



Published in final edited form as:

Neuron. 2011 January 13; 69(1): 170–182. doi:10.1016/j.neuron.2010.11.041.

Heterogeneous Coding of Temporally Discounted Values in the Dorsal and Ventral Striatum during Intertemporal Choice

Xinying Cai¹, Soyoun Kim¹, and Daeyeol Lee²

Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510

SUMMARY

In choosing between different rewards expected after unequal delays, humans and animals often prefer the smaller but more immediate reward, indicating that the subjective value or utility of reward is depreciated according to its delay. Here, we show that the neurons in the primate caudate nucleus and ventral striatum modulate their activity according to temporally discounted values of rewards with a similar time course. However, neurons in the caudate nucleus encoded the difference in the temporally discounted values of the two alternative targets more reliably than the neurons in the ventral striatum. In contrast, the neurons in the ventral striatum largely encoded the sum of the temporally discounted values, and therefore, the overall goodness of available options. These results suggest a more pivotal role for the dorsal striatum in action selection during intertemporal choice.

INTRODUCTION

The outcomes expected from various actions vary in multiple dimensions and can often create a conflict. Accordingly, the ability to combine appropriately the information about multiple attributes of action outcomes is critical for choosing the actions most beneficial to the animal. For example, during intertemporal choice between a small but more immediate reward and a large but more delayed reward, people and animals often choose the smaller reward if the difference in magnitude is too small or if the difference in delay is sufficiently large. This indicates that the subjective value of a delayed reward is reduced compared to when the same reward is immediately available. Formally, how steeply the reward value decreases with its delay is given by a temporal discount function. A temporally discounted value for a delayed reward is then given by the magnitude of reward multiplied by its discount function. Humans and many other species of animals tend to choose the reward with the maximum temporally discounted value (Frederick et al., 2002; Green and Myerson, 2004; Kalenscher and Pennartz, 2008; Hwang et al., 2009).

Disruption in this ability to combine appropriately the information about the magnitude and delay of reward characterizes the maladaptive choice behaviors observed in many psychiatric disorders (Madden et al., 1997; Vuchinich and Simpson, 1998; Mitchell, 1999; Kirby and Petry, 2004; Reynolds, 2006). Nevertheless, how temporally discounted values are computed in the brain and used for decision making is not well understood. In particular,

²Correspondence: daeyeol.lee@yale.edu.

¹These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

previous neuroimaging and lesion studies have highlighted the role of the basal ganglia in decision making involving temporal delays (Cardinal et al., 2001; McClure et al., 2004, 2007; Tanaka et al., 2004; Hariri et al., 2006; Kable and Glimcher, 2007; Wittmann et al., 2007; Weber and Huettel, 2008; Gregorios-Pippas et al., 2009; Pine et al., 2009; Luhmann et al., 2008; Ballard and Knutson, 2009; Bickel et al., 2009; Xu et al., 2009), but precisely how its different subdivisions contribute to intertemporal choice is not clear. Although previous neurophysiological studies in primates (Apicella et al., 1991; Schultz et al., 1992; Williams et al., 1993; Bowman et al., 1996; Hassani et al., 2001; Cromwell and Schultz, 2003) have found that the signals related to the direction of the animal's movement and expected reward tend to be more strongly represented in the dorsal and ventral striatum, respectively, how the activity in different subdivisions of the striatum is coordinated during intertemporal choice has not been investigated. In this study, we found that neurons in both the caudate nucleus and ventral striatum encoded temporally discounted values. However, neurons in the ventral striatum tended to represent the sum of the temporally discounted values for the two targets, whereas those in the caudate nucleus additionally encoded the signals necessary for selecting the action with the maximum temporally discounted value, namely, the relative difference in the temporally discounted values of the two alternative rewards. Therefore, the primate dorsal striatum might play a more important role in decision making for delayed rewards.

RESULTS

Intertemporal choice behavior in monkeys

Two monkeys (H and J) were trained to perform an intertemporal choice task, in which they chose between two different amounts of juice that is either available immediately or delayed (Kim et al., 2008; Hwang et al., 2009). The magnitude and delay of each reward was indicated by the color of the target and the number of small yellow dots around it (Figure 1A, top; see Experimental Procedures). Both animals chose the small reward more often as the delay for the small and large reward decreased and increased, respectively, indicating that they integrated both reward magnitude and delay to determine its choice. The choice behavior during this task was modeled using exponential and hyperbolic discount functions. We found that among 61 and 116 sessions tested for monkeys H and J, respectively, the hyperbolic discount function provided the better fit in 55.7% and 98.3% of the sessions (Figure 1B). The median value of k parameter was 0.18 and 0.25 s^{-1} for monkey H and J, corresponding to the half-life ($1/k$) of 5.6 and 4.0 s, respectively.

Striatal activity related to temporally discounted values

Single-neuron activity was recorded from 93 neurons in the caudate nucleus (CD; 32 from monkey H, 61 from monkey J) and 90 neurons in the ventral striatum (VS; 33 from monkey H, 57 from monkey J) during the intertemporal choice task (Figure 1C). In addition, each of these neurons was also tested during the control task, in which the animal was required to shift its gaze according to the color of the central fixation target (Figure 1A, bottom). Although the visual stimuli were similar for the two tasks, the reward delay and magnitude was fixed for all targets during the control task, which made it possible to distinguish between the activity changes related to the temporally discounted values and those related to visual features of the computer display (see below).

To analyze the neural activity during the intertemporal choice task, we estimated the temporally discounted values for both targets in each trial using the discount function estimated from the animal's behavior (see Experimental Procedures). We then examined the activity of each neuron during the 1-s cue period by applying a series of regression models that include the temporally discounted values of the two targets or various linear

combinations of them in addition to the position of the animal's chosen target. Among these variables, the difference between the temporally discounted values of the two targets was of particular interest, since this corresponds to the decision variable used to fit the animal's choice in the behavioral model. Therefore, we first applied a model including the sum of the discounted values for the leftward and rightward targets, their difference, and the difference in the discounted values for the chosen and unchosen targets (model 1). This analysis showed that many neurons in the CD significantly changed their activity according to the difference in the temporally discounted values for the leftward and right targets (Figure 2; Table 1). Overall, the neurons in the CD were more likely to encode the difference in the discounted values (24 neurons, 25.8%) than those in the VS (10 neurons, 11.1%; χ^2 -test, $p < 0.05$). Similarly, the percentage of neurons encoding the position of the target chosen by the animal was significantly higher in the CD (24 neurons, 25.8%) than in the VS (5 neurons, 5.6%; χ^2 -test, $p < 0.0005$). The fraction of neurons encoding the animal's choice was not significantly above the chance level in the ventral striatum (binomial test, $p = 0.47$).

In addition to the difference in the temporally discounted values for the leftward and rightward targets, some neurons in both CD and VS encoded their sum and the difference in temporally discounted values for the chosen and unchosen targets. For example, the CD neuron illustrated in Figure 2 significantly decreased its activity with the sum of the temporally discounted values (Figure 2C), whereas one of the two VS neurons illustrated in Figure 3 significantly increased its activity with the same variable (Figure 3B). The other VS neuron in Figure 3 decreased its activity significantly as the temporally discounted value of the chosen target increased relative to that of the unchosen target (Figure 3F). Neurons in the VS were more likely to encode the sum in the temporally discounted value of the two targets than their difference (χ^2 -test, $p < 10^{-3}$), whereas the proportion of the neurons in the CD significantly modulating their activity according to these two variables was not significantly different ($p = 0.57$). In addition, the percentage of neurons encoding the sum of the discounted values for the two targets was higher in the VS (31 neurons, 34.4%) than in the CD (20 neurons, 21.5%), although this difference was only marginally significant (χ^2 -test, $p = 0.051$). More neurons in the VS (12 neurons, 13.3%) encoded the difference in the temporally discounted values for the chosen and unchosen targets than in the CD (7 neurons, 7.5%), but this difference was not statistically significant (χ^2 -test, $p = 0.20$). In addition, there was no significant bias for the neurons in either CD or VS to increase or decrease their activity as the temporally discounted value of the target in the contralateral visual field increased relative to that of the target in the ipsilateral visual field (binomial test, $p > 0.1$). We also found that the number of neurons significantly modulating their activity according to various types of temporally discounted values was largely unaffected when the reaction time and peak velocity of the saccade were included in the regression model (Table S1).

These results suggest that the signals related to the temporally discounted values for the two targets are combined differently in the caudate nucleus and ventral striatum. In the caudate nucleus, neurons often encoded the difference between the temporally discounted values of the two alternative targets, suggesting that the activity might increase with the value of one target and decrease with the value of the other target. In contrast, neurons in the ventral striatum largely encoded the sum of temporally discounted value of the two targets, suggesting that their activity might be influenced similarly by the temporally discounted values of both targets. To test these predictions more directly, we applied a regression model that includes the temporally discounted values of the leftward and rightward targets (model 2; see Experimental Procedures). For the CD neuron illustrated in Figure 2, this analysis found that the regression coefficient for the temporally discounted value of the left target was significantly negative (t-test, $p < 10^{-15}$), whereas the regression coefficient for the right target was significantly positive ($p < 0.05$). We found that the number of neurons showing the significant effects of temporally discounted values for both targets was 9 for both CD and

VS (Figure 4A). In both areas, this was significantly more than expected when the activity of each neuron was influenced by the temporally discounted values of the two targets independently (χ^2 -test, $p < 0.05$). Furthermore, among the neurons that significantly modulated their activity according to both variables, 6 neurons in the CD but only 1 neuron in the VS showed opposite signs in the corresponding regression coefficients. This difference was statistically significant (χ^2 -test, $p < 0.05$), confirming the results described above that the neurons in the CD tended to encode the difference in the temporally discounted values of the two alternative targets more frequently than the VS neurons. We also found that the regression coefficients associated with the temporally discounted values of the left and right targets were significantly more positive than the values obtained from the permutation test (see Experimental Procedures) in the VS ($p < 10^{-4}$), but not in the CD ($p = 0.58$; Figure 4A).

To test whether neurons in the striatum combine the signals related to the temporally discounted values for the chosen and unchosen targets, we also applied a regression model that includes these two values separately (model 3). We found that neurons in the CD were more likely to encode the temporally discounted value for the chosen target ($n = 22$ neurons) than for the unchosen target ($n = 9$ neurons; χ^2 -test, $p < 0.01$; Figure 4B). In the VS, 26 and 21 neurons significantly modulated their activity according to the temporally discounted value of the chosen and unchosen targets, respectively, and this difference was not significant (χ^2 -test, $p > 0.4$). We also found that 6 and 9 neurons in the CD and VS, respectively, significantly modulated their activity according to the temporally discounted values for both chosen and unchosen targets (Figure 4B). For the CD, this was significantly more than expected when the temporally discounted values of chosen and unchosen targets influenced the activity of each neuron independently (χ^2 -test, $p < 0.005$). In addition, most neurons encoding the temporally discounted values for both chosen and unchosen targets showed the same signs for their regression coefficients (4 and 7 neurons in the CD and VS, respectively). For both CD and VS, the correlation coefficient between the regression coefficients for the temporally discounted values of the chosen and unchosen targets was significantly more positive than the values obtained from the permutation test ($p < 10^{-4}$; Figure 4B).

Striatal activity related to discounted values vs. visual features

To test whether activity seemingly related to temporally discounted values might reflect the effects of different target colors or number of yellow dots used to indicate the reward magnitude and delay, we analyzed the activity recorded during the control task. During the control task, the delay and magnitude of reward were fixed for all targets. Therefore, the activity of neurons encoding temporally discounted values should be unrelated to the “fictitious” temporally discounted values that are computed as if the magnitude and delay of reward during the control task varied with the target color and number of yellow dots. Indeed, many of the neurons in the CD and VS that changed their activity according to the difference in the temporally discounted values for the leftward and rightward targets (Figures 2B and 2C), their sum (Figure 3B), or the difference in the values for the chosen and unchosen targets (Figure 3F) did not change their activity according to the fictitious temporally discounted values in the control task.

The number of CD neurons encoding the difference in the fictitious temporally discounted values for the leftward and rightward targets in the control task ($n = 8$, 8.6%) was significantly smaller than that in the intertemporal choice task ($n = 24$, 25.8%; χ^2 -test, $p < 0.005$; Table S2). In addition, the number of VS neurons encoding the sum of the fictitious temporally discounted values ($n = 15$, 16.7%) was significantly lower than that in the intertemporal choice task ($n = 31$, 34.4%, χ^2 -test, $p < 0.01$). In addition, we found that for both CD and VS, the average magnitude of the standardized regression coefficients related

to the sum and difference in the temporally discounted values was significantly larger for the intertemporal choice than for the control task (Figure 5; Table S2). Moreover, for the majority of the neurons that showed significant interactions between the temporally discounted values and the task (model 4), the standardized regression coefficients associated with the temporally discounted values were smaller for the control task than for the intertemporal choice task, when they were estimated by applying the original regression model separately to these two separate groups of trials (Figure 5; Table S2). Therefore, value-related activity in the striatum during the intertemporal choice did not simply reflect the visual features used to indicate the reward parameters. In contrast to the activity changes related to temporally discounted values, neural activity in the CD related to the animal's choice was largely comparable for the intertemporal choice and control tasks. For example, the number of CD neurons that modulated their activity according to the animal's choice was 24 and 25 during the intertemporal choice and control tasks, respectively (Figure 2B). The number of VS neurons encoding the animal's choice increased significantly during the control task (18 neurons, 20%) compared to the result obtained for the intertemporal choice task (5 neurons, 5.6%; χ^2 -test, $p < 0.01$).

Striatal activity related to multiple reward parameters

By definition, the temporally discounted value of the reward from a given target increases with its magnitude and decreases with its delay. Therefore, the activity of any neuron that is correlated with either the magnitude or delay of a reward, but not necessarily both, would be also correlated with its temporally discounted value. To test whether the activity of striatal neurons seemingly related to the temporally discounted values was modulated by both of these reward parameters, we applied a regression model that includes the position of the large-reward target, the magnitude of the reward chosen by the animal, the reward delays for the two alternative targets, and the delay of the chosen reward (model 5; see Experimental Procedures). We found that many neurons in the CD and VS indeed significantly changed their activity according to reward magnitudes and delays. For example, a neuron in the CD illustrated in Figure 2B increased its activity with the reward delay for the leftward target (t -test, $p < 10^{-8}$). It also decreased its activity with the reward delay for the rightward target, although this was not statistically significant ($p = 0.2$). The activity of the same neuron increased significantly when the reward for the rightward target was large ($p < 10^{-10}$), suggesting that the activity of this neuron related to the temporally discounted values did not merely result from the signals related to either the magnitude or delay of reward alone. Similarly, the VS neuron illustrated in Figure 3F increased its activity as the reward delay increased for the target chosen by the animal ($p < 0.05$) and decreased its activity when the animal chose the large reward ($p < 0.01$).

The same regression analysis showed that the position of the large-reward target significantly changed the activity of 29 (31.2%) and 16 (17.8%) neurons in the CD and VS, respectively. In addition, the magnitude of the reward chosen by the animal significantly influenced the activity of 16 (17.2%) and 14 (15.6%) neurons in the CD and VS, respectively. The effect of reward delay was significant in 11 (11.8%) and 16 (17.8%) neurons in the CD and VS, respectively. In addition, the neurons significantly changing their activity according to reward delays were more likely to encode the position of the large-reward target (χ^2 -test, $p < 0.005$). Overall, 9 of 11 CD neurons (81.8%) showing the significant effect of delay also encoded the position of the large-reward target, whereas this was true for 7 of 16 neurons (43.8%) in the VS. Similarly, the effect of reward delay for the chosen target was significant in 10 neurons in the CD (10.8%) and 11 neurons in the VS (12.2%). The neurons significantly changing their activity according to the delay of chosen reward were also more likely to encode the magnitude of the chosen reward (χ^2 -test,

$p < 0.005$). Overall, 5 of 10 (50%) neurons in the CD and 8 of 11 (72.7%) neurons in the VS with the significant effect of chosen delay also encoded the magnitude of chosen reward.

For neurons encoding the temporally discounted value of the reward from a particular target, their activity should be modulated oppositely by the magnitude and delay of the reward. To test whether striatal neurons combine the information about the magnitude and delay of reward in their activity appropriately to encode its temporally discounted values, we examined neuron-target pairs that showed significant effects of both reward magnitude and delay. For the majority of such cases in both the CD and VS, the regression coefficients associated with the position of the large-reward target and reward delay showed appropriate signs expected for the temporally discounted values (10/10 and 8/10 cases for CD and VS, respectively). The results were relatively unchanged when the level of statistical significance was relaxed to $p = 0.1$ in order to reduce the likelihood of type II error (15/15 and 10/12 cases for CD and VS). In addition, all of 13 neurons (5 in the CD, 8 in the VS) that showed the significant effects of the magnitude and delay of the chosen reward showed opposite signs for their regression coefficients. When the criterion for statistical significance was relaxed to $p = 0.1$, the number of neurons increased to 17 (8 in the CD, and 9 in the VS), but all of them still showed opposite signs for the regression coefficients related to the magnitude and delay of the chosen reward.

Temporal evolution of value signals in the striatum

To investigate how the signals related to the temporally discounted values changed during the cue period, we applied the regression model (model 1) including the animal's choice and multiple variables related to the temporally discounted values to the spike rates estimated with a 200-ms sliding window shifted in 25-ms steps. The time course of signals related to the sum and difference in the temporally discounted values for the left and right targets emerged immediately and nearly simultaneously in the CD and VS. This was true regardless of whether the results from these two areas were compared using the fraction of neurons showing significant effects of each variable (Figure 6A) or the proportion of the variance in neural activity attributed to a given variable (coefficient of partial determination, CPD; Figure 6B). Average CPD for the difference in the temporally discounted values reached their maximum values 200 and 175 ms from the cue onset for the CD and VS, whereas the values for the sum reached their maximum 225 ms and 250 ms from the cue onset for the CD and VS, respectively (Figure 6B). In contrast, signals related to the difference in temporally discounted values for the chosen and unchosen targets and the animal's choice arose more slowly and gradually during the cue period (Figure 6). In both CD and VS, the latencies of the signals related to the sum and difference in the temporally discounted value for the left and right targets were both shorter than those related to the animal's choice (Kolmogorov-Smirnov tests, $p < 0.05$; Figure S1, A and B). The latencies of the signals related to the difference in the discounted values for the chosen and unchosen targets and the animal's choice were not statistically different in either CD ($p > 0.3$) or VS ($p > 0.2$), and none of the signals related to the values or choice showed significant differences in their latencies between the CD and VS ($p > 0.1$).

It has been shown that the signals related to the value of chosen option arise in the primate orbitofrontal cortex immediately after the stimulus onset (Padoa-Schioppa and Assad, 2006), whereas other studies found that similar signals might develop more gradually in the striatum (Lau and Glimcher, 2008; Kim et al., 2009b) as well as in the rodent frontal cortex (Sul et al., 2010). We found that the time course of these so-called chosen value signals might change depending on whether the sum of the temporally discounted values for the two targets was included in the regression model or not. In particular, when the sum of the temporally discounted values was omitted from the model, activity changes related to the temporally discounted values of the chosen target appeared much earlier (see Figure S1, C

and D). Therefore, it is important to distinguish the neural activity related to the value of the chosen target from those related to the sum of the values for alternative targets.

Cell types and value coding in the striatum

In the present study, the neurons were not classified into distinct categories, since the distribution of baseline firing rates and spike widths, which have been linked to anatomical cell types in the striatum (Apicella, 2007; Berke, 2008; Gage et al., 2010), did not show clear boundaries (Figure 7). Nevertheless, we tested whether the neural activity related to temporally discounted values varied according to the baseline firing rate. We divided the neurons depending on whether their baseline activity during the last 1-s of the inter-trial interval was higher than 3 spikes/s, since this criterion was often used to identify tentative medium spiny neurons (Schultz et al., 1992; Hassani et al., 2001; Cromwell and Schultz, 2003). The baseline activity was larger than this threshold for many of the neurons tested in our study, and this was more likely in the CD (60 neurons, 64.5%) than in the VS (34 neurons, 37.8%; χ^2 -test, $p < 0.001$). The average baseline firing rate in the CD (9.6 ± 1.1 spikes/s) was also significantly higher than that in the VS (4.6 ± 0.7 spikes/s; t-test, $p < 10^{-3}$). Despite this possible difference in the proportion of inhibitory interneurons in the CD and VS, the proportion of neurons that significantly modulated their activity according to the sum of the temporally discounted values or their difference did not vary significantly with the average firing rates in either CD or VS (Table S3). For some neurons (56 and 65 neurons in CD and VS, respectively), we also recorded their spike waveforms and measured spike widths (Figure 7A). To test whether striatal activity related to temporally discounted values changes with spike width, we compared the percentage of neurons showing significant modulations related to the temporally discounted values, separately for the neurons with spikes width longer or shorter than the median spike width in each area (0.28 and 0.30 ms for the CD and VS, respectively). Similar to the results based on baseline firing rate, the proportion of neurons with significant modulations related to temporally discounted values did not differ for these two groups, in either the CD or VS (Figure 7B; Table S3).

DISCUSSION

Intertemporal choices of humans and other animals are relatively well accounted for by temporal discounting models, suggesting that the subjective value or utility of reward is discounted by its delay. We found that neurons in the primate striatum encode the subjective value of reward temporally discounted by its delay. Previous studies have shown that the magnitude and delay of the reward expected from the animal's action influence the activity of some neurons in the ventral striatum of domestic chicks (Izawa et al., 2005) and rodents (Roesch et al., 2009). However, these studies have not demonstrated the antagonistic effects of reward magnitude and delay, which is required for computing temporally discounted values. To our knowledge, the results from the present study provide the first evidence for signals related to temporally discounted values at the level of individual neurons in the striatum during intertemporal choice. We also found that two different types of signals related to temporally discounted values are partially segregated in the dorsal and ventral striatum. First, the signals related to the difference between the temporally discounted values for the two alternative targets, which reliably predicts the animal's choice, were more robust and found more frequently in the dorsal striatum. Second, the signals related to the direction of the animal's eye movement during intertemporal choice were found only in the dorsal striatum. Therefore, the dorsal striatum is likely to play a more important role in choosing a particular action based on temporally discounted values than the ventral striatum.

Previous single-neuron recording studies in the primate striatum have also shown that signals related to specific movements are largely confined to the dorsal striatum, including the caudate nucleus and putamen, whereas reward-related signals tend to be distributed

evenly across different subdivisions of the striatum (Apicella et al., 1991; Schultz et al., 1992; Williams et al., 1993; Bowman et al., 1996; Hassani et al., 2001; Cromwell and Schultz, 2003; Kawagoe et al., 1998; Ding and Hikosaka, 2006; Kobayashi et al., 2007). In some of these studies, the position of the target associated with a large reward was fixed for a block of trials during an instructed delay task, while the direction of the required movement was selected randomly (Kawagoe et al., 1998; Ding and Hikosaka, 2006; Kobayashi et al., 2007). These studies have found that some neurons in the caudate nucleus change their activity according to the position of the target associated with a large reward. In reinforcement learning theory, the value of reward expected from a particular action are referred to as action values (Sutton and Barto, 1998), and could be used to select an action to maximize reward intake. Indeed, it has been shown that during a free-choice task, some neurons in the dorsal striatum change their activity according to the action values of specific movements (Samejima et al., 2005; Lau and Glimcher, 2008; Kim et al., 2009b). These results suggest that the dorsal striatum might play an important role in selecting an action with the most desirable outcomes, when the likelihood of reward from each action needs to be estimated from experience (O'Doherty et al., 2004; Tricomi et al., 2004; Kimchi and Laubach, 2009). The results from the present study show that the dorsal striatum might also contribute to intertemporal choice by encoding the difference in the temporally discounted values for alternative outcomes. In addition, neurons in both CD and VS encoded the sum of the temporally discounted values with a time course similar to their difference, suggesting that the signals related to the temporally discounted values of the two targets were combined heterogeneously across different striatal neurons, similar to the activity related to action values in the posterior parietal cortex (Seo et al., 2009). Moreover, neurons in the VS tended to encode the sum of the temporally discounted values more often than the CD neurons. Therefore, this difference between the CD and VS is consistent with the actor-critic model of the basal ganglia in which the ventral striatum uses the state value functions to guide the action selection in the dorsal striatum (O'Doherty et al., 2004; Atallah et al., 2007).

In contrast to the signals related to the sum and difference of temporally discounted values associated with the two alternative targets, the signals related to the animal's choice and its temporally discounted value increased more gradually during the cue period. The time course of these two signals was similar, suggesting that striatal activity encoding the subjective value of the chosen action is closely related to the process of action selection. Neural activity related to the reward expected from the action chosen by the animal have been found in both the dorsal and ventral striatum (Apicella et al., 1991; Schultz et al., 1992; Williams et al., 1993; Bowman et al., 1996; Hassani et al., 2001; Cromwell and Schultz, 2003; Roesch et al., 2009; Kawagoe et al., 1998; Ding and Hikosaka, 2006; Kobayashi et al., 2007; Kim et al., 2009b). For example, it has been shown that some striatal neurons change their activity similarly in anticipation of reward, regardless of the direction of the movement produced by the animal (Hassani et al., 2001; Cromwell and Schultz, 2003; Ding and Hikosaka, 2006; Kobayashi et al., 2007) or regardless of whether the animal is required to execute or withhold a particular movement in a go/no-go task (Schultz et al., 1992). Similarly, during a free-choice task in which the reward probabilities were dynamically adjusted, some neurons in the striatum tracked the probability of reward expected from the action chosen by the animal, and these so-called chosen-value signals tend to emerge in the striatum largely after the animal executes its chosen action and approximately when the outcome from the animal's action is revealed (Lau and Glimcher, 2008; Kim et al., 2009b). During reinforcement learning, chosen-value signals can be used to compute reward prediction error, namely the difference between the expected and actual rewards, and play an important role in updating the animal's decision-making strategies. Therefore, when the outcomes of chosen actions are uncertain and the chosen values can be estimated only through experience, signals related to chosen values and outcomes might be combined in the striatum to compute reward prediction errors (Kim et al., 2009b). In the present study, the

signals related to the temporally discounted value of reward developed in both divisions of the striatum before the animal's choice was revealed, even though the outcome of the animal's choice was already known. This suggests that striatal signals related to the value of chosen action might be an integral part of the action selection process rather than only contributing to the computation of reward prediction errors.

Although the present study focused on the signals related to the temporally discounted values in the striatum, signals related to reward delays also exist in other brain areas. In particular, neurons in areas directly connected with the striatum, such as the prefrontal cortex (Kim et al., 2008; Roesch and Olson, 2005; Roesch et al., 2006), ventral tegmental area, and substantia nigra pars compacta (Roesch et al., 2007; Kobayashi and Schultz, 2008), often modulate their activity according to the delay of expected reward. The properties and time course of signals related to the temporally discounted values in the dorsal striatum are also similar to those in the dorsolateral prefrontal cortex identified during intertemporal choice task (Kim et al., 2008, 2009a), suggesting that the fronto-cortico-striatal network plays an important role in evaluating the desirability of alternative outcomes and selecting actions optimally (Haber et al., 2006). Nevertheless, whether and how each of these multiple brain areas makes a unique contribution to the decision making process requires further studies. For example, compared to the value signals in the striatum, chosen value signals might arise in the orbitofrontal cortex more rapidly and immediately after the alternative options are specified (Padoa-Schioppa and Assad, 2006), raising the possibility that chosen value signals are first computed in the prefrontal cortex and transmitted to the striatum. However, the time course of the chosen value signals might change depending on other variables included in the regression model. In addition, the precise time course of value signals is likely to vary across trials, so the value-related signals in multiple brain areas need to be monitored simultaneously in order to understand their precise temporal relationship.

The functions of different classes of striatal neurons in decision making also remain poorly understood. The majority of the neurons in the striatum are the projection neurons referred to as medium spiny neurons (MSN). In addition, the striatum contains several different types of inhibitory interneurons that can be distinguished neurochemically. They include cholinergic aspiny neurons, parvalbumin-positive neurons, calretinin-positive interneurons, and neurons that express neuropeptide Y and somatostatin (Tepper and Bolam, 2004; Kreitzer, 2009). We found that the baseline firing rate was higher in the CD than in the VS, and this might due to the lack of parvalbumin-positive neurons in the ventral striatum (Parent et al., 1996; Waldvogel and Faull, 1993), since parvalbumin positive neurons tend to display higher firing rates than MSN (Berke, 2008; Berke et al., 2004; Sharott et al., 2009). However, in the present study, the signals related to the temporally discounted values did not vary with the firing rates or spike widths. Given their anatomical and biochemical specificity, it would be important to understand better the contribution of different classes of striatal neurons in value coding and action selection, which needs to be further examined in future studies.

EXPERIMENTAL PROCEDURES

Animal preparations

Two male rhesus monkeys (H and J; body weight, 9.3~10.6 kg) were used. During the experiment, the animal was seated in a primate chair with its head fixed and faced a computer screen. The animal's eye position was monitored with a video-based eye tracking system with a 225Hz sampling rate (ET-49, Thomas Recording, Germany). Single-unit activity was recorded from the dorsal and ventral striatum using a multielectrode recording system (Thomas Recording, Giessen, Germany) and a multichannel acquisition processor

(Plexon Inc, Dallas, TX). All neurons were recorded from the right hemisphere (68 and 90 neurons in the CD and VS, respectively), except 25 neurons recorded from the caudate nucleus of the left hemisphere in monkey H. All the procedures were approved by the Institutional Animal Care and Use Committee at Yale University and conformed to the Public Health Services Policy on Humane Care and Use of Laboratory Animals and the Guide for the Care and Use of Laboratory Animals.

Intertemporal choice task

The animal performed an intertemporal choice task and a control task in alternating blocks of 40 trials. During the intertemporal choice task, the animals began each trial by fixating a white square presented at the center of a computer screen. After a 1-s fore-period, two peripheral targets were presented, and the animal was required to shift its gaze towards one of the targets within 1 s, when the central square was extinguished after a 1-s cue period. One of the peripheral targets was green (TS) and delivered a small reward (0.26 ml of apple juice) when it was chosen, whereas the other target was red (TL) and delivered a large reward (0.4ml of apple juice). Each target was surrounded by a variable number of yellow dots ($n=0, 2, 4, 6, \text{ or } 8$) which indicated the delay (1 s/dot) before reward delivery after the animal fixated its chosen target. During this reward delay period, the animal was required to fixate the chosen target while the yellow dots disappeared one at a time, but was allowed to re-fixate the target within 0.3 s without any penalty. The inter-trial interval was 1 s after the animal chose TL, but was padded to compensate for the difference in the reward delays for the two targets after the TS was chosen, so that the onset of the next trial was not influenced by the animal's choice. The reward delay was 0 or 2 s for TS, and 0, 2, 4, 6 or 8 s for TL. Each of the 10 possible delay combinations for the two targets was presented 4 times in a given block in a pseudo-random order with the position of the TL counter-balanced. The control task was identical to the intertemporal choice task, except for the following two changes. First, the central fixation target was either green or red, and this indicated the color of the peripheral target the animal was required to choose. Second, the animal was always rewarded by the same amount of reward (0.13ml) without any delay after it fixated the correct target.

Analysis of behavioral data

The temporally discounted value of the reward from target x is denoted as $DV(A_x, D_x)$, where A_x and D_x indicate the magnitude and delay of the reward from target x . In the model used to analyze the animal's choices, the probability that the animal would choose TS was given by the logistic function of the difference in the temporally discounted values for the two targets, as follows.

$$p(\text{TS}) = \sigma[\beta \{DV(A_{\text{TS}}, D_{\text{TS}}) - DV(A_{\text{TL}}, D_{\text{TL}})\}],$$

where the function $\sigma[z] = \{1 + \exp(-z)\}^{-1}$ corresponds to the logistic transformation, and β is the inverse temperature parameter. The temporally discounted value was determined using a hyperbolic discount function,

$$DV(A_x, D_x) = A_x / (1 + k D_x),$$

or an exponential discount function,

$$DV(A_x, D_x) = A_x \exp(-k D_x),$$

where the parameter k determines the steepness of the discount function. The model parameters (k and β) were estimated using a maximum likelihood procedure as in the previous studies (Kim et al., 2008, 2009a).

Analysis of neural data

We analyzed all the neurons recorded in the caudate nucleus and ventral striatum, as long as they were recorded for more than 2 blocks (80 trials) during the intertemporal choice task. Except for 2 neurons, all neurons were tested at least for 3 blocks (120 trials). The average number of intertemporal choice trials tested for each neuron was 167.4 ± 3.7 and 162.4 ± 4.1 for the CD and VS, respectively. The spike rate during the 1-s cue period was analyzed by applying a series of regression models. For each trial, we first estimated the temporally discounted values by multiplying the magnitude of reward from each target and the discount function (hyperbolic or exponential) for its delay that provided the best fit to the behavioral data in the same session. Next, we used a regression model to test whether the activity was influenced by the difference between the temporally discounted values of the left and right targets ($DV_L - DV_R$), since this is equivalent to the decision variable used by the behavioral model described above. This regression model also included the sum of the temporally discounted values ($DV_{\text{sum}} = DV_L + DV_R$), and the difference in the temporally discounted values for the chosen and unchosen targets ($DV_{\text{chosen}} - DV_{\text{unchosen}}$), in addition to the animal's choice ($C=0$ and 1 for the leftward and rightward choice). In other words,

$$S = a_0 + a_1 DV_{\text{sum}} + a_2 (DV_L - DV_R) + a_3 (DV_{\text{chosen}} - DV_{\text{unchosen}}) + a_4 C, \quad (\text{model 1})$$

where S denotes the spike rate during the cue period. The same model was also applied to the control trials with temporally discounted values replaced by fictitious values calculated as if the reward magnitude and delays were indicated by the target color and the number of yellow dots as in the intertemporal choice task.

In the above regression model, we used the difference in the temporally discounted values for the chosen and unchosen targets rather than the temporally discounted value of the chosen value, in order to minimize the correlation among the regressors. To test whether the neural activity is modulated according to the temporally discounted values of individual targets, we also applied the following two models.

$$S = a_0 + a_1 DV_L + a_2 DV_R + a_3 (DV_{\text{chosen}} - DV_{\text{unchosen}}) + a_4 C, \quad (\text{model 2})$$

$$S = a_0 + a_1 DV_{\text{chosen}} + a_2 DV_{\text{unchosen}} + a_3 (DV_L - DV_R) + a_4 C. \quad (\text{model 3})$$

The set of independent variables in each of these 3 models forms the basis for the same vector space. Therefore, these three models account for the same amount of variance in the neural activity, and are used to test the statistical significance for different independent variables. To test whether the regression coefficients associated with the temporally discounted values of individual targets are significantly correlated, we repeatedly ($n=10,000$) shuffled the spike counts randomly across trials and estimated the p-value from

the frequency of such shuffles in which the correlation coefficient between the regression coefficients exceeded the value obtained from the original data (Figure 4).

To test whether the activity related to temporally discounted values differs for the intertemporal choice and control tasks, we applied a regression model that includes a series of interaction terms between the dummy variable indicating the task performed by the animal and other variables related to the animal's choice and temporally discounted values as follows.

$$S = a_0 + a_1 (DV_L + DV_R) + a_2 (DV_L - DV_R) + a_3 (DV_{\text{chosen}} - DV_{\text{unchosen}}) + a_4 C \\ + a_5 T + a_6 T \times (DV_L + DV_R) + a_7 T \times (DV_L - DV_R) \\ + a_8 T \times (DV_{\text{chosen}} - DV_{\text{unchosen}}) + a_9 T \times C, \quad (\text{model 4})$$

where T denotes the task (0 and 1 for the choice and control task, respectively).

To test whether the activity was modulated by the magnitude and delay of reward expected from a given target, we also applied the following regression model.

$$S = a_0 + a_1 M + a_2 D_L + a_3 D_R + a_4 M_{\text{chosen}} + a_5 D_{\text{chosen}} = a_6 C, \quad (\text{model 5})$$

where M denotes the position of the large-reward target (0 and 1 for the trials in which the large reward was assigned to the leftward and rightward targets, respectively), D_L (D_R) the delay of the reward from the left (right) target, and M_{chosen} and D_{chosen} the magnitude and delay of the reward chosen by the animal. The statistical significance of each regression coefficient was determined with a t-test ($p < 0.05$), and the significance for the effect of the reward delays (D_L and D_R) was adjusted for multiple comparison using the Bonferroni correction.

The standardized regression coefficient (SRC) for the i -th regressor x_i is defined as $a_i \cdot s_i / s_y$, where a_i is the raw regression coefficient, and s_i and s_y the standard deviations of x_i and the dependent variable y . To quantify how strongly neural activity was influenced by a set of regressors, we used the coefficient of partial determination (CPD). The CPD for X_i is defined as the following.

$$\text{CPD}(X_i) = \{ \text{SSE}(X_{-i}) - \text{SSE}(X_{-i}, X_i) \} / \text{SSE}(X_{-i}),$$

where $\text{SSE}(X)$ refers to the sum of squared errors in a regression model that includes a set of regressors X , and X_{-i} a set of all the regressors included in the full model except X_i .

To compare the time course of neural signals related to the sum of the temporally discounted values, their difference, the difference in the temporally discounted values for the chosen and unchosen targets, and the animal's choice (model 1) within each region of the striatum and between the CD and VS, we applied the same regression analysis using a 200-ms window shifted in 25-ms steps. To estimate the latency of signals related to temporally discounted values, we examined the results from this regression analysis in which the center of the window started 0.1 s after cue onset and stopped 0.3 s after the fixation offset. For each neuron, we then defined the latency for a given variable as the first time in which the CPD related to each of these variables exceeds $4 \times$ standard deviation above the mean of the CPD during the baseline period (fore-period) in 3 consecutive time steps. This analysis

produced a latency histogram for each variable separately for CD and VS, and the statistical significance of the difference between two such histograms was evaluated using the Kolmogorov-Smirnov test ($p < 0.05$; Figure S1).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

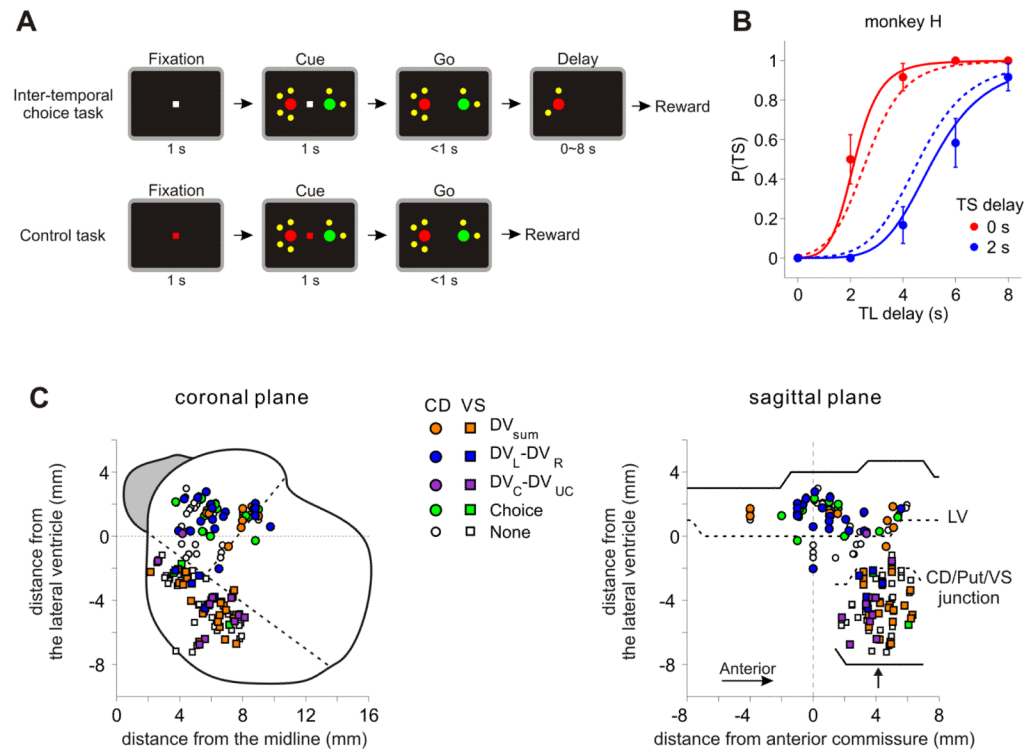
We thank Mark Hammond and Patrice Kurnath for technical assistance. This study was supported by the National Institute of Health (RL1 DA024855, P01 NS048328, and P30 EY000785).

References

- Apicella P. Leading tonically active neurons of the striatum from reward detection to context recognition. *Trends Neurosci.* 2007; 30:299–306. [PubMed: 17420057]
- Apicella P, Ljungberg T, Scarnati E, Schultz W. Responses to reward in monkey dorsal and ventral striatum. *Exp Brain Res.* 1991; 85:491–500. [PubMed: 1915708]
- Atallah HE, Lopez-Paniagua D, Rudy JW, O'Reilly RC. Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat Neurosci.* 2007; 10:126–131. [PubMed: 17187065]
- Ballard K, Knutson B. Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage.* 2009; 45:143–150. [PubMed: 19071223]
- Berke JD. Uncoordinated firing rate changes of striatal fast-spiking interneurons during behavioral task performance. *J Neurosci.* 2008; 28:10075–10080. [PubMed: 18829965]
- Berke JD, Okatan M, Skurski J, Eichenbaum HB. Oscillatory entrainment of striatal neurons in freely moving rats. *Neuron.* 2004; 43:883–896. [PubMed: 15363398]
- Bickel WK, Pitcock JA, Yi R, Angtuaco EJC. Congruence of BOLD response across intertemporal choice conditions: fictive and real money gains and losses. *J Neurosci.* 2009; 29:8839–8846. [PubMed: 19587291]
- Bowman EM, Aigner TG, Richmond BJ. Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J Neurophysiol.* 1996; 75:1061–1073. [PubMed: 8867118]
- Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW, Everitt BJ. Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science.* 2001; 292:2499–2501. [PubMed: 11375482]
- Cromwell HC, Schultz W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol.* 2003; 89:2823–2838. [PubMed: 12611937]
- Ding L, Hikosaka O. Comparison of reward modulation in the frontal eye field and caudate of the macaque. *J Neurosci.* 2006; 26:6695–6703. [PubMed: 16793877]
- Frederick S, Loewenstein G, O'Donoghue T. Time discounting and time preference: a critical review. *J Econ Lit.* 2002; 40:351–401.
- Gage GJ, Stoetzer CR, Wiltchko AB, Berke JD. Selective activation of striatal fast-spiking interneurons during choice execution. *Neuron.* 2010; 67:466–479. [PubMed: 20696383]
- Green L, Myerson J. A discounting framework for choice with delayed and probabilistic rewards. *Psychol Bull.* 2004; 130:769–792. [PubMed: 15367080]
- Gregorios-Pippas L, Tobler PN, Schultz W. Short-term temporal discounting of reward value in human ventral striatum. *J Neurophysiol.* 2009; 101:1507–1523. [PubMed: 19164109]
- Haber SN, Kim KS, Maily P, Calzavara R. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J Neurosci.* 2006; 26:8368–8376. [PubMed: 16899732]

- Hariri AR, Brown SM, Williamson DE, Flory JD, de Wit H, Manuck SB. Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *J Neurosci*. 2006; 26:13213–13217. [PubMed: 17182771]
- Hassani OK, Cromwell HC, Schultz W. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J Neurophysiol*. 2001; 85:2477–2489. [PubMed: 11387394]
- Hwang J, Kim S, Lee D. Temporal discounting and inter-temporal choice in rhesus monkeys. *Front Behav Neurosci*. 2009; 3:9. [PubMed: 19562091]
- Izawa EI, Aoki N, Matsushima T. Neural correlates of the proximity and quantity of anticipated food rewards in the ventral striatum of domestic chicks. *Eur J Neurosci*. 2005; 22:1502–1512. [PubMed: 16190903]
- Kable JW, Glimcher PW. The neural correlates of subjective value during intertemporal choice. *Nat Neurosci*. 2007; 10:1625–1633. [PubMed: 17982449]
- Kalenscher T, Pennartz CMA. Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Prog Neurobiol*. 2008; 84:284–315. [PubMed: 18207301]
- Kawagoe R, Takikawa Y, Hikosaka O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci*. 1998; 1:411–416. [PubMed: 10196532]
- Kim S, Hwang J, Lee D. Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron*. 2008; 59:161–172. [PubMed: 18614037]
- Kim S, Hwang J, Seo H, Lee D. Valuation of uncertain and delayed rewards in primate prefrontal cortex. *Neural Networks*. 2009a; 22:294–304. [PubMed: 19375276]
- Kim H, Sul JH, Huh N, Lee D, Jung MW. Role of striatum in updating values of chosen actions. *J Neurosci*. 2009b; 29:14701–14712. [PubMed: 19940165]
- Kimchi EY, Laubach M. Dynamic encoding of action selection by the medial striatum. *J Neurosci*. 2009; 29:3148–3159. [PubMed: 19279252]
- Kirby KN, Petry NM. Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction*. 2004; 99:461–471. [PubMed: 15049746]
- Kobayashi S, Kawagoe R, Takikawa Y, Koizumi M, Sakagami M, Hikosaka O. Functional differences between macaque prefrontal cortex and caudate nucleus during eye movements with and without reward. *Exp Brain Res*. 2007; 176:341–355. [PubMed: 16902776]
- Kobayashi S, Schultz W. Influence of reward delays on responses of dopamine neurons. *J Neurosci*. 2008; 28:7837–7846. [PubMed: 18667616]
- Kreitzer AC. Physiology and pharmacology of striatal neurons. *Annu Rev Neurosci*. 2009; 32:127–147. [PubMed: 19400717]
- Lau B, Glimcher PW. Value representations in the primate striatum during matching behavior. *Neuron*. 2008; 58:451–463. [PubMed: 18466754]
- Luhmann CC, Chun MM, Yi DJ, Lee D, Wang XJ. Neural dissociation of delay and uncertainty in intertemporal choice. *J Neurosci*. 2008; 28:14459–14466. [PubMed: 19118180]
- Madden GJ, Petry NM, Badger GJ, Bickel WK. Impulsive and self-control choices in opioid-dependent patients and non-drug-using control patients: drug and monetary rewards. *Exp Clin Psychopharmacol*. 1997; 5:256–262. [PubMed: 9260073]
- McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD. Time discounting for primary rewards. *J Neurosci*. 2007; 27:5796–5804. [PubMed: 17522323]
- McClure SM, Laibson DI, Loewenstein G, Cohen JD. Separate neural systems value immediate and delayed monetary rewards. *Science*. 2004; 306:503–507. [PubMed: 15486304]
- Mitchell SH. Measures of impulsivity in cigarette smokers and non-smokers. *Psychopharmacol*. 1999; 146:455–464.
- O’Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*. 2004; 304:452–454. [PubMed: 15087550]
- Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature*. 2006; 441:223–226. [PubMed: 16633341]
- Parent A, Fortin M, Côté PY, Cicchetti F. Calcium-binding proteins in primate basal ganglia. *Neurosci Res*. 1996; 25:309–334. [PubMed: 8866512]

- Pine A, Seymour B, Roiser JP, Bossaerts P, Friston KJ, Curran HV, Dolan RJ. Encoding of marginal utility across time in the human brain. *J Neurosci*. 2009; 29:9575–9581. [PubMed: 19641120]
- Reynolds BA. Review of delay-discounting research with humans: relations to drug use and gambling. *Behav Pharmacol*. 2006; 17:651–667. [PubMed: 17110792]
- Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*. 2007; 10:1615–1624. [PubMed: 18026098]
- Roesch MR, Olson CR. Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. *J Neurophysiol*. 2005; 94:1469–1497. [PubMed: 15817652]
- Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G. Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J Neurosci*. 2009; 29:13365–13376. [PubMed: 19846724]
- Roesch MR, Taylor AR, Schoenbaum G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron*. 2006; 51:509–520. [PubMed: 16908415]
- Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. *Science*. 2005; 310:1337–1340. [PubMed: 16311337]
- Schultz W, Apicella P, Scarnati E, Ljungberg T. Neuronal activity in monkey ventral striatum related to the expectation of reward. *J Neurosci*. 1992; 12:4595–4610. [PubMed: 1464759]
- Seo H, Barraclough DJ, Lee D. Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *J Neurosci*. 2009; 29:7278–7289. [PubMed: 19494150]
- Sharott A, Moll CKE, Engler G, Denker M, Grün S, Engel AK. Different subtypes of striatal neurons are selectively modulated by cortical oscillations. *J Neurosci*. 2009; 29:4571–4585. [PubMed: 19357282]
- Sul JH, Kim H, Huh N, Lee D, Jung MW. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*. 2010; 66:449–460. [PubMed: 20471357]
- Sutton, RS.; Barto, AG. Reinforcement learning: an introduction. Cambridge: MIT Press; 1998.
- Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci*. 2004; 7:887–893. [PubMed: 15235607]
- Tepper JM, Bolam JP. Functional diversity and specificity of neostriatal interneurons. *Curr Opin Neurobiol*. 2004; 14:685–692. [PubMed: 15582369]
- Tricomi EM, Delgado MR, Fiez JA. Modulation of caudate activity by action contingency. *Neuron*. 2004; 41:281–292. [PubMed: 14741108]
- Vuchinich RE, Simpson CA. Hyperbolic temporal discounting in social drinkers and problem drinkers. *Exp Clin Psychopharmacol*. 1998; 6:292–305. [PubMed: 9725113]
- Waldvogel HJ, Faull RLM. Compartmentalization of parvalbumin immunoreactivity in the human striatum. *Brain Res*. 1993; 610:311–316. [PubMed: 8319092]
- Weber BJ, Huettel SA. The neural substrates of probabilistic and intertemporal decision making. *Brain Res*. 2008; 1234:104–115. [PubMed: 18710652]
- Williams GV, Rolls ET, Leonard CM, Stern C. Neuronal responses in the ventral striatum of the behaving macaque. *Behav Brain Res*. 1993; 55:243–252. [PubMed: 8395182]
- Wittmann M, Leland DS, Paulus MP. Time and decision making: differential contribution of the posterior insular cortex and the striatum during a delay discounting task. *Exp Brain Res*. 2007; 179:643–653. [PubMed: 17216152]
- Xu L, Liang ZY, Wang K, Li S, Jiang T. Neural mechanism of intertemporal choice: from discounting future gains to future losses. *Brain Res*. 2009; 1261:65–74. [PubMed: 19185567]

**Figure 1.**

Intertemporal choice task and the locations of recorded neurons. **A.** Spatio-temporal sequences of the intertemporal choice and control tasks. **B.** Probability of choosing the small-reward target (TS) plotted as a function of the delays for the large-reward (TL) and small-reward (TS) targets in an example session. These data were best fit by the hyperbolic discount function with $k=0.23 \text{ s}^{-1}$. Solid and dotted lines indicate the predictions from the best-fitting hyperbolic and exponential discount functions, respectively. Error bars, SEM. **C.** Locations of neurons recorded in the striatum projected onto coronal and sagittal planes. Colors indicate the variables that significantly modulated the activity of each neuron (DV_{sum} , sum of temporally discounted values; DV_L/DV_R , temporally discounted value for leftward/rightward target; DV_C/DV_{UC} , temporally discounted value for chosen/unchosen target). When the neuron encoded multiple variables, the variable with the maximum coefficient of partial determination (CPD) is indicated. The outline of the striatum shown in the coronal plane was obtained from an MR image corresponding to the level indicated by the arrow in the sagittal plane. Dotted lines in the coronal plane, border between the caudate nucleus (CD), putamen, and ventral striatum (VS); dotted lines in the sagittal plane, ventral tip of the lateral ventricle (LV) and the T-junctions between CD, putamen, and VS, relative to the LV at the level of anterior commissure.

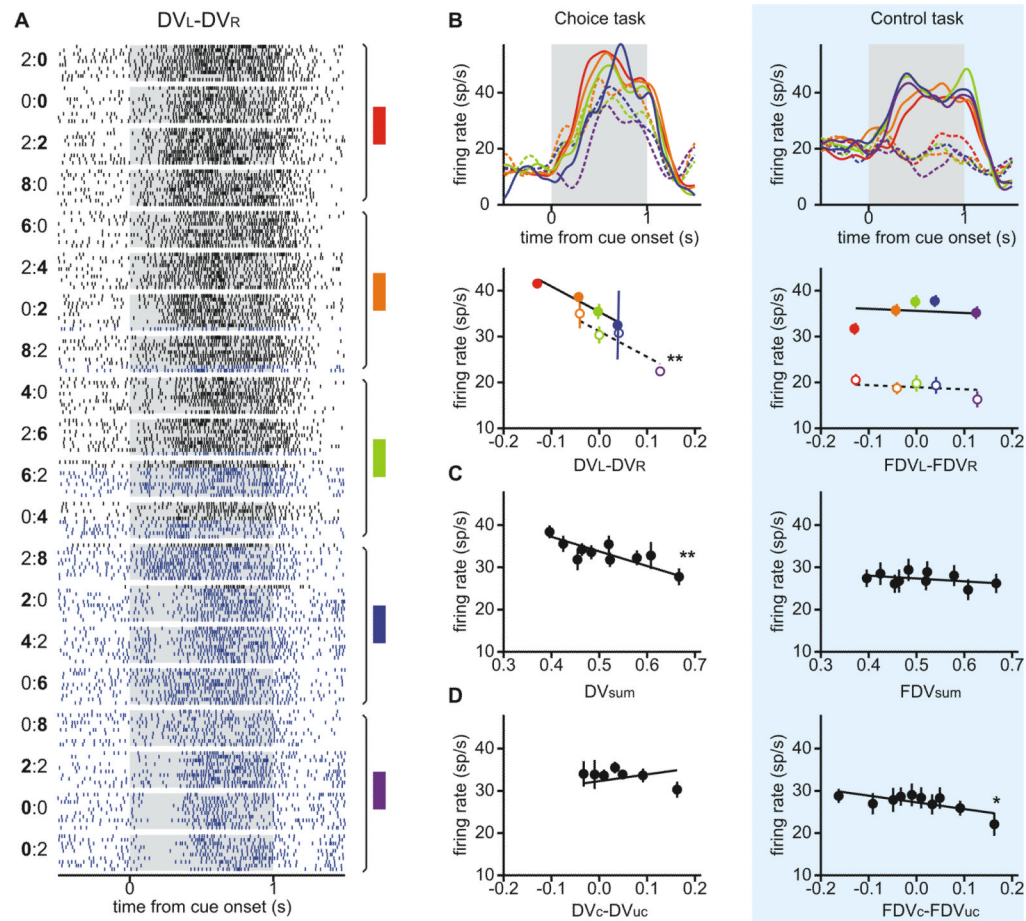


Figure 2.

An example neuron in the caudate nucleus encoding the sum of the temporally discounted values for the two targets and their difference. **A.** Raster plots for trials first grouped by the magnitudes and delays of rewards from the two targets and sorted by the difference in the temporally discounted values (DV) between them. A pair of numbers to the left indicate the reward delays for the two targets with bold typeface used to indicate the delay for the large reward (e.g., “0:4” corresponds to 0 and 4 s delays for leftward small-reward target and rightward large-reward target). Blue and black rasters indicate the trials in which the animal chose the left and right targets, respectively. Colored rectangles and vertical line segments to the right indicate a set of trials grouped together to calculate average activity shown in **B.** **B.** Spike density functions (SDF; top) and firing rates during the cue period (bottom) averaged according to the difference in the temporally discounted values for the two targets during intertemporal choice (left) and control (right) tasks. Empty (filled) circles and dotted (solid) lines denote the activity in trials in which the animal chose the left (right) target. FDV, fictitious temporally discounted value. **C.** Firing rates during the cue period averaged according to the sum of the temporally discounted values during the two tasks. **D.** Firing rates during the cue period averaged according to the difference in the temporally discounted values for the chosen and unchosen targets. Lines in **B–D** are derived from a regression model (model 1) by fixing the values of other regressors at their means. Asterisks indicate that the relationship was statistically significant (*, $p < 0.05$; **, $p < 0.001$). Error bars, SEM.

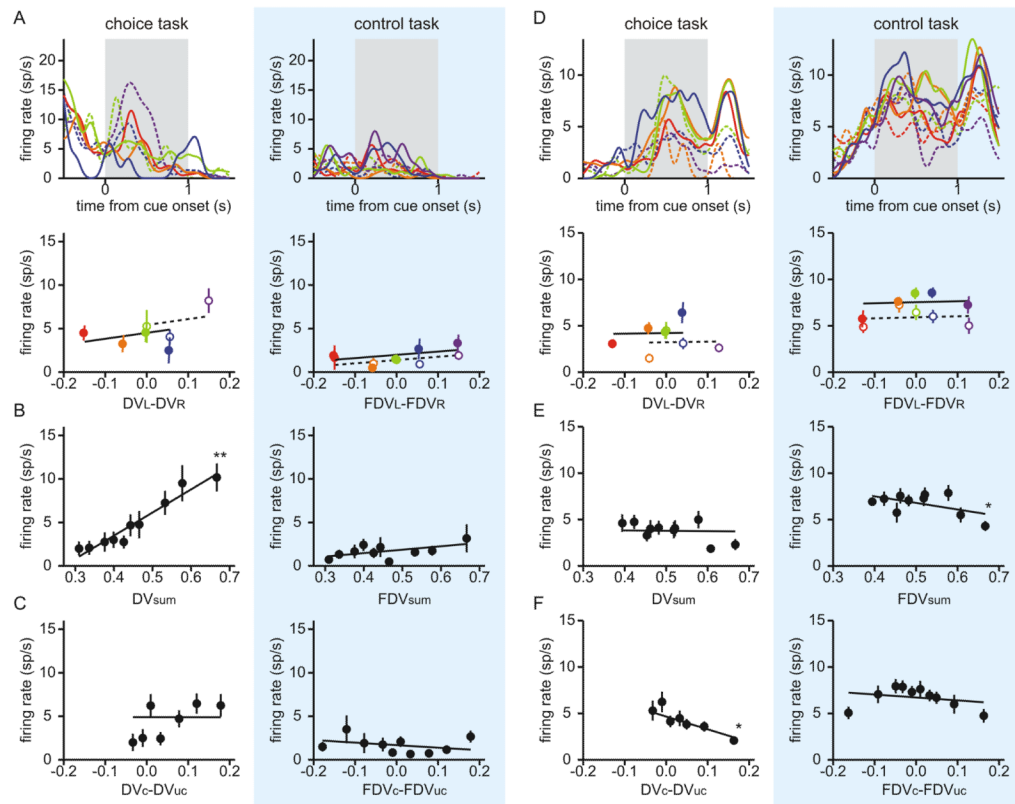


Figure 3.

Two example neurons in the ventral striatum encoding the sum of the temporally discounted values for the two targets (A–C) or the difference in the discounted values for chosen and unchosen targets (D–F). Same format as in Figure 2B–D.

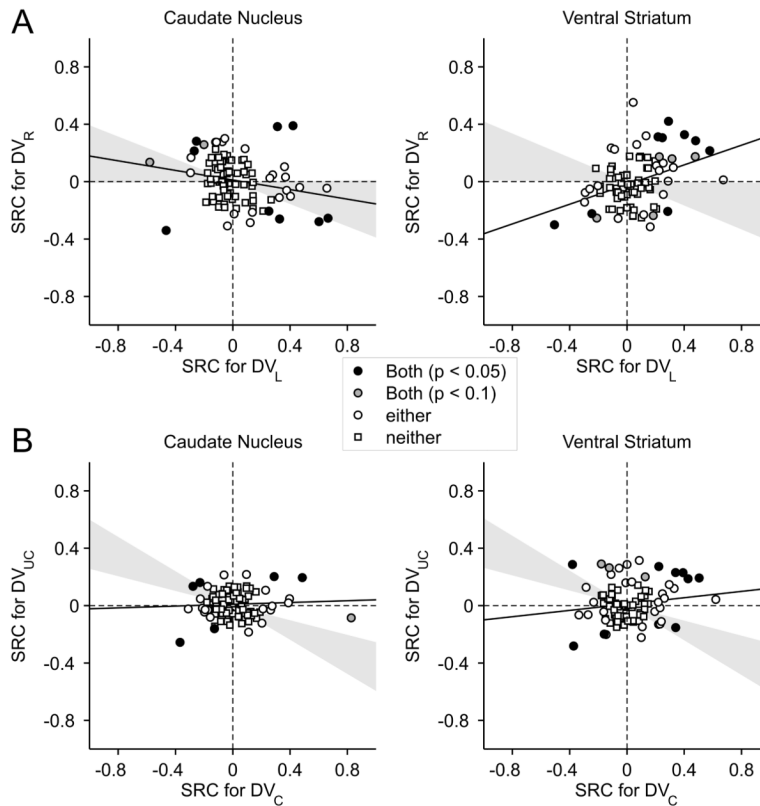


Figure 4.

Population summary of activity related to temporally discounted values in the CD and VS. Scatter plots show the standardized regression coefficients (SRC) associated with the temporally discounted values of the left and right targets (**A**) or chosen and unchosen targets (**B**). Circles correspond to the neurons for which the effect of the discounted value was significant for at least one of the variables ($p < 0.05$), whereas squares indicate the neurons in which the effect was not significant for either variable. Circles filled in gray and black indicate the neurons in which the effect was significant for both variables at the significance level of 0.1 and 0.05, respectively. Gray area corresponds to the 95% confidence interval for the correlation coefficient obtained from the permutation test.

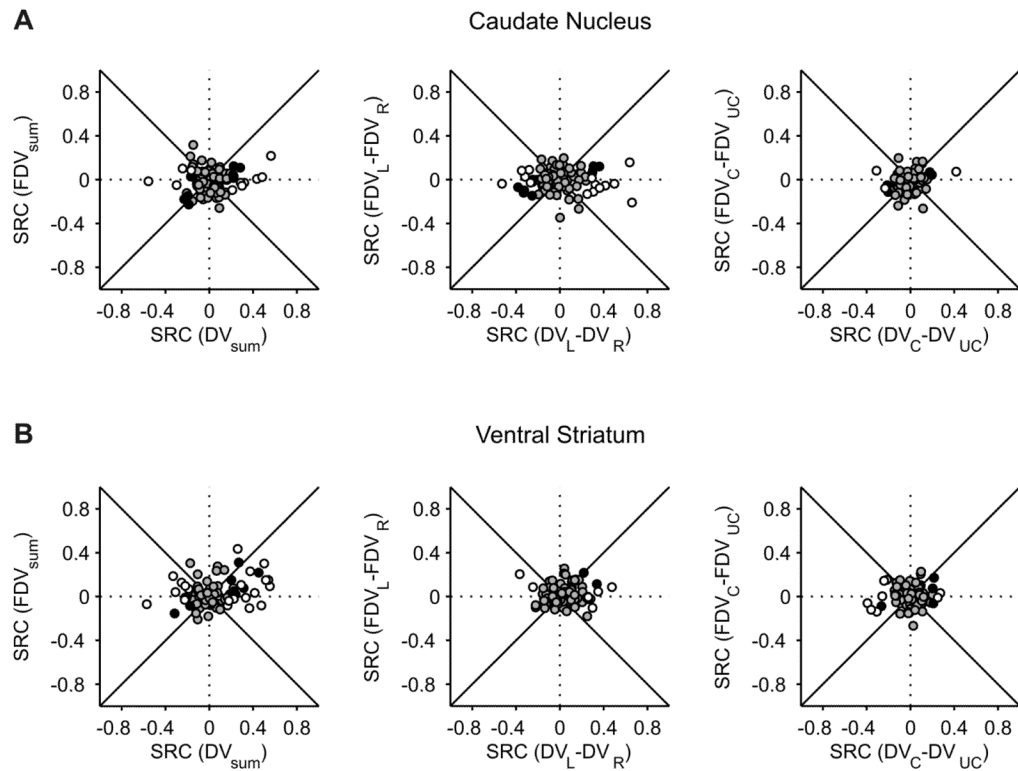


Figure 5. Activity related to temporally discounted values during intertemporal choice task (abscissa) and fictitious temporally discounted values during control (ordinate) task for the caudate nucleus (**A**) and ventral striatum (**B**). The results are shown separately for the sum of the temporally discounted values for left and right targets (left), their difference (middle), and the difference in the temporally discounted value for chosen and unchosen target (right). Empty circles, neurons showing significant interaction between temporally discounted values and task; black disks, neurons showing only the main effect of temporally discounted values; gray disks, neurons without any significant effects of temporally discounted values.

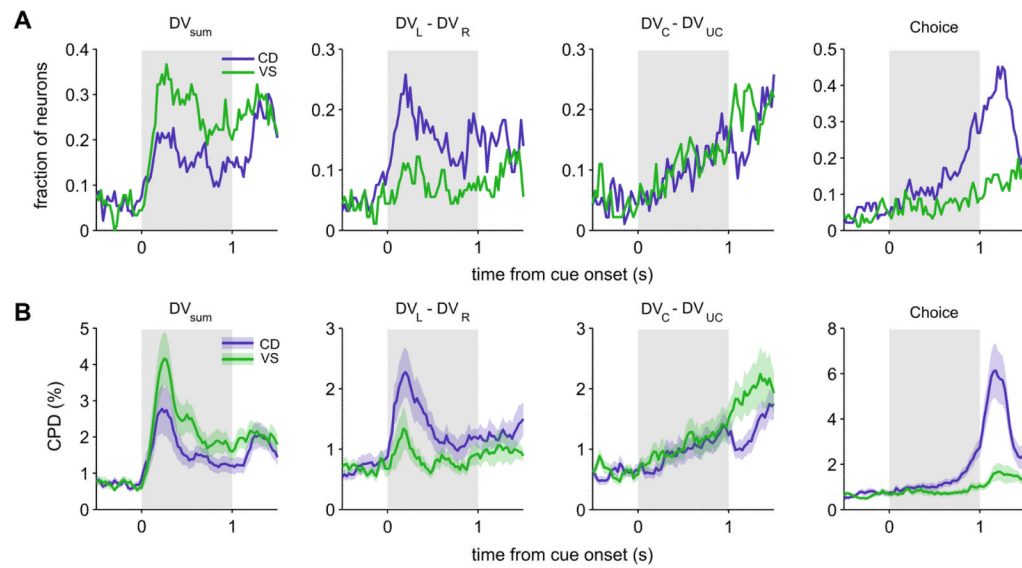


Figure 6. Time course of neural activity related to the animal's choice and temporally discounted values. **A.** Fraction of neurons in CD and VS that significantly modulated their activity according to the sum of the temporally discounted values for left and right targets, their difference, the difference in the temporally discounted values for chosen and unchosen targets, and the animal's choice. **B.** Population average of the coefficient of partial determination (CPD) for the same variables. Shaded areas, \pm SEM.

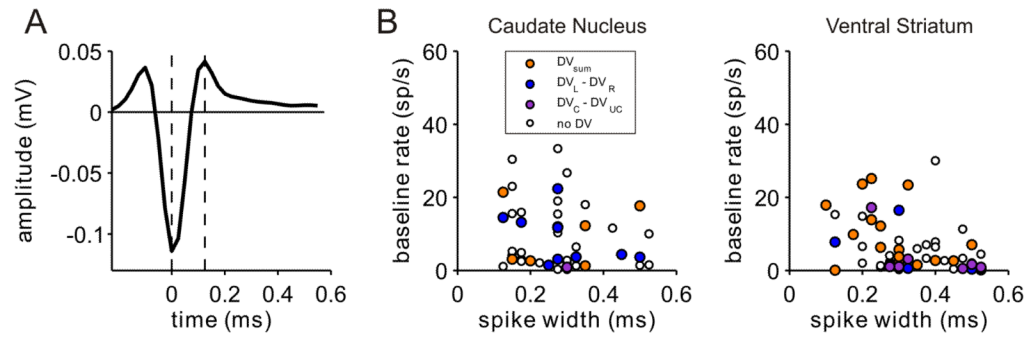


Figure 7. Effects of spike width and baseline firing rate. **A.** Example waveform of a neuron recorded in CD. The spike width (distance between the vertical dotted lines) was 0.125 ms for this neuron. **B.** Relationship between baseline firing rate and spike width. Colors indicate the variable that modulated the activity of each neuron most strongly according to CPD.

Table 1

Number of neurons (and their percentages) in the caudate nucleus (CD) and ventral striatum (VS) that significantly modulated their activity only according to the sum of the temporally discounted values of the left and right targets (Σ), their difference (Δ_{LR}), the difference in the temporally discounted values of the chosen and unchosen targets (Δ_{CU}), the animal's choice (C), and their various combinations.

	CD	VS
Σ	7 (7.5)	20 (22.2)
Δ_{LR}	8 (8.6)	5 (5.6)
Δ_{CU}	1 (1.1)	5 (5.6)
C	9 (9.7)	1 (1.1)
$\Sigma+\Delta_{LR}$	4 (4.3)	2 (2.2)
$\Sigma+\Delta_{CU}$	2 (2.2)	4 (4.4)
$\Sigma+C$	3 (3.2)	2 (2.2)
$\Delta_{LR}+\Delta_{CU}$	1 (1.1)	1 (1.1)
$\Delta_{LR}+C$	5 (5.4)	0 (0.0)
$\Delta_{CU}+C$	1 (1.1)	0 (0.0)
$\Sigma+\Delta_{LR}+\Delta_{CU}$	0 (0.0)	1 (1.1)
$\Sigma+\Delta_{LR}+C$	4 (4.3)	1 (1.1)
$\Sigma+\Delta_{CU}+C$	0 (0.0)	1 (1.1)
$\Delta_{LR}+\Delta_{CU}+C$	2 (2.2)	0 (0.0)
$\Sigma+\Delta_{LR}+\Delta_{CU}+C$	0 (0.0)	0 (0.0)
None	46 (49.5)	47 (52.2)
Total	93 (100)	90 (100)