nature neuroscience

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards

Matthew R Roesch[1,4], Donna J Calu[2,4] & Geoffrey Schoenbaum[1,3]

The dopamine system is thought to be involved in making decisions about reward. Here we recorded from the ventral tegmental area in rats learning to choose between differently delayed and sized rewards. As expected, the activity of many putative dopamine neurons reflected reward prediction errors, changing when the value of the reward increased or decreased unexpectedly. During learning, neural responses to reward in these neurons waned and responses to cues that predicted reward emerged. Notably, this cue-evoked activity varied with size and delay. Moreover, when rats were given a choice between two differently valued outcomes, the activity of the neurons initially reflected the more valuable option, even when it was not subsequently selected.

Dopamine is thought to be essential for reinforcement learning[1–4]. This hypothesis is rooted in single-unit studies, which show that the firing patterns of dopamine neurons accord well with a signal that reports errors in reward prediction[5–15]. This is true of firing for rewards as well as for cues that come to predict rewards. This signal is thought to regulate learning and to guide decision-making; however, few studies have investigated this issue in behavioral tasks that involve active decision-making[11,14].

To address this issue, we recorded single-unit activity in the ventral tegmental area (VTA) of rats performing a variant of a time-discounting task. It is known that animals prefer immediate over delayed reward (time discounting)[16–25], but the role of dopamine in this impulsive behavior remains unclear[26–30]. We designed a task to isolate changes in neural activity related to reward size from changes in neural activity related to time to reward by independently varying when (after a short or long delay) or what (big or small) reward was to be delivered at the end of the trial. As the rats learned to respond to the more valuable odor cue, many dopamine neurons developed cue-selective activity during odor sampling, firing more strongly for cues that predicted either the more immediate reward or the larger one. Notably, these same neurons exhibited activity that was consistent with prediction errors during unexpected delivery or omission of reward. When the rats were given the opportunity to choose between two differently valued rewards, the cue-evoked activity in the dopamine neurons initially reflected the value of the best available option. This was true even when that option was not subsequently selected. Only after cue-offset (after the decision was made) did this activity change to reflect the value of the option that would eventually be chosen.

RESULTS

Neurons were recorded in a delay discounting task[31] (**Fig. 1a**,**b**). On each trial, rats responded to one of two adjacent wells after sampling an odor at a central port (**Fig. 1c**). Rats were trained to respond to three 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 in each recording session, we manipulated either the length of the delay that preceded reward delivery (**Fig. 1a**; blocks 1,2) or the size of the reward (**Fig. 1b**; blocks 3,4). All trials were normalized to be the same duration by adjusting the inter-trial interval. As shown in **Figure 1d–f**, the rats changed their behavior on both free and forced-choice trials across these training blocks, choosing the higher value reward more often on free-choice trials (**Fig. 1e**,**f**) and with greater accuracy and shorter latency on forced-choice trials (**Fig. 1d**). Thus the rats perceived the differently delayed and sized rewards as having different values and could rapidly learn to change their behavior within each trial block.

We recorded 258 neurons from the VTA and substantia nigra pars compacta (SN) in nine rats during these sessions (**Fig. 2a**). Waveforms from these neurons (**Fig. 2b**) were analyzed for features characteristic of dopamine neurons[12–14,32,33] (**Fig. 2c**). This analysis identified 36 neurons that met established electrophysiological criteria for dopamine neurons (**Fig. 2c**), including 2 in the SN (see **Supplementary Data** online for analyses of SN cells) and 34 in the VTA. To confirm the validity of the waveform criteria, we recorded an additional 18 neurons before and after intravenous infusion of the dopamine agonist apomorphine, which inhibits the activity of dopamine neurons[13,34,35]. Neurons whose firing was suppressed in response to apomorphine

[1]Department of Anatomy and Neurobiology, [2]Program in Neuroscience and [3]Department of Psychiatry, University of Maryland School of Medicine, 20 Penn Street, HSF-2 S251, Baltimore, Maryland 21201, USA. [4]These authors contributed equally to this work. Correspondence should be addressed to M.R.R. (mroes001@umaryland.edu).

Received 30 August; accepted 22 October; published online 18 November 2007; doi:10.1038/nn2013

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

clustered with the 36 putative dopamine neurons that were identified by the waveform analysis (**Fig. 2c**).

Prediction error signaling in cue-responsive neurons

The firing activity in 33 of the 34 dopaminergic neurons in the VTA was significantly modulated during odor sampling (compared with baseline; *t*-test; $P < 0.05$). Consistent with previous reports[2,5–13], this population included many neurons ($n = 19$) that increased firing in response to reward. Notably, activity in the 19 dopamine neurons that responded to either cue or reward (cue/reward-responsive dopamine neurons) reflected the positive and negative reward prediction errors that are inherent in our task design, whereas activity in the remaining 14 cue-responsive neurons did not. This was most apparent in the transition between blocks 3 and 4 (**Fig. 1b**); at this transition, the reward in one well is made larger, by delivery of several unexpected boluses of sucrose, and the reward in the other well is made smaller, by omission of several expected sucrose boluses. Many of the cue/reward-responsive dopamine neurons showed suppressed firing when expected boluses were omitted and increased firing when unexpected boluses were delivered. These shifts in firing were maximal immediately after the transition and diminished with learning. These effects are illustrated by the single-unit example shown in **Figure 3a**.

Signaling of prediction errors can also be appreciated in the heat plot in **Figure 4b**, which shows the population response of the 19 cue/reward-responsive dopamine neurons on the first and last 10 forced-choice trials in each direction, for each of the four training blocks shown in **Figure 1a**. Positive prediction errors occurred at the beginning of every trial block in which the reward at a particular well increased in value. This occurred at the transition from 'small' to 'big' (**Fig. 4b**, 4^{bg}, white arrow) and at transitions from 'long' to 'short' (**Fig. 4b**, 2^{sh}, white arrow) and from 'long' to 'big' (**Fig. 4b**, 3^{bg}, white arrow). In each case, neural activity increased significantly during the initial trials of the new block, when reward was delivered unexpectedly, and then declined, as the rats learned to expect the reward. The decline in activity across these trial blocks is illustrated for the individual neurons in **Fig. 3b** ('reward delivery'), which plots the average firing in each neuron during the 500 ms after reward delivery in the first five and the last fifteen trials of these training blocks (2^{sh}, 4^{bg}, 3^{bg} in **Fig. 4b**). These neurons were significantly more likely to fire more in early trials than in late trials (**Fig. 3b**; chi-square, $P < 0.05$), and neurons with a statistically significant decline in activity outnumbered those that showed an increase (**Fig. 3b**; chi-square, $P < 0.0017$). Furthermore, activity averaged across the entire population declined significantly within these trial blocks (**Fig. 3b**; *t*-test, $P < 0.007$).

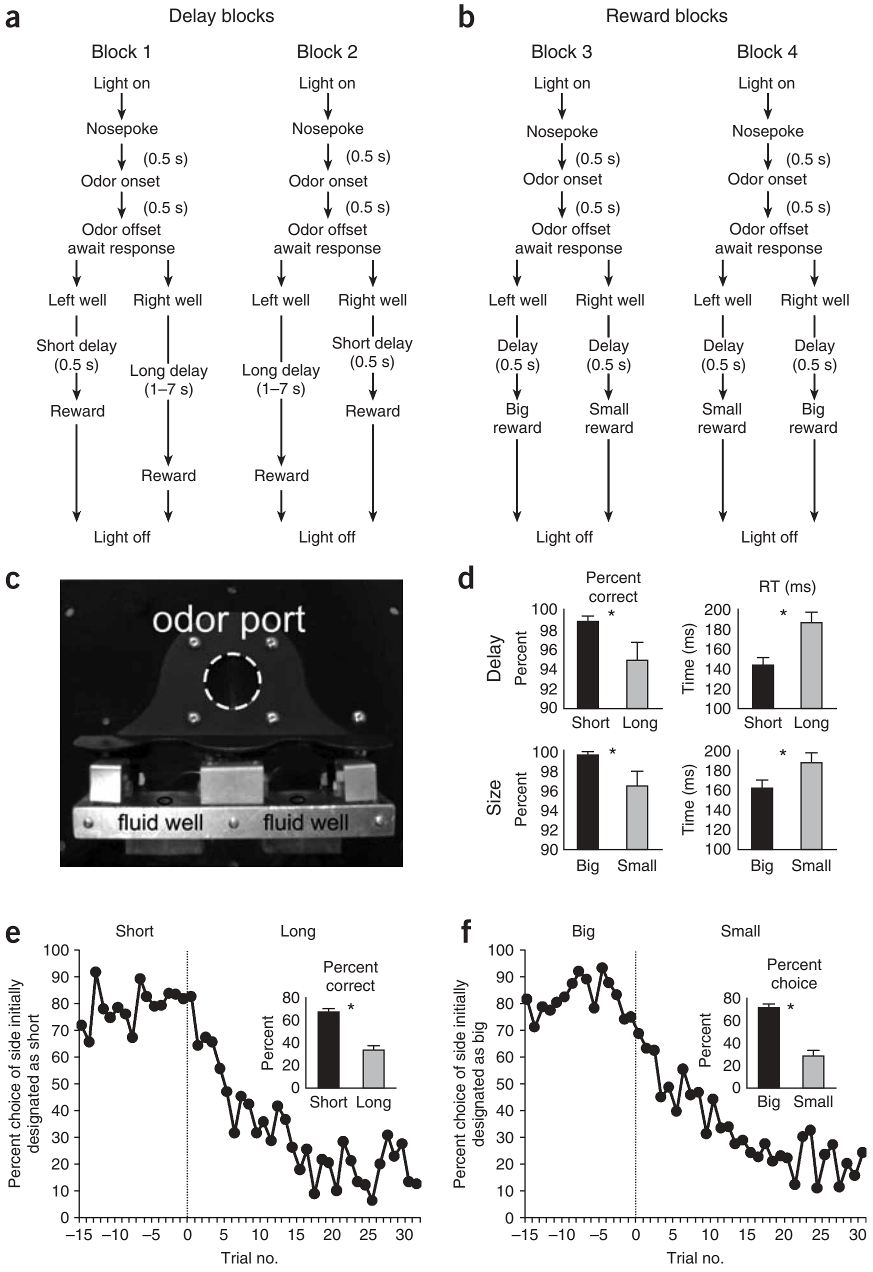


Figure 1 Choice task in which delay and size of reward were manipulated. (**a**,**b**) Sequence of events in delay blocks (**a**) and reward blocks (**b**). At the start of each recording session one well was randomly designated as short and the other as long (block 1). In the second block of trials these contingencies were switched (block 2). In blocks 3 and 4, we held the delay constant while manipulating the size of the reward. At least 60 trials were collected per block. (**c**) Picture of apparatus used in task, showing odor port (~2.5 cm diameter) and the two fluid wells. (**d**) The impact of delay length and reward size on behavior on forced-choice trials. Bar graphs show percent correct (left) and reaction time (RT, right) across all recording sessions. (**e**,**f**) The impact of delay length and reward size on behavior on free-choice trials that were interleaved within the forced-choice trials. Line graphs show choice behavior before and after the switch from short to long (**e**) and from big to small (**f**) reward; inset bar graphs show average percent choice for short versus long (**e**) or big versus small (**f**) across all free-choice trials. Asterisks indicate planned comparisons revealing statistically significant differences (*t*-test, $P < 0.05$). Error bars, s.e.m.

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

Negative prediction errors are also seen in **Figure 4b**, at the beginning of trial blocks in which an expected reward was omitted. This occurred at the transition from 'big' to 'small' (**Fig. 4b**, 4^{sm}, gray arrow) and at the transition from 'short' to 'long' at the time when the short reward would have been delivered (see **Supplementary Data** online). In each case, neural activity declined significantly at the time of reward omission (for example, at the time when the reward would have been delivered on 'short' trials or at the time when the second bolus would have been delivered on 'big' trials), and this decline was greatest in the initial trials and then lessened through the remaining trials in each block, as the rats learned to expect reward omission. This is illustrated for individual neurons in **Figure 3b** ('reward omission'), which plots the average firing for each neuron during the 500-ms period after reward omission in the first five and last fifteen trials of these training blocks (2^{lo}, 4^{sm} in **Fig. 4b**). These neurons were significantly more likely to fire more in late trials than in early trials (**Fig. 3b**; chi-square, $P < 0.001$), and neurons with a statistically significant increase in activity outnumbered those that showed a decrease (**Fig. 3b**; chi-square, $P < 0.02$). Furthermore, activity averaged across the entire population increased significantly within these trial blocks (**Fig. 3b**; *t*-test, $P < 0.009$).

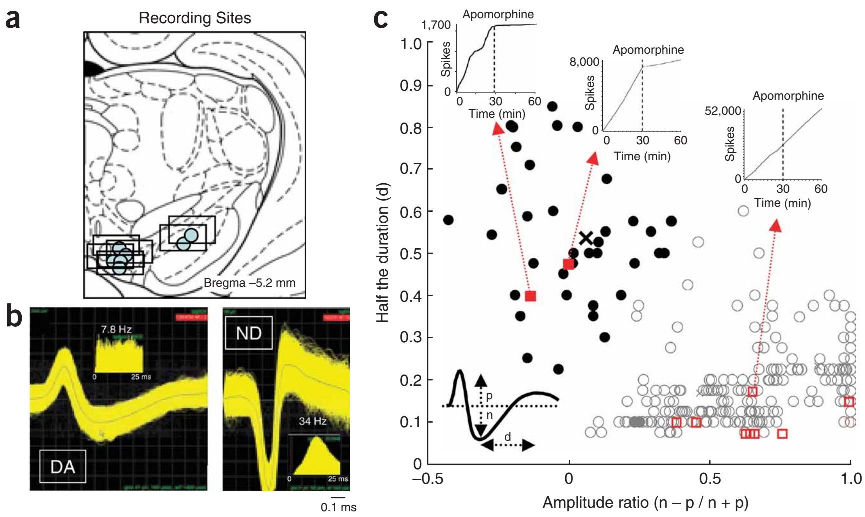


Figure 2 Locations, representative waveforms and classification of putative dopamine neurons. (**a**) Location of the electrode track in each rat; boxes indicate approximate extent of lateral (and anteroposterior) spread of the wires (~1 mm) centered on the final position (dot). (**b**) Example waveforms for putative dopamine (DA) and nondopamine (ND) neurons. (**c**) Results of cluster analysis based on spike duration (d in waveform inset, *y*-axis, ms) and the amplitude ratio (*x*-axis) of the initial negative (n in inset) and positive (p in inset) segments. The center and variance of each cluster was computed without data from the neuron of interest, and then that neuron was assigned to a cluster if it was within 3 s.d. of the center. Neurons that met this criterion for more than one cluster were not classified. This process was repeated for each neuron. Putative dopamine neurons are shown in black; neurons that classified with other clusters, no clusters or more than one cluster are shown as open symbols. Neurons recorded before and after intravenous infusion of apomorphine are shown in red. Inset cumulative sum plots show the effects of apomorphine on baseline firing in two DA neurons and one ND neuron.

By contrast, the 14 cue-responsive dopamine neurons that did not respond to reward showed no evidence of prediction error signaling. This is illustrated in **Figure 3c**, which plots the average firing in response to reward in these neurons early and late in training blocks in which the reward at a particular well increased (**Fig. 3c**, 'reward delivery') or decreased (**Fig. 3c**, 'reward omission') in value unexpectedly. Unlike cue/reward-responsive neurons, these putative dopamine neurons showed no change in firing in response to changes in expected reward. Indeed, signals related to prediction errors were not observed in any of the other populations identified in **Figure 2c** (see **Supplementary Figs. 1–3** online).

Effect of delay and reward size on cue-evoked activity

The reward- and nonreward-responsive dopamine neurons also differed in what information they encoded during cue-sampling. Consistent with their role in signaling prediction errors, the cue/reward-responsive dopamine neurons fired on the basis of the value of the reward that a cue predicted. In each trial block, learning was associated with increased firing in response to the cue that predicted the more valuable reward and decreased firing in response to the cue that predicted the less valuable reward. This is evident in the single-unit example shown in **Figure 4a**, and in the population response shown in **Figure 4b**.

Notably, increased firing in response to the cue that predicted the more valuable reward with learning was seen in both 'size' and 'delay' blocks (**Fig. 4c,d**). The correlation between size and delay is further illustrated in **Figure 5a**, which plots the difference in cue-evoked activity between high and low value trials for each cue/reward-responsive neuron in the first five and last fifteen trials in each block. This difference score was calculated separately for size (big minus small) and delay (short minus long) blocks. During the early trials in each block, these scores were seldom different from zero (**Fig. 5a**), indicating that the value of the reward that the cues predicted caused little or no difference in firing; during the late trials in each trial block, these scores were typically above zero (**Fig. 5a**), indicating greater firing in response to the cue that predicted the more valuable reward. Of the 19 cue/reward-responsive neurons, 9 (47%) showed significant enhancement of firing (*t*-test, $P < 0.05$) under short and big conditions (black dots) as compared with long and small conditions, respectively. Of the remaining neurons, one neuron fired significantly more strongly for short alone (compared to long; *t*-test, $P < 0.05$; blue dots) and three showed significant enhancement under big conditions alone (compared to small; *t*-test, $P < 0.05$; green dots). None exhibited a significant preference for less valuable rewards (long > short or small > big; *t*-test, $P > 0.05$). Furthermore, difference scores from blocks in which value differed owing to delay and reward size were highly correlated (**Fig. 5a**). Notably, the changes in cue-evoked firing were proportional to the delay and were not related to the rate of titration on delay trials (**Supplementary Data**).

By contrast, the 14 cue-responsive neurons that did not respond to reward showed no evidence of value encoding during cue sampling. This is evident in **Figure 5c**, which shows that not a single one of these neurons fired differently depending on the value of the predicted reward.

Notably, cue-selective activity in the cue/reward-responsive neurons reflected the relative rather than absolute value of the rewards that were available in a given block. This is most apparent in a comparison of firing in the 'short' reward (2^{sh}) and 'small' reward (3^{sm}) conditions.

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

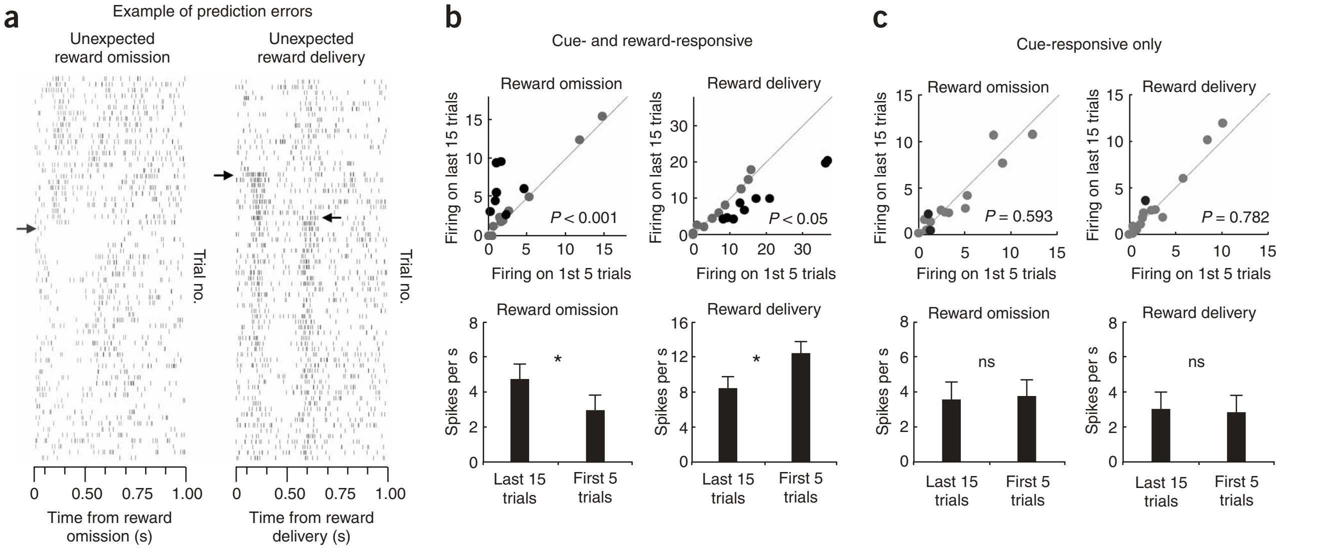


Figure 3 Activity during reward reflects prediction errors in a subpopulation of cue-responsive dopamine neurons. (**a**) Left, example of error signaling (negative prediction error) when an expected reward was omitted during the transition from a 'big' to a 'small' block (gray arrow). Right, example of error signaling (positive prediction error) from the neuron shown in **a** when reward was instituted during the transition from a 'small' to a 'big' block (first black arrow). For this neuron, an additional third bolus was delivered several trials later (second black arrow) to further illustrate prediction error encoding. Activity is aligned to the onset of the first unexpected reward. Raster display includes free- and forced-choice trials. Consistent with encoding of errors, activity changes were transient, diminishing as the rats learned to expect (or not expect) reward at that time during each trial block. These effects are quantified during reward omission (left) and reward delivery (right) for dopamine neurons that were cue- and reward-responsive (**b**; $n = 19$) and dopamine neurons that were cue-responsive only (**c**; $n = 14$) by comparing the average firing rate of each neuron during the 500 ms after an expected reward was omitted (left) or an unexpected reward was instituted (right) in the first five versus the last fifteen trials in the appropriate trial blocks (see text). Black dots represent neurons in which the difference in firing was statistically significant (*t*-test; $P < 0.05$). *P*-values in scatter plots indicate results of chi-square tests comparing the number of neurons above and below the diagonal in each plot. Bar graphs represent average firing rates for each population. Asterisks indicate planned comparisons revealing statistically significant differences (*t*-test, $P < 0.05$). Error bars, s.e.m.

Even though the same odor was presented in both blocks, and it predicted the same amount of reward, these neurons fired more to the cue when it predicted a 'short' reward ($n = 13$; 68%) than when it predicted a 'small' reward ($n = 0$) (**Fig. 5b**; chi-square, $P < 0.001$). That is, although the absolute value of the reward in these two blocks of trials was the same (the 'short' and 'small' rewards were identical), the neurons responded more to this cue when it predicted immediate reward versus delayed reward than when it predicted a small versus a big reward. This effect was not found in the 14 cue-responsive dopamine neurons that did not respond to reward (**Fig. 5d**; chi-square, $P = 0.593$).

Finally, as expected from previous reports, activity in the cue/reward-responsive dopamine neurons did not depend on cue identity or response direction. To illustrate this, we computed for each neuron the value index (high – low/high + low) during odor sampling for each direction. The value selectivity was strongly correlated across direction ($r^2 = 0.598$; $P < 0.001$), indicating that dopamine neurons encoded value for responses made in both directions (**Supplementary Fig. 4** online).

Effect of available options on cue-evoked activity

Few studies have investigated the activity of dopamine neurons in the context of decision-making. Popular computational theories propose that dopamine neurons should encode the value of the best available option (Q-learning) to promote optimal learning, even when less valuable options are subsequently selected[36]. Other 'Actor-Critic' models suggest that dopamine neurons should report the value of the cue averaged over the available options. However, at least one recent study has presented evidence contrary to both these ideas, showing that cue-evoked activity in dopamine neurons directly reflects the value of the option that will subsequently be chosen, a result that supports learning of Q-values by a SARSA (state-action-reward-state-action) type of learning algorithm rather than the more conventional Q-learning or Actor-Critic algorithms[14].

To determine whether signaling of cue value by dopamine neurons in our study also depended on the value of the option selected we compared cue-evoked activity in the cue/reward-responsive neurons on free-choice versus forced-choice trials. Free-choice trials were similar to the forced-choice trials in that the value of the rewards available at the two wells was the same; the only difference was that both rewards were simultaneously available, so the rat had the opportunity to exploit its previous knowledge and select the more valuable reward or explore the less valuable alternative. The availability of both rewards was signaled by presentation of a third 'free-choice' odor cue. This cue was presented throughout training on 7/20 trials, so the rats had the same exposure to this cue as to the other two cues and had the same opportunity to learn about the rewards predicted by this cue over the course of each trial block.

To control for learning, we analyzed data from trials after behavior reflected the contingencies in the current block (>50% choice of more valuable option), so that our analysis would not be contaminated by responses based on the contingencies from the preceding trial block. 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 average population responses on these trials, collapsed across direction, are shown separately for size and delay blocks in **Figure 6**.

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

Consistent with the results already described (**Figs. 4** and **5**), cue-evoked activity was higher on the forced-choice trials when the cue predicted the high-value reward than when it predicted the low-value reward (**Fig. 6a,c**). However, on free-choice trials, this difference did not exist (**Fig. 6b,d**) during cue sampling (gray bar; 500 ms). Instead, cue-evoked activity on free-choice trials was the same as that on the high-value forced-choice trials, regardless of whether the rat ultimately chose the high-value reward or the low-value reward.

Statistical comparisons of activity during cue sampling are shown in **Figure 6**. Activity was significantly stronger on free-choice trials when the low-value reward was selected than on forced-choice trials that ended in presentation of the same low-value reward (*t*-test, *P*'s < 0.007). Furthermore, cue-evoked activity on free-choice, low-value trials was not significantly different from cue-evoked activity on forced-choice trials that ended in presentation of the high-value reward (*t*-test, *P*'s > 0.1). Thus, during cue sampling on free-choice trials, dopamine neurons signaled the value of the best available option, even when it was not ultimately selected. We found similar effects when we compared cue-evoked activity on correct and incorrect forced-choice trials (**Supplementary Fig. 5** online).

Of course, activity in these neurons did ultimately change to reflect the value of the reward that was chosen on low-value free-choice trials. This transition occurred within the first 100 ms after odor offset for both delay and size blocks (*t*-test, $P < 0.05$). This time period immediately precedes movement from the odor port to the fluid well, indicating that the activity of these neurons might change to reflect the unfolding of the decision process on free-choice trials.

The simplest explanation for the unique profile of activity on low-value free-choice trials is that the dopamine neurons initially responded to the value of the best available reward, regardless of whether this outcome was subsequently selected, and then changed to reflect the value of the chosen option, once that decision had been made. However, several other possible explanations for the difference in neural activity between free- and forced-choice trials must be

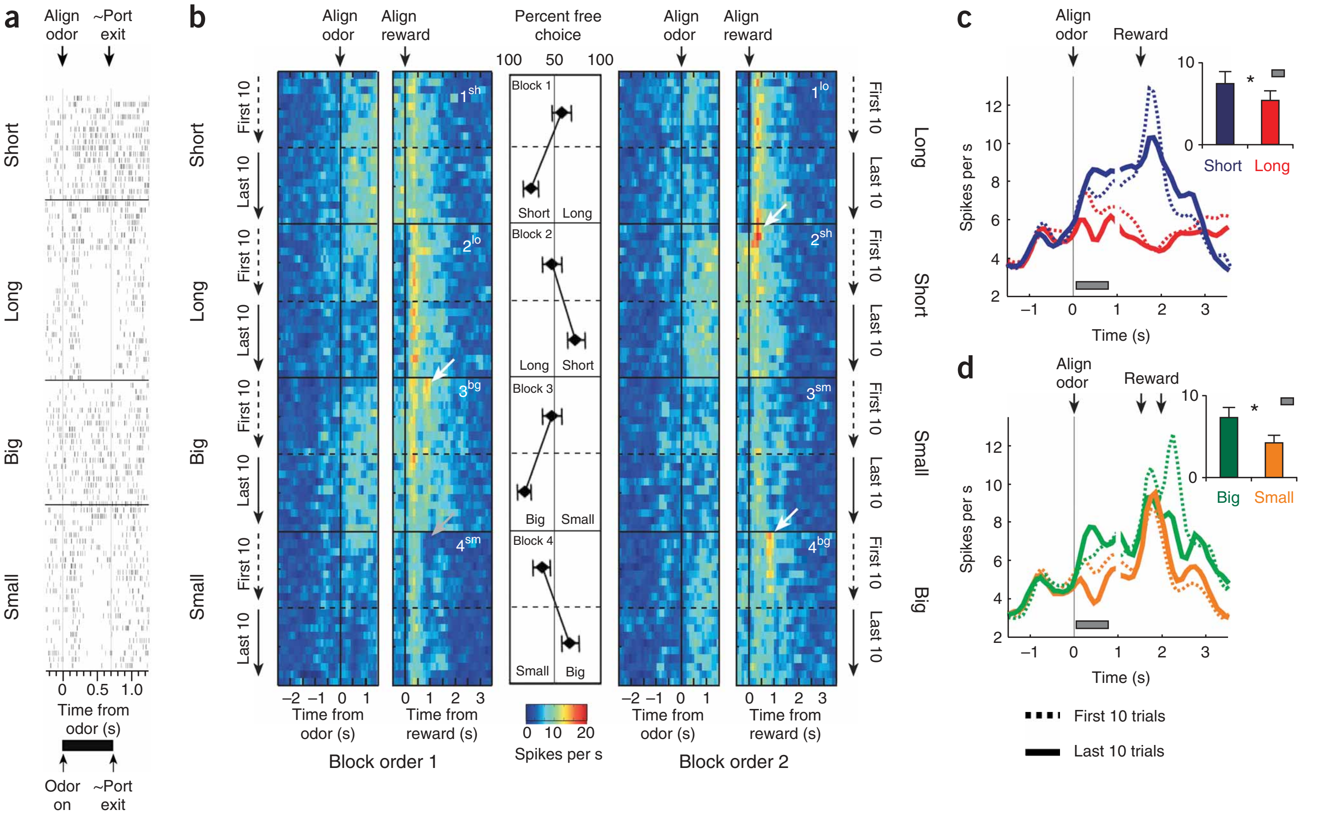


Figure 4 Cue-evoked activity in reward-responsive dopamine neurons reflects the value of the predicted rewards. (**a**) Single-unit example of cue-evoked activity in dopamine neurons on forced-choice trials. Initially the odor predicted that the reward would be delivered immediately ('short'). Subsequently, the same odor predicted a delayed reward ('long'), an immediate but large reward ('big'), and finally an immediate but small reward ('small'). Note, the 'short' and 'small' conditions were identical (1 bolus of reward after 500 ms) but differed in their relative value because 'short' was paired with 'long' in the opposite well whereas 'small' was paired with 'big'. (**b**) Heat plots showing average activity of all cue/reward-responsive dopamine neurons ($n = 19$) during the first and last twenty (10 per direction) forced-choice trials in each training block (**Fig. 1**; blocks 1–4). Activity is shown, aligned on odor onset ('align odor') and reward delivery ('align reward'). Blocks 1–4 are shown in the order performed (top to bottom). During block 1, rats responded after a 'long' delay or a 'short' delay to receive reward (starting direction (left or right) was counterbalanced in each block and is collapsed here). In block 2, the locations of the 'short' delay and 'long' delay were reversed. In blocks 3 and 4, delays were held constant but the size of the reward ('big' or 'small') varied. Line display between heat plots shows the rats' behavior on free-choice trials that were interleaved within the forced-choice trials from which the neural data were taken. Evidence for encoding of positive and negative prediction errors described in **Figure 2** can also be observed here whenever reward is unexpectedly delivered (white arrows) or omitted (gray arrow, analysis epoch). (**c**,**d**) Line graphs summarizing the data shown in **b**. Lines representing average firing rate are broken 1 s after cue onset and 500 ms before reward delivery so that activity can be aligned on both events. Insets: Bar graphs represent average firing rates (gray bar). Blue, short; red, long; green, big; orange, small; dashed, first 10; solid, last 10. Asterisks indicate planned comparisons revealing statistically significant differences (*t*-test, $P < 0.05$). Error bars, s.e.m.

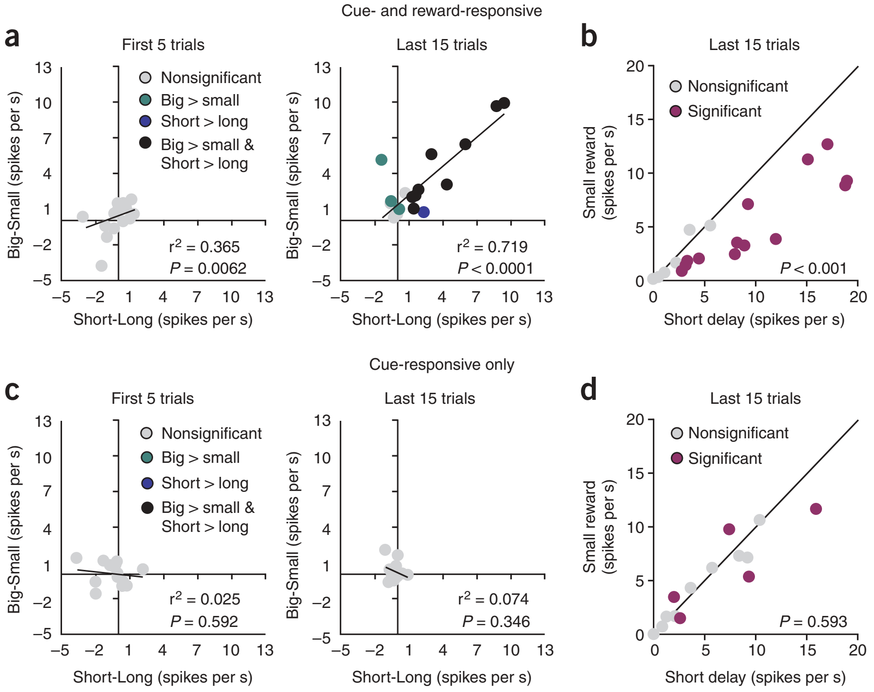
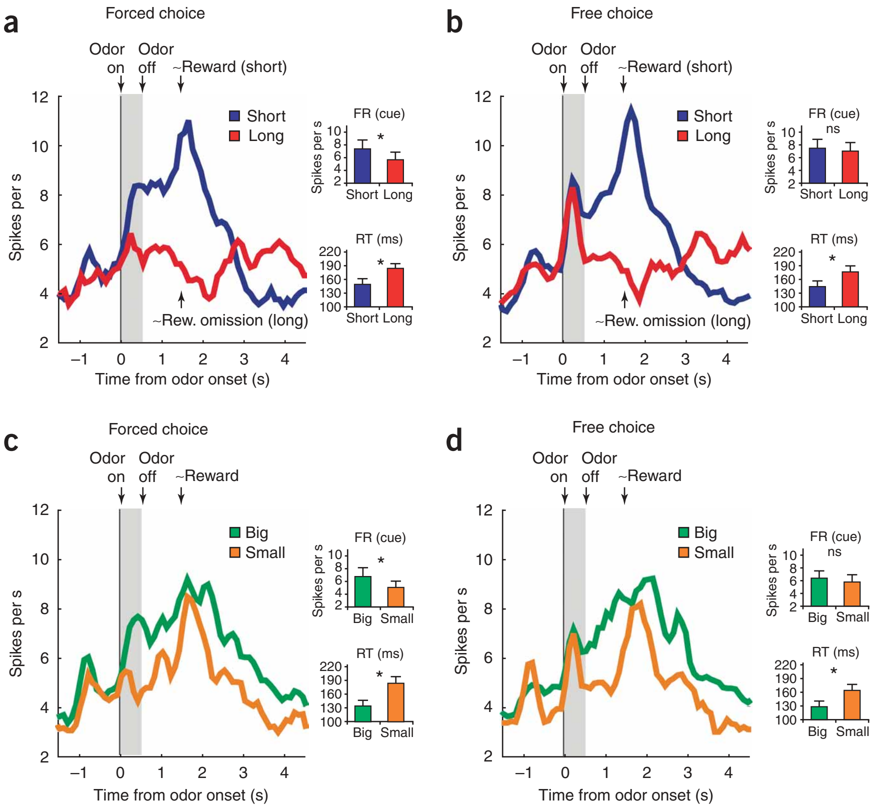


Figure 5 Cue-evoked activity in reward-responsive dopamine neurons covaries with the delay and size of the predicted reward and its relative value. (**a**) Comparison of the difference in firing rate on high- and low- value trials for each cue/reward-responsive neuron ($n = 19$), calculated separately for 'delay' (short-long) and 'reward' blocks (big-small). Colored dots represent those neurons that showed a significant difference in firing between 'high' and 'low' conditions (*t*-test; $P < 0.05$; blue: delay; green: reward; black: both reward and delay). Difference scores were significantly higher in the last fifteen trials of each block (right), indicating that cue selectivity developed with learning. Furthermore, the scores calculated from 'delay' and 'reward' blocks were significantly correlated after learning (right), indicating that encoding of cue value co-varied across the two value manipulations. (**b**) Average firing rate for the same cue/reward-responsive neurons ($n = 19$) under 'short' versus 'small' conditions. Purple dots represent neurons that showed a significant difference in firing between 'short' and 'small' conditions (*t*-test, $P < 0.05$). Neurons were significantly more likely to fire more strongly for a short than for a small reward (chi-square, $P < 0.001$). (**c**,**d**) Same analysis as in **a** and **b** for cue-responsive neurons that did not respond to reward ($n = 14$).

considered. One possible explanation is that rats might have taken longer to decide where to respond on free-choice trials than on forced-choice trials, thereby altering the time course of the dopamine signal on free-choice trials. A second possible explanation is that the free-choice odor elicited a stronger dopaminergic response—even on trials when the rat selected the low-value reward—because it typically led to acquisition of the better outcome, and therefore, was more motivating.

To test whether these factors contributed to our results, we computed the reaction time for each trial from the time of odor offset to the animal's decision to exit the odor port. This measure would be

Figure 6 Cue-evoked activity on free-choice trials reflects the more valuable option. (**a**–**d**) Figures show average activity of all cue/reward-responsive dopamine neurons ($n = 19$) on forced- and free-choice trials, collapsed across direction, for 'delay' (**a**,**b**) and 'reward' blocks (**c**,**d**). 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 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- and free-choice trials in each condition, aligned to odor onset. Bar graphs represent the average firing rate (FR) between odor onset to odor offset (top) and the average reaction time (bottom). Blue, short; red, long; green, big; orange, small. Long-forced versus short-free, *t*-test, $P = 0.002$; long-forced versus long-free, *t*-test, $P = 0.002$; long-forced versus short-forced, *t*-test, $P = 0.001$; short-forced versus short-free, *t*-test, $P = 0.641$; short-forced versus long-free, *t*-test, $P = 0.431$; long-free versus short-free, *t*-test, $P = 0.220$; small-forced versus big-free, *t*-test, $P = 0.004$; small-forced versus small-free, *t*-test, $P = 0.006$; small-forced versus big-forced, *t*-test, $P = 0.002$; big-forced versus big-free, *t*-test, $P = 0.244$; big-forced versus small-free, *t*-test, $P = 0.104$; small-free versus big-free, *t*-test, $P = 0.221$.



© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

influenced by a number of variables relevant to the two possible explanations described above, including the difficulty of the decision and the motivation level of the rat. For example, if free-choice trials were simply more difficult or required more processing time, then reaction times on free-choice trials should be longer. Similarly if the free-choice odor was more motivating, then reaction times on free-choice trials should be shorter.

Contrary to both of these predictions, we found that the rats showed similar reaction times on free- and forced- choice trials of the same value (**Fig. 6**). A two-factor ANOVA (value × trial type) revealed a main effect of value ($F_{1,18} = 14.6$, $P < 0.0003$) but no effect nor any interaction with trial type (F's < 0.15, P's > 0.7). Subsequent contrast testing showed significant differences between differently valued trials, regardless of trial type (P's < 0.017), whereas there were no significant differences in any comparisons between similarly valued trials (P's $>$ 0.675). Thus, the high firing rates on low-value, free-choice trials did not seem to reflect higher motivation or better learning for the odor cue that signaled these trials, nor did the difference in activity on free-versus forced-choice trials reflect any difference in the time course of responding on these two trial types. Notably, the difference in reaction times on high-value and low-value free-choice trials also indicates that the rats knew the likely outcomes of their responses, even when they chose the less valuable reward.

DISCUSSION

We monitored the activity of dopamine neurons in rats performing a time-discounting task in which we independently manipulated the timing and size of rewards. Delay length and reward size were counterbalanced across spatial location and cue identity within each session, and the rats changed their behavior as we manipulated these variables, reliably showing their preference for big over small and immediate over delayed rewards.

As described in primates, dopamine neurons fired more strongly for cues that predicted larger rewards[11,32]. Here we replicated those findings in rats and also showed that the same dopamine neurons encode the relative value of an immediate versus delayed reward. This common representation is interesting because choice of the big over the small reward led to more sucrose over the course of the recording, whereas selection of the immediate over the delayed reward did not. Thus, encoding of prediction errors in VTA dopamine neurons reflects the subjective or relative value of rewards rather than their actual value. This is consistent with studies in other settings in primates[32].

Generally, changes in firing in response to delayed rewards might reflect any of a number of variables. For example, delayed rewards might be discounted because of the cost of lost opportunities or because of the uncertainty of the future reward. Although the delayed reward was always delivered in our task, future rewards are thought to be inherently uncertain. In addition, the timing of the delayed reward was titrated to discourage but not eliminate responding. This was necessary so that we could compare neural activity at the two wells on free-choice trials in each block. However, titrating the delay caused the timing of delayed reward to be less consistent than that of the immediate reward, particularly at the beginning of each delay block, when the time to reward on one side increased rapidly to its new value. Comparison of cue-evoked activity on trials of different delays showed that activity was inversely related to delay length and was not related to the frequency of titration (**Supplementary Fig. 6**). Thus it seems unlikely that the influence of delayed reward on neural activity observed here is an artifact of titration *per se*, although changes in activity might still reflect the more general uncertainty that is inherent in delayed rewards. Whatever the underlying cause, the reduced signaling by dopamine neurons would presumably result in weaker learning for cues that predicted delayed reward over time, leading to apparently impulsive responding for immediate reward.

We also investigated neural activity on trials in which rats could choose to respond for either the high-value or the low-value reward. Reinforcement learning models suggest a number of different ways in which dopamine neurons could represent value in this situation: (i) dopamine neurons could report the value of the option that the animal is going to select (Q value achieved by SARSA learning), (ii) dopamine neurons could represent the average value of all available options (V-learning) or (iii) dopamine neurons could report the value of the best possible option independent of which is ultimately selected (Q-learning). It is unclear which of these predictions is supported by neural data.

Our results are clearly inconsistent with the first alternative, known as Q-value or SARSA, which predicts that cue-evoked activity on free-choice and forced-choice trials should be similar for trials that end in selection of the same reward. Though true for high-value trials, this was not true for the low-value trials. On low-value trials, activity was much higher on free-choice than on forced-choice trials. Indeed, activity on low-value, free-choice trials did not differ statistically from activity on high-value, forced-choice trials.

Our results could be viewed as consistent with the second alternative, known as V-learning, as this model predicts that cue-evoked activity on free-choice trials should be the same, regardless of which reward is ultimately selected. In accord with the predictions of this model, we found that activity was the same on free-choice trials, regardless of subsequent behavior. However this model also indicates that activity should reflect the average 'Pavlovian' values of the cues. In our task, the rats received the high-value reward on more than 99% of the high-value, forced-choice trials. By contrast, they selected the high-value reward on only ~70% of the free-choice trials, opting for the low-value reward on the other ~30% of the trials. Thus the 'Pavlovian' value of the free-choice cue, averaged proportionally across the available rewards, should have been lower than that of the high-value forced-choice cue. Yet, contrary to the predictions of V-learning, cue-evoked activity on these trial types did not differ. Of course, this conclusion depends on the assumption that the activity of dopamine neurons has not saturated at a level corresponding to roughly 70% of the value of the high-value, forced choice cue. Nevertheless although we cannot definitively eliminate this interpretation, it seems to us that our data are not fully consistent with the predictions of V-learning.

Instead, our results seem to be most consistent with the third alternative, known as Q-learning, which predicts that cue-evoked activity on free-choice trials will reflect the value of the best available option. By allowing error signals to reflect the best available option, this model dissociates error signaling from subsequent actions. As a result, learning for antecedent events is not penalized when animals choose to explore less valuable alternatives. Such exploratory behavior would allow animals to recognize when existing knowledge needs to be updated to reflect changing conditions.

Notably, whether our data are interpreted as supporting V-learning or Q-learning, our results are at odds with a recent report in monkeys[14], which did not support these two models. In that study, monkeys were trained to associate different cues displayed on a computer monitor with different probabilities of reward. On some trials, only one cue was presented and the monkey had to select it to obtain reward, as in our forced-choice trials. On other trials, two cues were presented together and the monkey could select either cue to obtain its associated reward, as in our free-choice trials. The authors reported that cue-evoked activity on free-choice trials always reflected the value of the cue that

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

was ultimately selected. These results were interpreted as showing that the activity of dopamine neurons complied with the predictions of SARSA learning.

Several procedural differences between our study and this recent report might account for the divergent findings discussed above. First, it is possible that dopamine signaling is different in rats and in monkeys. However, our results and those of at least one other lab[13] are fully consistent with the proposal that dopamine neurons encode prediction errors in rats as in primates. A second possibility is that the use of a unique odor cue on free-choice trials in the current study biased toward signaling of the better option, as it required the rat to attend to only a single item. The simultaneous presentation of two separate reward-predictive cues in the previous report might have allowed the monkey to attend preferentially to the cue that was to drive subsequent responding. A third possibility is the difference in training. Monkeys are often highly over-trained, completing thousands of trials each day for many months; by comparison, the rats in our study were only lightly trained by recording standards, completing perhaps 5,000 trials over the entire experiment. In addition, the rats had to update their existing knowledge several times within each session to reflect changing reward contingencies. These conditions presumably put a premium on exploratory behavior that might not have been present for the monkeys in the delayed choice task; it might be particularly important to decouple error signaling from subsequent actions in this context.

However, another far more intriguing possibility is that the divergent findings reflect a legitimate difference between the encoding properties of dopamine neurons in SN and those in the VTA. These areas project to very different neural targets[37,38]. Neurons in the SN project preferentially to the dorsal striatum and other areas that have been implicated in habitual and instrumental learning, in which animals learn about responses that predict reward[39–41]. By contrast, neurons in the VTA project preferentially to the ventral striatum and limbic areas that have been implicated in learning what cues predict reward[42–46]. This difference parallels the different firing properties of dopamine neurons in these two circuits — firing on free-choice trials in rat VTA neurons, reported here, signals the potential value of the cue, irrespective of the response selected, whereas firing on free-choice trials in monkey SN neurons signals the value of the response[14]. The unique properties of these circuits for encoding instrumental versus Pavlovian associations might reflect, in part, the different teaching signals received from midbrain dopamine neurons[3].

Of course, activity on free-choice trials did eventually change to reflect the value of chosen reward. This transition occurred after cue sampling, just as the animal exited the odor port and selected a well, thereby tracking the unfolding decision process. Indeed, viewed from another perspective, the dopamine neurons are signaling the best available option both during cue sampling and also thereafter. On high-value, free-choice trials the value of the best option remains unchanged by the decision, whereas on low-value, free-choice trials, the value of the best available option is higher before than after the decision. Thus, after the decision is made, the best option available on free-choice trials is identical to that on forced-choice trials. Neural activity in dopamine neurons in the rat VTA closely tracks this change. Consistent with speculation in ref. 14, this indicates that dopamine neurons are early recipients of information concerning the intention to take a particular action.

METHODS

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

Surgical procedures and histology. Surgical procedures followed guidelines for aseptic technique. Electrodes were manufactured and implanted as in prior recording experiments[31]. All rats had a drivable bundle of 10 25-µm diameter FeNiCr wires (Stablohm 675, California Fine Wire) chronically implanted dorsal to the VTA in the left hemisphere at 5.2 mm posterior to bregma, 0.7 mm laterally, and 7.0 mm ventral to the brain surface. Wires were cut with surgical scissors to extend ~1 mm beyond the cannula and electroplated with platinum (H_2PtCl_6, Aldrich) to an impedance of ~300 kΩ. Cephalexin (15 mg kg^{-1}, post-operative) was administered twice daily for two weeks post-operatively.

For experiments involving apomorphine infusions, sterilized silastic catheters (Dow Corning) were also implanted using published procedures[47]. An incision was made lateral to the midline to expose the jugular vein. The catheter was inserted into the jugular vein and secured using silk sutures. The catheter then passed subcutaneously to the top of the skull where it was connected to the 22-gauge cannula (Plastics One) head mount, which was anchored on the skull using screws and grip cement. Buprenorphine (0.1 mg kg^{-1}, subcutaneous) was administered post-operatively, and the catheters were flushed every 24–48 h with an antibiotic gentamicin/saline solution (0.1 ml at 0.08 mg per 50 ml).

The final electrode position was marked by passing a 15-µA current through each electrode. The rats were then perfused, and their brains removed and processed for histology[31].

Dopamine cell identification. Neurons were screened for wide waveform and amplitude characteristics, and then tested with a non-specific dopamine agonist, apomorphine (0.60–1.0 mg kg^{-1}, intravenous). The apomorphine test consisted of ~30 min of baseline recording, apomorphine infusion, and ~30 min post-infusion recording.

Time-discounting choice task. Recording was conducted in aluminum chambers approximately 18 inch on each side, with sloping walls narrowing to an area of 12 inch by 12 inch at the bottom. A central odor port was located above the two fluid wells. 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. Odors were chosen from compounds obtained from International Flavors and Fragrances.

Trials were signaled by illumination of the panel lights inside the box. When these lights were on, a nosepoke into the odor port resulted in delivery of the odor cue to a small hemicylinder behind this opening. One of three 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 a reward, a second odor instructed the rat to go to the right to get a reward, and a third odor indicated that the rat could obtain a reward at either well. Odors were presented in a pseudorandom sequence so that the free-choice odor was presented on 7/20 trials and the left/right odors were presented in equal numbers. In addition, the same odor could be presented on no more than three consecutive trials.

Once the rats were trained to perform this basic task, we introduced blocks in which we independently manipulated the size of the reward and the delay preceding reward delivery. For recording, one well was randomly designated as short and the other long at the start of the session (**Fig. 1a**, block 1). In the second block of trials these contingencies were switched (**Fig. 1a**, block 2). The length of the delay under long conditions abided by 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 on fewer than 8 out of the previous 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 constant while manipulating the size of reward (**Fig. 1b**). The reward was a 0.05-ml bolus of 10% sucrose solution. The reward magnitude used in delay blocks was the same as the reward used in the reward blocks. For big reward, an additional bolus was delivered after 500 ms. The additional bolus after a delay was used to make

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience

it extremely apparent that the reward got bigger. Occasionally a third bolus was added to further demonstrate positive prediction errors.

Single-unit recording. Wires were screened for activity daily; if no activity was detected, the rat was removed, and the electrode assembly was advanced by 40 or 80 μm. 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 Plexon Multichannel Acquisition Processor systems. Signals from the electrode wires were amplified 20× by an op-amp headstage (Plexon Inc, HST/8o50-G20-GR), located on the electrode array. Immediately outside the training chamber, the signals were passed through a differential pre-amplifier (Plexon Inc, PBX2/16sp-r-G50/16fp-G50), where the single unit signals were amplified 50× and filtered at 150—9,000 Hz. The single unit signals were then sent to the Multichannel Acquisition Processor box, where they were further filtered at 250—8,000 Hz, digitized at 40 kHz and amplified at 1–32×. Waveforms (>2.5:1 signal-to-noise) were extracted from active channels and recorded to disk by an associated workstation

Data analysis. Units were sorted using Offline Sorter software from Plexon Inc. Sorted files were then processed and analyzed in Neuroexplorer and Matlab. To examine activity related to reward delivery or omission, we studied activity 500 ms after reward delivery or omission. We chose 500 ms because no other trial event occurred until at least 500 ms after reward delivery (or omission). Initial analysis of cue-related activity was confined to activity starting after odor onset and ending at the odor port exit. Later analysis of cue related activity examined activity during the 500 ms while the odor was on. Analysis of free-choice versus forced-choice trials included trials on which the more valuable option was chosen more than 50% of the time. Furthermore, we paired each free-choice trial with the immediately preceding and following forced-choice trial of the same value. This procedure allowed us to control for the fact that low-value choices might be more frequent early in a block of trials. We used *t*-tests to measure difference between trial types ($P < 0.05$). Additionally, Pearson Chi-square tests ($P < 0.05$) were used to compare the proportions of neurons.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank Y. Niv, P. Shepard, G. Morris and W. Schultz for thoughtful comments on this manuscript, and S. Warrenburg at International Flavors and Fragrances for his assistance in obtaining odor compounds. This work was supported by grants from the US National Institute on Drug Abuse (R01-DA015718, G.S.; K01-DA021609, M.R.R.), the National Institute of Mental Health (F31-MH080514, D.J.C.), the National Institute on Aging (R01-AG027097, G.S.) and the National Institute of Neurological Disorders and Stroke (T32-NS07375, M.R.R.).

AUTHOR CONTRIBUTIONS

M.R.R., D.J.C. and G.S. conceived the experiments. M.R.R. and D.J.C. carried out the recording work and assisted with electrode construction, surgeries and histology. The data were analyzed by M.R.R. and G.S., who also wrote the manuscript with assistance from D.J.C.

Published online at http://www.nature.com/natureneuroscience
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Wise, R.A. Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**, 483–494 (2004).
2. Schultz, W. Getting formal with dopamine and reward. *Neuron* **36**, 241–263 (2002).
3. Dayan, P. & Balleine, B.W. Reward, motivation and reinforcement learning. *Neuron* **36**, 285–298 (2002).
4. Day, J.J., Roitman, M.F., Wightman, R.M. & Carelli, R.M. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* **10**, 1020–1028 (2007).
5. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* **72**, 1024–1027 (1994).
6. Fiorillo, C.D., Tobler, P.N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
7. Tobler, P.N., Dickinson, A. & Schultz, W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* **23**, 10402–10410 (2003).
8. Hollerman, J.R. & Schultz, W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* **1**, 304–309 (1998).
9. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
10. Montague, P.R., Dayan, P. & Sejnowski, T.J. A framework for mesencephalic dopamine systems based on predictive hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
11. Bayer, H.M. & Glimcher, P.W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
12. Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y. & Hikosaka, O. Dopamine neurons can represent context-dependent prediction error. *Neuron* **41**, 269–280 (2004).
13. Pan, W.X., Schmidt, R., Wickens, J.R. & Hyland, B.I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* **25**, 6235–6242 (2005).
14. Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* **9**, 1057–1063 (2006).
15. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Reward-predicting activity of dopamine and caudate neurons—a possible mechanism of motivational control of saccadic eye movement. *J. Neurophysiol.* **91**, 1013–1024 (2004).
16. Cardinal, R.N., Pennicott, D.R., Sugathapala, C.L., Robbins, T.W. & Everitt, B.J. Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science* **292**, 2499–2501 (2001).
17. Evenden, J.L. & Ryan, C.N. The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology (Berl.)* **128**, 161–170 (1996).
18. Herrnstein, R.J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* **4**, 267–272 (1961).
19. Ho, M.Y., Mobini, S., Chiang, T.J., Bradshaw, C.M. & Szabadi, E. Theory and method in the quantitative analysis of "impulsive choice" behaviour: implications for psychopharmacology. *Psychopharmacology (Berl.)* **146**, 362–372 (1999).
20. Mobini, S. *et al.* Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology (Berl.)* **160**, 290–298 (2002).
21. Kahneman, D. & Tverskey, A. Choices, values and frames. *Am. Psychol.* **39**, 341–350 (1984).
22. Kalenscher, T. *et al.* Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. *Curr. Biol.* **15**, 594–602 (2005).
23. Lowenstein, G.E.J. *Choice Over Time* (Russel Sage Foundation, New York, 1992).
24. Thaler, R. Some empirical evidence on dynamic inconsistency. *Econ. Lett.* **8**, 201–207 (1981).
25. Winstanley, C.A., Theobald, D.E., Cardinal, R.N. & Robbins, T.W. Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J. Neurosci.* **24**, 4718–4722 (2004).
26. Cardinal, R.N., Winstanley, C.A., Robbins, T.W. & Everitt, B.J. Limbic corticostriatal systems and delayed reinforcement. *Ann. NY Acad. Sci.* **1021**, 33–50 (2004).
27. Kheramin, S. *et al.* Effects of orbital prefrontal cortex dopamine depletion on intertemporal choice: a quantitative analysis. *Psychopharmacology (Berl.)* **175**, 206–214 (2004).
28. Wade, T.R., de Wit, H. & Richards, J.B. Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology (Berl.)* **150**, 90–101 (2000).
29. Cardinal, R.N., Robbins, T.W. & Everitt, B.J. The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology (Berl.)* **152**, 362–375 (2000).
30. Roesch, M.R., Takahashi, Y., Gugsa, N., Bissonette, G.B. & Schoenbaum, G. Previous cocaine exposure makes rats hypersensitive to both delay and reward magnitude. *J. Neurosci.* **27**, 245–250 (2007).
31. Roesch, M.R., Taylor, A.R. & Schoenbaum, G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* **51**, 509–520 (2006).
32. Tobler, P.N., Fiorillo, C.D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
33. Kiyatkin, E.A. & Rebec, G.V. Heterogeneity of ventral tegmental area neurons: single-unit recording and iontophoresis in awake, unrestrained rats. *Neuroscience* **85**, 1285–1309 (1998).
34. Bunney, B.S., Aghajanian, G.K. & Roth, R.H. Comparison of effects of L-dopa, amphetamine and apomorphine on firing rate of rat dopaminergic neurones. *Nat. New Biol.* **245**, 123–125 (1973).
35. Skirboll, L.R., Grace, A.A. & Bunney, B.S. Dopamine auto- and postsynaptic receptors: electrophysiological evidence for differential sensitivity to dopamine agonists. *Science* **206**, 80–82 (1979).
36. Niv, Y., Daw, N.D. & Dayan, P. Choice values. *Nat. Neurosci.* **9**, 987–988 (2006).
37. Haber, S.N., Fudge, J.L. & McFarland, N.R. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J. Neurosci.* **20**, 2369–2382 (2000).
38. Joel, D. & Weiner, I. The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* **96**, 451–474 (2000).
39. Yin, H.H., Knowlton, B.J. & Balleine, B.W. Lesions of dorsolateral striatum preserve outcome expectancy, but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* **19**, 181–189 (2004).
40. O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).

41. Knowlton, B.J., Mangels, J.A. & Squire, L. A neostriatal habit learning system in humans. *Science* **273**, 1399–1402 (1996).
42. Hatfield, T., Han, J.S., Conley, M., Gallagher, M. & Holland, P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J. Neurosci.* **16**, 5256–5265 (1996).
43. Gallagher, M., McMahan, R.W. & Schoenbaum, G. Orbitofrontal cortex and representation of incentive value in associative learning. *J. Neurosci.* **19**, 6610–6614 (1999).
44. Baxter, M.G., Parker, A., Lindner, C.C.C., Izquierdo, A.D. & Murray, E.A. Control of response selection by reinforcer value requires interaction of amygdala and orbitofrontal cortex. *J. Neurosci.* **20**, 4311–4319 (2000).
45. Gottfried, J.A., O'Doherty, J. & Dolan, R.J. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301**, 1104–1107 (2003).
46. Parkinson, J.A., Cardinal, R.N. & Everitt, B.J. Limbic cortical-ventral striatal systems underlying appetitive conditioning. *Prog. Brain Res.* **126**, 263–285 (2000).
47. Lu, L. *et al.* Central amygdala ERK signaling pathway is critical to incubation of cocaine craving. *Nat. Neurosci.* **8**, 212–219 (2005).

© 2007 Nature Publishing Group http://www.nature.com/natureneuroscience