-value outcome and only when that outcome was in a particular direction (Fig. 5B). This is quantified in Figure 5C, which plots the difference between the preferred outcome/response (e.g., high-value-left) and nonpreferred outcome/response (e.g., low-value-right) of the cell on forced-choice trials (*x*-axis) versus the same calculation from data on free-choice trials (*y*-axis). By definition, values are all shifted above zero on the *x*-axis, because firing in these neurons was always higher for the preferred outcome/response on forced-choice trials. Importantly, values were also shifted above zero on free-choice trials (*y*-axis; Wilcoxon's; $\mu = 0.2879$; $z = 4$; $p < 0.001$). This indicates that neural activity was the same for a particular value and response, although the two trial types (free and forced) involved different odors (Fig. 5C). This pattern suggests that neural signals in VS neurons reflect the value of a particular yet-to-be-executed motor response and is not cue specific. This pattern also indicates that signaling in VS reflects the value of the response that is going to be executed, because firing

differed on free-choice trials when different responses were made, although the high-value reward was always available to be selected.

Activity after the decision was stronger in anticipation of the delayed reward

Lesions or other manipulations of VS make animals more likely to abandon a larger, delayed or higher-cost reward in favor of a smaller, more immediate or lower-cost reward (Cousins et al., 1996; Cardinal et al., 2001, 2004; Winstanley et al., 2004; Bezzina et al., 2007; Floresco et al., 2008; Kalenscher and Pennartz, 2008). These studies suggest that VS may be important for maintaining information about reward after the decision has been made. Consistent with this, we found that activity in the cue-responsive VS neurons described above was also elevated during the delay in our task, especially on correct trials. This is apparent in Figures 3 and 5, which show that activity was higher after the response in the preferred direction of the cell on long-delay (red) compared with short-delay (blue) trials. To quantify this effect, Figure 6, **A** and **B**, plots the distribution of delay indices (short – long/short + long) during the 3 s (minimum delay after learning) after the behavioral response in the preferred and nonpreferred direction of the cell. Delay indices were shifted significantly below

zero, indicating higher firing after responding to the delayed well (Wilcoxon's test; $\mu = -0.158$; $z = 2.2$; $p < 0.024$), and the counts of neurons exhibiting this pattern [$n = 24$ (55%)] significantly outnumbered those showing the opposite effect [$n = 6$ (14%)]. Notably, the increased firing after responding to the delayed well always preceded reward, because it occurred before the minimum delay after learning (3 s).

Interestingly, the difference in firing between short- and long-delay trials after the behavioral response was also correlated with reaction time (Fig. 6C) ($p < 0.005$; $r^2 = 0.190$). However, the direction of this correlation was the opposite of that between reaction times and cue-evoked activity described previously. Thus, slower responding on long-delay trials resulted in stronger firing rates after well entry and before reward delivery. If activity in VS during decision making reflects motivation, as we have suggested, then activity during this period may reflect the exertion of increased will to remain in the well to receive reward or expectation of reward, rather than signaling of other variables such as disappointment. Perhaps loss of this signal after lesions or inactivation of VS reduces the rat's capacity to maintain motivation toward the delayed reward. This suggests that it is necessary for VS to fire more in the delay to keep the rat in the well waiting for reward. Unfortunately, there were too few trials in which the rat left the fluid port prematurely to test this hypothesis.

Inhibitory responses in VS were not correlated with motivation

Finally, we asked whether the 76 neurons (30% of total neurons recorded) that were inhibited during odor sampling re-

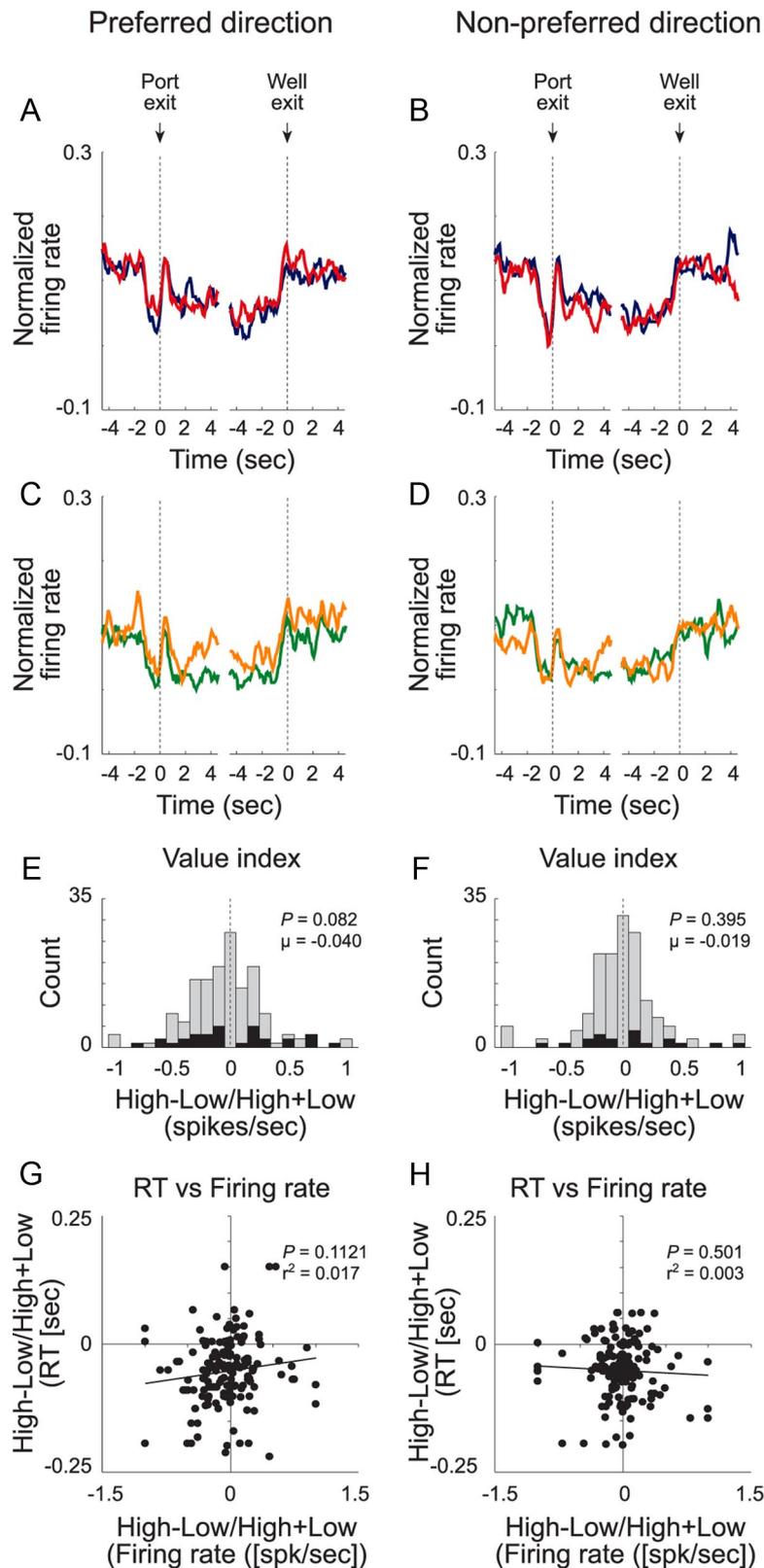


Figure 7. Neurons inhibited during the trial were not significantly modulated by motivation value. *A–D*, Curves representing normalized population firing rate during performance of forced-choice trials for the 76 neurons inhibited during odor sampling as a function of time under the eight task conditions (short, blue; long, red; big, green; small, orange). Data are aligned on odor port exit and well exit (trial end). Preferred and nonpreferred directions are represented in left and right columns, respectively. *E, F*, Distribution of value indices (short – long/short + long and big – small/big + small) in preferred (*E*) and nonpreferred (*F*) directions representing the difference in firing (odor on to port exit) when high- and low-value outcomes were expected. Significance measured by Wilcoxon's test. Black bars represent neurons that showed a significant difference (*t* test; $p < 0.05$) within single neurons. *G, H*, Correlations in the preferred (*G*) and nonpreferred (*H*) direction between value indices (short – long/short +

lected motivational value. Inhibitions in VS activity during performance of behavioral tasks have been described previously (Carelli and Deadwyler, 1994; Nicola et al., 2004; Taha and Fields, 2006; Robinson and Carelli, 2008) and might reflect the inhibition of inappropriate behaviors during task performance (i.e., leaving odor port or fluid well early), which might be more critical when a better reward is at stake. Here we address whether or not these neurons were modulated by expected reward value.

The average firing rates over these neurons are illustrated in Figure 7, *A* and *B*. As defined in the analysis, activity was inhibited during odor sampling. As the rat moved down to the well, activity briefly returned to baseline but then quickly returned to an inhibited state after entering the well and then subsequently returned to baseline during well exit. As for excitatory neurons, we asked whether the motivational level of the animal modulated neural firing in these neurons. Distributions of value indices were not significantly shifted from zero (Fig. 7*E, F*) (Wilcoxon's test; z values < 2 ; p values > 0.082), and approximately equal numbers of neurons fired more strongly and weakly for high-value reward (Fig. 7*E, F*, black bars). Furthermore, activity in these neurons was not correlated with reaction time. Thus, inhibitions observed during task performance were not modulated by value as observed for excitations.

VS activity during reward delivery was not modulated by unexpected reward

Previously, in rats performing this same task, we have shown that dopamine neurons fire more strongly at the beginning of trial blocks when an unexpected reward was delivered and less strongly in trial blocks when an expected reward was omitted (Roesch et al., 2007a). Such activity is thought to represent bidirectional prediction error encoding.

Of the sample of 257 VS neurons, activity in 41 neurons was responsive to reward delivery [*t* test comparing baseline with reward delivery (1 s) over all trials collapsed across condition; $p < 0.05$]. Of those, 12 were also cue responsive as defined above. Analysis of prediction errors revealed that few VS neurons seem to signal errors in reward prediction. For exam-

←
 long and big – small/big + small) computed for firing rate (during odor sampling) and reaction time (RT) (speed at which rats exited the odor port after sampling the odor).

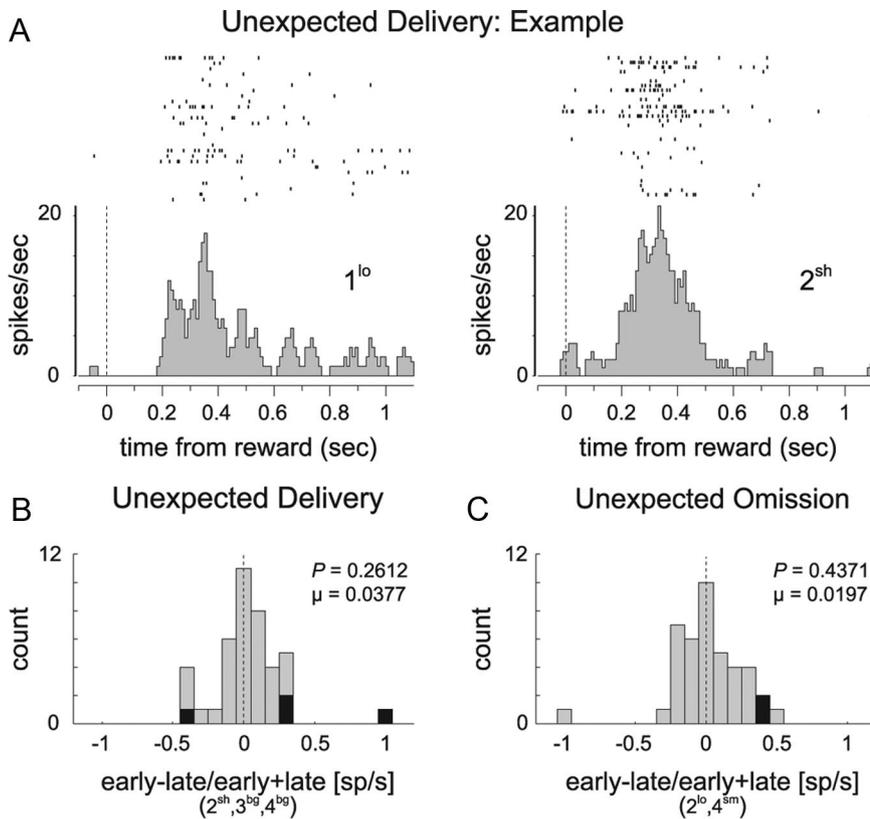


Figure 8. VS activity during reward delivery was not modulated by unexpected reward. **A**, Activity of a single VS neuron, aligned on reward delivery for long-delay trials in the first block (left) and short-delay trials from the second block (right). Rasters indicate action potentials and are present in order performed. Reward delivery at the start of short delay trials in block 2 were unexpected. **B**, Distribution of indices (early – late/early + late) representing the difference in firing to reward delivery (500 ms) during the first 5 (early) and last 15 (late) trials in trial blocks 2–4 (Fig. 1). In trial block 2, reward was delivered earlier than expected (2^{sh}), and, in trial blocks 3 and 4, reward was bigger than expected (3^{bg} and 4^{bg}). **C**, Distribution of indices (early – late/early + late) representing the difference in firing to reward omission (500 ms) during the first 5 and last 15 trials in trial blocks 2^{lo} and 4^{sm} in which reward was unexpectedly delayed (lo, long) or was smaller (sm, small). Analysis and figures shown here include those neurons that fired more strongly during reward delivery (1 s after reward) compared with baseline and included both free- and forced-choice trials. *p* values are derived from Wilcoxon's test. Filled bars indicated significance at the single-cell level (first 5 versus last 15 trials; *t* test; *p* < 0.05).

ple, the single cell illustrated in Figure 8A fired more strongly when reward was delivered unexpectedly; firing was maximal immediately after a new reward was instituted and diminishing with learning. However, this example was the exception rather than the rule. This is illustrated across the population in Figure 8, B and C, which shows the contrast in activity (early vs late) for all of the reward-responsive VS neurons ($n = 41$). This contrast is plotted separately for blocks involving unexpected delivery and omission of reward. Neither distribution was shifted significantly above zero, indicating no difference in firing early, after a change in reward, compared with later, after learning (Fig. 8B,C) (Wilcoxon's test; *z* values < 2; *p* values > 0.2610).

Discussion

Here we show that single neurons in VS integrate information regarding value and impending response during decision making and influence the motivational level associated with responding in a given direction. Cues predicting high-value outcomes had a profound impact on behavior, decreasing reaction time and increasing accuracy. This behavioral effect was correlated with integration of value and impending response during cue sampling in VS neurons. This result is broadly consistent with proposals that VS acts as a limbic–motor interface (Mogenson et al., 1980) and with a number of recent reports showing that VS signals

information about impending outcomes at the time a decision is made (Carelli, 2002; Setlow et al., 2003; Janak et al., 2004; Nicola, 2007; Ito and Doya, 2009; van der Meer and Redish, 2009).

Although these results are correlational in nature, they are in agreement with results from several studies in which pharmacological methods were used to show a more causal relationship between VS function and behavior (Berridge and Robinson, 1998; Cardinal et al., 2002a; Nicola, 2007). One set of studies in particular examined the impact of several different VS manipulations on rats' latencies to respond for different quantities of reward (Hauber et al., 2000; Gierler et al., 2003). In this simple reaction time task, discriminative stimuli presented early in each trial predicted the magnitude of the upcoming reward. As in our task, rats were faster to respond when reward was larger. Manipulations of glutamate and dopamine transmission in VS disrupted changes in the speed of responding to stimuli predictive of the upcoming reward magnitude. This is consistent with correlations between reaction time and firing in VS reported above.

Interestingly, the same group reported that lesions or inactivation of VS had no impact on latency measures, suggesting that complete disruption of VS allows for other areas to motivate behavioral output (Brown and Bowman, 1995; Gierler et al., 2004). This may explain why in some sessions in the current study VS activity was not selective for the upcoming reward, yet there remained a weak difference in response latencies. Notably, rats continued

to choose the more preferred outcome during free-choice trials, consistent with reports that VS is not required for choosing a large over a small reward (Cousins et al., 1996).

Interestingly, our results suggest that VS may play multiple, potentially conflicting roles in delay discounting tasks. On one hand, activity during the decision is higher preceding an immediate reward and seems to invigorate behavior toward the more valued reward. On the other hand, once a decision to respond for the delayed reward had been made, activity in VS neurons increased, as if maintaining a representation of the anticipated reward. Most of the delay discounting literature suggests that the latter function is the one of importance; lesions or other manipulations of VS make animals more likely to abandon a larger, delayed reward in favor of a smaller, more immediate reward (Cousins et al., 1996; Cardinal et al., 2001, 2004; Winstanley et al., 2004; Bezzina et al., 2007; Floresco et al., 2008; Kalenscher and Pennartz, 2008). However, we would speculate that different training procedures might change the relative contributions of these two functions. For example, if animals were highly trained to reverse behaviors based on discounted reward, as in the recording setting used here, they might be less reliant on VS to maintain the value of the discounted reward. In this situation, the primary effect of VS manipulations might be to reduce the ele-

vated motivation elicited by cues predicting more immediate reward.

Another notable aspect of these data is that VS neurons integrated activity regarding value (size and delay) and response, both during forced- and free-choice behavior. Anticipation of differently valued rewards has been shown previously to affect firing in other regions of striatum. For example, many neurons in oculomotor regions of caudate (dorsal medial striatum) encode both direction and motivational value and are thought to be critical in the development of response biases toward desired goals (Lauwereyns et al., 2002). These data differ from our results in several ways. First, neurons in caudate typically exhibit a contralateral bias, firing more strongly for saccades made in the direction opposite to the recording hemisphere. In VS, approximately equal numbers of neurons preferred leftward and rightward movement. These results are consistent with deficits observed after pharmacological manipulations of these areas (Carli et al., 1989). Second, activity in many neurons in caudate has been reported to reflect available movement–reward associations even when the relevant response is not subsequently executed (Lauwereyns et al., 2002; Samejima et al., 2005; Lau and Glimcher, 2008). Such “action-value” or “response-bias” correlates were not present in VS. In this, our results are consistent with recent findings by Ito and Doya (2009), which showed that representations of action value are less dominant in rat VS compared with other types of information. Thus, whereas activity in dorsal striatum (DS) may be critical in representing the value of available actions (behaviorally independent action value), activity in VS seems to be more closely tuned to representing the value of the upcoming response (behaviorally dependent action value). Such activity may reflect an “action-specific reward value” (Samejima et al., 2005), because it is specific for value for only one of the two actions. Practically speaking, such a representation could invigorate or motivate a specific behavior (left or right) through downstream motor areas via some sort of winner-take-all mechanism (Pennartz et al., 1994; Redgrave et al., 1999; Nicola, 2007; Taha et al., 2007).

Another possibility is that the correlates observed in VS incorporate information about the expected outcome itself. Such representations would allow behavior to change spontaneously in response to changes in the value of the outcome. Such information might be acquired through inputs from orbitofrontal cortex or basolateral amygdala, both of which send information to VS and are implicated in signaling of information about expected outcomes (Hatfield et al., 1996; Schoenbaum et al., 1998; Gallagher et al., 1999; Gottfried et al., 2003; Ambroggi et al., 2008). Interestingly, data regarding the role of VS in these behavioral settings is sparse and often contradictory. This is also somewhat true of our own results; because we recorded during presentation of the differently valued outcomes (i.e., during learning), we cannot distinguish signaling such outcome representations from cached estimates of response value.

Critically, such firing cannot represent “cue value” because the signal integrating value and impending response in VS neurons is not present when the rats choose to respond in the opposite direction. Moreover, we have shown previously that responses to the low-value well on these trials are not mistakes; the rats’ response latencies on these trials indicate that they know they are responding for the less valuable outcome (Roesch et al., 2007a). As illustrated in Figure 5B, the elevated cue-evoked activity on trials in which the rats responded in the preferred direction of the neuron (bold lines) was not evident when the rat chose to go in the opposite direction (dashed lines). This was true despite the fact

that, on these trials, the rats sampled the same odor and had available the same outcome in the preferred direction. Notably, this result differs from what we have reported previously for cue-evoked activity in dopamine neurons in this same task; these neurons signaled the value of the best available option on free-choice trials, even when it was not selected (Roesch et al., 2007a). Thus, firing in ventral tegmental area dopamine neurons reflects the value of the better option during decision making, whereas activity in VS neurons tracks the value of the action that is ultimately chosen.

This notion is consistent with the idea that VS plays a critical role in actor–critic models, optimizing long-term action selection through its connections with midbrain dopamine neurons. In this model, the “critic” stores and learns values of states, which in turn are used to compute prediction errors necessary for learning and adaptive behavior. The “actor” stores and forms a policy on which actions should be selected (Joel et al., 2002; Montague et al., 2004). Recently, the functions of critic and actor have been attributed to ventral and dorsal lateral striatum, respectively, based on connectivity, pharmacology, and functional magnetic resonance imaging (fMRI) (Everitt et al., 1991; Cardinal et al., 2002a; O’Doherty et al., 2004; Voorn et al., 2004; Balleine, 2005; Pessiglione et al., 2006).

Our single-unit results fit well with this hypothesis. Neurons in VS signal the value of the upcoming decision, which may in turn impact downstream dopamine neurons that subsequently modify both the actor (DS) and the critic (VS). In this regard, it is noteworthy that analysis of neural activity in VS during learning in this task revealed no evidence that VS neurons encode the actual reward prediction errors, which are proposed to stamp in associative information. This is consistent with recent suggestions that the strong error signal in VS often reported in human fMRI studies reflects input from other areas and is not an output signal from this region (Knutson and Gibbs, 2007).

Finally, we also found that many neurons were inhibited during task performance. Previous studies have also reported long-lasting inhibitions during task performance and argued that these correlates may reflect inhibition of competing behaviors (e.g., locomotion, grooming, and running away) (Nicola et al., 2004; Taha and Fields, 2005; Taha and Fields, 2006). Unlike the instructive signals described above, it is thought that these inhibitory signals should be modulated by appetitive behavior but be independent of the specific response being performed (Taha and Fields, 2006). In our task, it necessary for rats to remain stationary in the odor port and then in the well to receive reward. During these two periods, activity of many VS neurons was inhibited, perhaps reflecting the need to suppress competing behaviors. However, if this is the case, it seems odd that inhibitory activity was no more pronounced on big-reward and short-delay conditions than on small-reward and long-delay trials. This suggests that activity in these cells is not influenced by the value of the reward at stake, despite the fact that the rats attend better and are more motivated on these trials. Of course, maintaining hold in the odor port and then in the fluid well was not difficult, as evidenced by the low number of early non-pokes. It is possible that increasing the requirement to remain still during these periods would provide more evidence for such a function.

References

- Amalric M, Koob GF (1987) Depletion of dopamine in the caudate nucleus but not in nucleus accumbens impairs reaction-time performance in rats. *J Neurosci* 7:2129–2134.
- Ambroggi F, Ishikawa A, Fields HL, Nicola SM (2008) Basolateral amygdala

- neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron* 59:648–661.
- Balleine BW (2005) Neural bases of food-seeking: affect, arousal and reward in corticostriatal limbic circuits. *Physiol Behav* 86:717–730.
- Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 28:309–369.
- Bezzina G, Cheung TH, Asgari K, Hampson CL, Body S, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2007) Effects of quinolinic acid-induced lesions of the nucleus accumbens core on inter-temporal choice: a quantitative analysis. *Psychopharmacology (Berl)* 195:71–84.
- Brog JS, Salyapongse A, Deutch AY, Zahm DS (1993) The patterns of afferent innervation of the core and shell in the “accumbens” part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluoro-gold. *J Comp Neurol* 338:255–278.
- Brown VJ, Bowman EM (1995) Discriminative cues indicating reward magnitude continue to determine reaction time of rats following lesions of the nucleus accumbens. *Eur J Neurosci* 7:2479–2485.
- Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW, Everitt BJ (2001) Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science* 292:2499–2501.
- Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002a) Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev* 26:321–352.
- Cardinal RN, Parkinson JA, Lachenal G, Halkerston KM, Rudarakanchana N, Hall J, Morrison CH, Howes SR, Robbins TW, Everitt BJ (2002b) Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behav Neurosci* 116:553–567.
- Cardinal RN, Winstanley CA, Robbins TW, Everitt BJ (2004) Limbic corticostriatal systems and delayed reinforcement. *Ann N Y Acad Sci* 1021:33–50.
- Carelli RM (2002) Nucleus accumbens cell firing during goal-directed behaviors for cocaine vs “natural” reinforcement. *Physiol Behav* 76:379–387.
- Carelli RM, Deadwyler SA (1994) A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. *J Neurosci* 14:7735–7746.
- Carli M, Jones GH, Robbins TW (1989) Effects of unilateral dorsal and ventral striatal dopamine depletion on visual neglect in the rat: a neural and behavioural analysis. *Neuroscience* 29:309–327.
- 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.
- Cousins MS, Atherton A, Turner L, Salamone JD (1996) Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. *Behav Brain Res* 74:189–197.
- Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol* 89:2823–2838.
- Di Chiara G (2002) Nucleus accumbens shell and core dopamine: differential role in behavior and addiction. *Behav Brain Res* 137:75–114.
- 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.
- Everitt BJ, Morris KA, O’Brien A, Robbins TW (1991) The basolateral amygdala-ventral striatal system and conditioned place preference: further evidence of limbic-striatal interactions underlying reward-related processes. *Neuroscience* 42:1–18.
- Floresco SB, St Onge JR, Ghods-Sharifi S, Winstanley CA (2008) Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cogn Affect Behav Neurosci* 8:375–389.
- Gallagher M, McMahan RW, Schoenbaum G (1999) Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci* 19:6610–6614.
- Gierler C, Bohn I, Hauber W (2003) The rat nucleus accumbens is involved in guiding of instrumental responses by stimuli predicting reward magnitude. *Eur J Neurosci* 18:1993–1996.
- Gierler C, Bohn I, Hauber W (2004) Transient inactivation of the rat nucleus accumbens does not impair guidance of instrumental behaviour by stimuli predicting reward magnitude. *Behav Pharmacol* 15:55–63.
- Gottfried JA, O’Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301:1104–1107.
- Groenewegen HJ, Russchen FT (1984) Organization of the efferent projections of the nucleus accumbens to pallidal, hypothalamic, and mesencephalic structures: a tracing and immunohistochemical study in the cat. *J Comp Neurol* 223:347–367.
- Gruber AJ, O’Donnell P (2009) Bursting activation of prefrontal cortex drives sustained up states in nucleus accumbens spiny neurons in vivo. *Synapse* 63:173–180.
- Gruber AJ, Hussain RJ, O’Donnell P (2009) The nucleus accumbens: a switchboard for goal-directed behaviors. *PLoS ONE* 4:e5062.
- Hassani OK, Cromwell HC, Schultz W (2001) Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J Neurophysiol* 85:2477–2489.
- Hatfield T, Han JS, Conley M, Gallagher M, Holland P (1996) Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci* 16:5256–5265.
- Hauber W, Bohn I, Gierler C (2000) NMDA, but not dopamine D₂, receptors in the rat nucleus accumbens are involved in guidance of instrumental behavior by stimuli predicting reward magnitude. *J Neurosci* 20:6282–6288.
- Heimer L, Zahm DS, Churchill L, Kalivas PW, Wohltmann C (1991) Specificity in the projection patterns of accumbal core and shell in the rat. *Neuroscience* 41:89–125.
- Holland PC, Straub JJ (1979) Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *J Exp Psychol Anim Behav Process* 5:65–78.
- Ikemoto S, Panksepp J (1999) The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res Brain Res Rev* 31:6–41.
- 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.
- Janak PH, Chen MT, Caulder T (2004) Dynamics of neural coding in the accumbens during extinction and reinstatement of rewarded behavior. *Behav Brain Res* 154:125–135.
- Joel D, Niv Y, Ruppin E (2002) Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw* 15:535–547.
- Kalenscher T, Pennartz CM (2008) Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Prog Neurobiol* 84:284–315.
- Knutson B, Gibbs SE (2007) Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology (Berl)* 191:813–822.
- Lau B, Glimcher PW (2008) Value representations in the primate striatum during matching behavior. *Neuron* 58:451–463.
- Lauwereyns J, Watanabe K, Coe B, Hikosaka O (2002) A neural correlate of response bias in monkey caudate nucleus. *Nature* 418:413–417.
- Mogenson GJ, Jones DL, Yim CY (1980) From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol* 14:69–97.
- Montague PR, Hyman SE, Cohen JD (2004) Computational roles for dopamine in behavioural control. *Nature* 431:760–767.
- Nicola SM (2007) The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology (Berl)* 191:521–550.
- 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.
- 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.
- Pennartz CM, Groenewegen HJ, Lopes da Silva FH (1994) The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomical data. *Prog Neurobiol* 42:719–761.
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442:1042–1045.
- Reading PJ, Dunnett SB (1991) The effects of excitotoxic lesions of the

- nucleus accumbens on a matching to position task. *Behav Brain Res* 46:17–29.
- Reading PJ, Dunnett SB, Robbins TW (1991) Dissociable roles of the ventral, medial and lateral striatum on the acquisition and performance of a complex visual stimulus-response habit. *Behav Brain Res* 45:147–161.
- Redgrave P, Prescott TJ, Gurney K (1999) The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* 89:1009–1023.
- Robbins TW, Everitt BJ (1996) Neurobehavioural mechanisms of reward and motivation. *Curr Opin Neurobiol* 6:228–236.
- Robbins TW, Giardini V, Jones GH, Reading P, Sahakian BJ (1990) Effects of dopamine depletion from the caudate-putamen and nucleus accumbens septi on the acquisition and performance of a conditional discrimination task. *Behav Brain Res* 38:243–261.
- Robinson DL, Carelli RM (2008) Distinct subsets of nucleus accumbens neurons encode operant responding for ethanol versus water. *Eur J Neurosci* 28:1887–1894.
- Roesch MR, Taylor AR, Schoenbaum G (2006) Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51:509–520.
- Roesch MR, Calu DJ, Schoenbaum G (2007a) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci* 10:1615–1624.
- Roesch MR, Calu DJ, Burke KA, Schoenbaum G (2007b) Should I stay or should I go? Transformation of time-discounted rewards in orbitofrontal cortex and associated brain circuits. *Ann N Y Acad Sci* 1104:21–34.
- Sage JR, Knowlton BJ (2000) Effects of US devaluation on win-stay and win-shift radial maze performance in rats. *Behav Neurosci* 114:295–306.
- Salamone JD, Correa M (2002) Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav Brain Res* 137:3–25.
- Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. *Science* 310:1337–1340.
- Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* 1:155–159.
- Setlow B, Schoenbaum G, Gallagher M (2003) Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* 38:625–636.
- Taha SA, Fields HL (2005) Encoding of palatability and appetitive behaviors by distinct neuronal populations in the nucleus accumbens. *J Neurosci* 25:1193–1202.
- Taha SA, Fields HL (2006) Inhibitions of nucleus accumbens neurons encode a gating signal for reward-directed behavior. *J Neurosci* 26:217–222.
- Taha SA, Nicola SM, Fields HL (2007) Cue-evoked encoding of movement planning and execution in the rat nucleus accumbens. *J Physiol* 584:801–818.
- van der Meer MA, Redish AD (2009) Covert expectation of reward in rat ventral striatum at decision points. *Front Integr Neurosci* 3:1.
- Voorn P, Vanderschuren LJ, Groenewegen HJ, Robbins TW, Pennartz CM (2004) Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci* 27:468–474.
- Wadenberg ML, Ericson E, Magnusson O, Ahlenius S (1990) Suppression of conditioned avoidance behavior by the local application of (–)sulpiride into the ventral, but not the dorsal, striatum of the rat. *Biol Psychiatry* 28:297–307.
- Wakabayashi KT, Fields HL, Nicola SM (2004) Dissociation of the role of nucleus accumbens dopamine in responding to reward-predictive cues and waiting for reward. *Behav Brain Res* 154:19–30.
- Winstanley CA, Theobald DE, Cardinal RN, Robbins TW (2004) Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J Neurosci* 24:4718–4722.
- Wright CI, Groenewegen HJ (1995) Patterns of convergence and segregation in the medial nucleus accumbens of the rat: relationships of prefrontal cortical, midline thalamic, and basal amygdaloid afferents. *J Comp Neurol* 361:383–403.
- 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.