Exp Brain Res (2011) 212:593–601
DOI 10.1007/s00221-011-2771-3

RESEARCH ARTICLE

Inter-individual discount factor differences in reward prediction are topographically associated with caudate activation

Keiichi Onoda · Yasumasa Okamoto · Yoshihiko Kunisato · Siori Aoyama · Kazuhiro Shishida · Go Okada · Saori C. Tanaka · Nicolas Schweighofer · Shuhei Yamaguchi · Kenji Doya · Shigeto Yamawaki

Received: 3 April 2011 / Accepted: 12 June 2011 / Published online: 22 June 2011

Abstract In general, humans tend to devalue a delayed reward. Such delay discounting is a theoretical and computational concept in which the discount factor influences the time scale of the trade-off between delay of reward and amount of reward. The discount factor relies on the individual's ability to evaluate the future reward. Using functional magnetic resonance imaging, we investigated brain mechanisms for reward valuation at different individual discount factors in a delayed reward choice task. In the task, participants were required to select small/immediate or large/delayed rewards to maximize the total reward over time. The discount factor for each participant individually was calculated from the behavioral data based on an exponential discounting model. The estimated value of a future reward increases as the expected delivery approaches, so the time course of these estimated values was computed based on each individual's discount factor; each was entered into the regression analysis as an explanatory (independent) variable. After the region of interest was narrowed anatomically to the caudate, a peak coordinate was detected in each individual. A correlation analysis revealed that the location of the peak along the dorsal–ventral axis in the right caudate was positively correlated with the discount factor. This implies that individuals who showed a larger discount factor had peak activations in a more dorsal part of the right caudate associated with future reward prediction. This evidence also suggests that a higher ability to delay reward prediction might be related to activation of the more dorsal caudate.

Keywords Delayed reward · Delay discounting · Caudate · Individual difference · fMRI

K. Onoda · S. Yamaguchi
Department of Neurology, Shimane University, Shimane, Japan

Y. Okamoto · Y. Kunisato · S. Aoyama · K. Shishida · G. Okada · S. Yamawaki (✉)
Department of Psychiatry and Neurosciences, Division of Frontier Graduate School of Biomedical Sciences, Hiroshima University, 1-2-3, Kasumi, Minami-ku, Hiroshima 734-8551, Japan
e-mail: yamawaki@hiroshima-u.ac.jp

S. C. Tanaka
Institute of Social and Economic Research, Osaka University, Osaka, Japan

N. Schweighofer
Department of Physical Therapy and Biokinesiology, University of Southern California, Los Angeles, CA, USA

K. Doya
Department of Computational Neurobiology, Advanced Telecommunication Research Institute International, Kyoto, Japan

Introduction

Reward processing can be investigated when it is situated within a more general framework of incentive processing. Evaluations of both immediate and future outcomes of one's actions are critical requirements for intelligent behavior. One can temporally distinguish incentive anticipation from outcomes, as well as qualitatively distinguish between positive and negative reward. The framework can also accommodate additional variables, such as whether a behavioral response is required prior to or after anticipation in order to obtain a reward, the duration of the anticipation, and the probability of the reward (Knutson and Cooper 2005).

Functional magnetic resonance imaging (fMRI) studies in the last decade have primarily examined neural activation in anticipation of a reward after a delay. Anticipation of monetary, social, and taste rewards all activate brain structures located in the reward system, particularly in the ventral striatum (Breiter et al. 2001; O'Doherty et al. 2002, 2003; McClure et al. 2004; Bromberg-Martin and Hikosaka 2009; Gregorios-Pippas et al. 2009). This implicates the ventral striatum as a universal mediator of reward prediction (Knutson and Cooper 2005). Recently, some studies suggested that the activation of the ventral striatum corresponds to the subjective discounted value (Kable and Glimcher 2007; Marco-Pallares et al. 2010), to an individual's discounting rate in a hyperbolic model (Hariri et al. 2006; Ballard and Knutson 2009), and to effort discounting (Botvinick et al. 2009).

On the other hand, there are some studies that suggest a contribution of the dorsal striatum to reward anticipation. For example, in a study by Wittmann and his colleagues, participants took part in a delay discounting task and were asked to decide between immediate and delayed reward (Wittmann et al. 2007; Wittmann and Paulus 2008). Their results revealed that the neural activation associated with shorter delays relative to longer delays took place in the ventral striatum, and the caudate showed a correlation between behaviorally determined discounting and activation for the contrast of immediate (<1 year) and delayed (>1 year) intervals. Furthermore, some remarkable studies demonstrated that two separate systems are involved in decisions based on temporal discounting (McClure et al. 2004, 2007). In these studies, parts of the limbic system associated with the midbrain dopamine system (including the ventral striatum) are preferentially activated by decisions for immediately available rewards; in contrast, regions of the lateral prefrontal cortex are engaged uniformly by decisions for delayed rewards. The evidence of delay-independent involvement of the striatum strengthens a model which assumes that different pathways of the cortico-striatum circuits may be specialized for reward prediction on different time scales (Doya 2002).

Using a reinforcement learning model, our group has previously investigated brain mechanisms for reward prediction at different time scales in a Markov decision task (Tanaka et al. 2004). Model-based regression analyses using reward prediction errors revealed graded maps of time scale within the striatum. The ventroanterior regions of the striatum were involved in predicting immediate rewards, and dorsoposterior regions were involved in predicting future rewards. These results suggest involvement of different parts of the striatum in reward prediction over different time scales. Furthermore, our group confirmed the separation of the time scale of reward prediction using a delayed reward choice task (Tanaka et al. 2007). In that task, participants were asked to choose an immediate/small or delayed/large reward and needed to actually wait for the presentation of the reward. A model-based analysis based on reinforced learning revealed that the activity of the ventral part of the striatum was correlated with reward prediction at shorter time scales. In contrast, the activity of the dorsal part of the striatum was correlated with reward prediction at longer time scales.

These studies used an exponential model (Samuelson 1937; Koopmans 1960; Lancaster 1963) in which the present value V of reward R after delay D is given by

$$V = R * G(D)$$

where $G(D)$ is a discounting function whose value decreases with increasing delay D. A steep rate of discounting results in an impulsive choice, defined as a more frequent choice of a more immediate reward (Ainslie 1975). A lower rate of discounting results in a less impulsive choice, defined as a more frequent choice of a later reward. In exponential discounting, the reward value V is given by

$$V = R * \gamma^D$$

where γ is the discount factor ($0 < \gamma < 1$). A large decay rate corresponds to a small discount factor (γ closer to 0), and vice versa (γ closer to 1). In previous research, we set γ at 0.6, 0.7, 0.8, 0.9, 0.95, and 0.99 and then found a graded map in the striatum: activity in the ventral portion was correlated with expected future rewards with small γ, while activity in the dorsal part was correlated with expected future rewards with large γ (Tanaka et al. 2007). Our group therefore suggested that there was differential involvement of the cortico-striatum circuits in reward prediction for different time scales. This evidence was obtained by surrogate γ at different, specific values. To investigate the hypothesis that there is a topographic distribution in reward prediction, we investigated the relationship between striatum activation and the inter-individual differences in γ. Given the topographic involvement of striatum in reward prediction, we hypothesized that a comparable result would be obtained from using not only the surrogate γ but also the individual γ.

We had two hypotheses. First, activation of the dorsal striatum would be involved with the time series' change of estimated reward value calculated from an individual's value of his or her discount factor γ. Second, the position of the activation peak would be correlated with the individual's γ.

Materials and methods

Thirty healthy right-handed adults gave informed consent to participate in the experiment, which was conducted with

the approval of the ethics committees of Hiroshima University. The participants averaged 21.7 ± 2.2 (SD) years old, and the female/male ratio was 16/14. Before starting the task, they were screened for previous psychiatric problems by using the mini mental scale (MINI).

The participants performed a multi-step delayed reward choice task (Fig. 1) while lying in an fMRI scanner. In this task, two "tanks" filled with different volumes of water were displayed side by side on a computer screen, and the participants chose between the two tanks. At the beginning of each trial, choosing the tank with the larger volume of water (range: 12–48) signified a choice of the stimulus with the delayed/large reward during that trial. Correspondingly, choosing the tank with smaller water volume (range: 4–10) signified a choice of the stimulus with the immediate/small reward. Every 2 s, the participant selected one tank by pressing a button on the corresponding side. The volume of the selected tank was then decreased at random (range: 2–6) in each step. The water volume did not decrease if the participant did not make a response. Updating the water volume was synchronized with the onset of each fMRI scan. The participants repeated the selection until their chosen tank was empty. (In theory, once a participant had chosen one tank, he or she was free to switch to the other even before a tank became empty. In practice, this would be a waste of time, and participants did not ever switch to a different tank until after the next inter-trial interval.) If the large tank was the one that was emptied, 40 G (virtual reward money) was displayed on that empty tank. For the smaller tank, 10 G was presented on the empty tank. The duration of the reward presentation was 2 s, and the following inter-trial interval (ITI) was also 2 s. A fixation point appeared at the center of the screen during the task, and we instructed the participants to fix their eyes there during the ITI. In the next trial, the initial volumes of the tanks were novel amounts (within the ranges specified above). The position of the tanks (large or small on left or right) was changed randomly at each trial. The participants were asked to maximize their total reward within a total task duration of 18 min and were instructed that the monetary payment would vary depending on their task performance. They performed a few training trials before the fMRI session. We also gave the participants an information that virtual money G was nearly equal Japanese yen. The mean total reward of the task was 2,083 G (SD: 136, min: 1,820, max: 2,310). However, regardless of the real earned reward, the participants were informed that their total reward would be 2,000 G after the fMRI session. The actual payment was fixed at ¥ 2,000. The average number of trials that the participants performed in a fMRI session was 100.6 (SD: 10.4, min: 77, max: 119).

Imaging data were acquired using a Siemens AG 1.5 T scanner. A time course series of 540 volumes per participant was acquired with echo planar imaging sequences (TR: 2000 ms, TE: 48 ms, FOV: 256 mm, matrix size: 128 × 128, slices: 20, thickness: 6 mm, flip angle: 90°). Functional scans lasted for the full duration of the experiment (18 min). Additionally, after this functional scanning, structural scans were acquired using T1-weighted gradient echo pulse sequences (TR: 12 ms, TE: 4.5 ms, FOV: 256 mm, flip angle: 20°).

Imaging data were analyzed using SPM8 software (Wellcome Department of Cognitive Neurology, London, UK). The first three volumes of each participant were discarded because the MRI signal was unsteady. Slice timing correction was performed for each individual set of

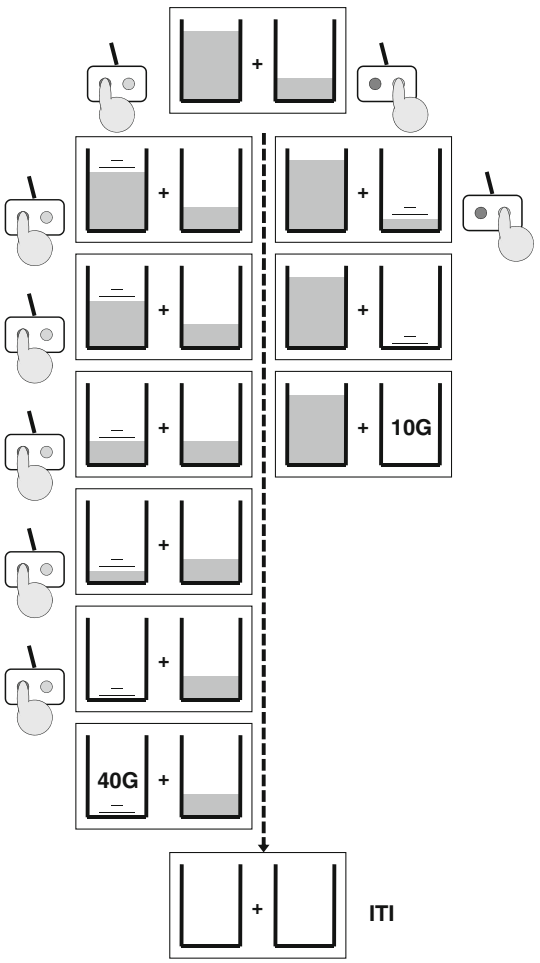

Fig. 1 Schema depicting the multi-step delayed reward choice task. The participant is asked to select either tank by pressing a *button* on the corresponding side. In the example shown on the *right*, the participant chooses the tank with the lower initial volume, and a small reward (10 G) is delivered after two steps. In the example shown on the *left*, the participant chooses the tank with the higher initial volume, and five steps are required to obtain a larger reward (40 G). After the reward was displayed, an inter-trial interval (ITI) preceded the next trial, with new water levels in the two tanks. The participant was asked to maximize the total reward within the scheduled time (approximately 18 min)

functional volumes. The sets were realigned to the first volume and spatially normalized to a standard template based on the Montreal Neurological Institute (MNI) reference brain using DARTEL (Ashburner 2007). Finally, these were smoothed using a 4 mm FWHM Gaussian kernel.

We firstly computed γ based on the choice behavior of each trial. The initial volume ranges of the tanks, M_S and M_L, were 4–10 and 12–48, respectively. The mean water volume decrements S were 4 (range: 2–6) in both tanks. The expected delays until receiving the small and large rewards D_S and D_L, respectively, are given by:

$$D_S = ts * M_S/S \text{ and } D_L = ts * M_L/S$$

where ts is the time step (2 s). Thus, the ranges of expected reward delays (D_S and D_L) for average values of water volume decrements were 2.0–5.0 and 6.0–24.0 s, respectively. Next, tank-selection behavior was modeled by using the softmax function (Sutton and Barto 1998). We computed tank selection probability using the inverse temperature $\beta > 0$, which controls the tendency to exploit or explore (Schweighofer et al. 2007). We applied the following equation:

$$P_L = \frac{e^{\beta * V_L}}{e^{\beta * V_L} + e^{\beta * V_S}}$$

where P_L is the probability of choosing the large reward, and V_L and V_S are the large and small reward values. V_L and V_S replaced $R*\gamma^D$ from the exponential model, where D is the abovementioned D_S and D_L. The γ and β were calculated by fitting the model to the subject's choices using a maximum likelihood method, namely the MATLAB Optimization Toolbox. The maximum likelihood of a hyperbolic model was also computed to compare the two models, and V in the above formula was replaced by $R/(1 + kD)$. We performed a paired t test between the models and confirmed that the exponential model fit the behavioral data better than the hyperbolic model (see "Results").

In the following calculation of the time series of the discounted reward values, we used the individual γs calculated by the above process. We assumed that the participants knew that the mean volume of water decreased. The expected delay D at each scan was calculated by the formula described above ($D = ts*M/S$) and rounded up to a 2 s unit. We defined estimated $V(t)$ as the discounted values for the expected delay in seconds until the reward:

$$V(t) = \sum\nolimits_t R * \gamma^D$$

Here, each term summed as part of $V(t)$ represents a discounted future reward R (10 or 40) at each scan time. Figure 2 (right panels) shows samples of $V(t)$ for three subjects. Because $V(t)$ decreases exponentially with a delay, the discount factor γ controls the time scale of reward prediction. A smaller γ results in steeper discounting of future rewards, leading to a higher probability of choice of the short-term reward.

We estimated each individual's $V(t)$ and used it as the explanatory (independent) variable in a general linear model (GLM) by multiplying the simple event regressor at the timing of the stimulus presentation of each step. The onset of each step corresponded with the onset of each fMRI scan. To remove any effects of factors other than $V(t)$, we added other variables into the 1st-level regression in each participant, namely: reward R (at reward delivery time) and response (long/short) at each step. All explanatory variables were convolved with a canonical hemodynamic response function. Images of parameter estimates were created for each subject and entered into a second-level group analysis using a one-sample t test at a threshold of uncorrected $p < 0.001$ and voxels >20.

To investigate the relationship between values of γ and activated striatum location during delayed reward prediction, we performed correlation analyses of γ versus the dorsal–ventral axis of that activated region. Individual peaks related to $V(t)$ were detected in the right caudate because the foregoing analysis revealed the significant contribution of the right caudate. Only one activation peak per participant within the right caudate was entered into the correlation analysis, and the t value of the individual peak was entered into a nuisance covariate in order to regress out the effect of the model fitness. The region defined by the WFU_PickAtlas toolbox (Maldjian et al. 2003) was used as the region of interest (ROI). Based on the studies by Tanaka et al. (2004, 2007), we predicted that an individual γ would be correlated with the dorsal–ventral axis of the activation peak in the ventral striatum.

Results

Figure 2 (left panels) shows the participants' samples of reward choice behavior as a function of delays and exponential reward discounting model as a function of the delay to reward. The figure reveals that a decrease of γ makes the discounting function steeper, and increases the choice of the small reward. Furthermore, the indifferent line seems to fit the exponential model better than the hyperbolic model. The logarithms of the maximum likelihood for the exponential and hyperbolic models were -27.1 ± 9.7 (mean $\pm$ SD) and -40.7 ± 12.3, respectively. A paired t test revealed that the maximum likelihood of the exponential model was higher than that of the hyperbolic model ($t(29) = 5.91$, $p < 0.001$), indicating that the exponential model better fits the data.

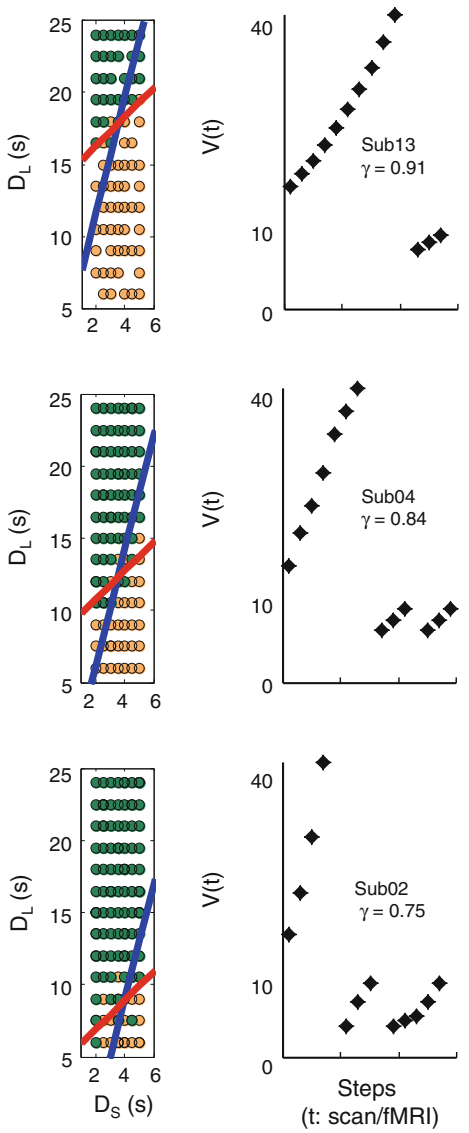

Fig. 2 *Left panels* examples of three participants' reward choices as a function of delays. At each trial, the participant had a choice between a large reward (40 G) and a small reward (10 G). DL and DS show the expected delays for the large and small reward, respectively. The *orange and green circles* show the choice of the large and small reward, respectively. The *red* indifference *line* was obtained by fitting an exponential discounting model, and the *blue line* was obtained by fitting a hyperbolic model. *Right panels* The time series shows samples of $V(t)$ calculated using the exponential model and the intrinsic γ of three participants. This figure shows that the choices differed as a function of the value of γ. The $V(t)$ was used as one of the independent variables in the general linear model for each individual

Mean estimated γ and the SD in the exponential model was 0.84 (SD 0.04, range: 0.75–0.91). We performed a whole-brain one-sample t test for activation related to $V(t)$, calculated based on each individual's γ. Activations of right caudate, left putamen, dorsolateral/ventrolateral prefrontal cortex, middle temporal gyrus, right primary somatosensory cortex, and bilateral primary motor cortex were significantly related to $V(t)$ (Table 1; Fig. 3).

Figure 4 shows individual peak activations related to reward prediction $V(t)$ in the right caudate. We performed correlation analyses between γ and the dorsal–ventral axis of the individual peak. This revealed a positive correlation ($r = 0.39$, $t(27) = 2.2$, $p = 0.037$) after controlling the effect of t value in right caudate (Fig. 5).

Discussion

The aim of the current study was to investigate the neural correlates of individuals' discounting of future rewards. We hypothesized that the discount factor γ, reflecting the relative value of long-term versus short-term prospects, would be correlated with the dorsal–ventral position of peak activation in the dorsal striatum during a delayed reward choice task. The model-based analysis revealed that the right caudate was involved in the reward prediction function $V(t)$ and that the activation peak position was correlated with the individual's γ.

Two different principal models have been proposed to characterize the shape of delay discounting: an exponential model (Samuelson 1937; Koopmans 1960; Lancaster 1963; Schweighofer et al. 2006, 2008) and a hyperbolic model (Ainslie 1975; Laibson 1997; Mazur 2001). Many studies agree that delayed rewards are discounted hyperbolically rather than exponentially in humans (Myerson and Green 1995; Green et al. 1997; Madden et al. 1999; Johnson and Bickel 2002; Reynolds and Schiffbauer 2004; Killeen 2009; McKerchar et al. 2009), rats (Mazur 2007; Mazur and Biondi 2009), and pigeons (Mazur and Biondi 2009). However, some researchers have argued that an exponential discounting model actually provides a better model of the delayed reward function (Sopher and Sheth 2005); furthermore, humans can adopt an optimal discounting strategy that varies in different situations (Schweighofer et al. 2006). Our behavioral data of the delayed reward choice task used in the current study fitted the exponential model better than the hyperbolic model. The standard view of temporal discounting, consistent with the discounting model of Samuelson (1937), is that discounting compensates for the risks associated with waiting for a delayed reward. If the hazard ratio is constant, an exponential decrease in value over time follows directly (Myerson and Green 1995). This assumes that decision-making mechanisms should prescribe optimal choice behavior. Because the task used in the current study had no risk with uncertainty, the participants may have been able to optimally behave to maximize their reward.

We found that activation of the dorsal and ventral striatum (caudate and putamen) was related to reward

Table 1 Activation areas related to $V(t)$

Region (BA)	x	y	z	Size	t
Caudate	22	14	18	53	4.3
Putamen	−26	2	2	30	5.0
DLPFC (8/9)	−46	18	46	67	4.0
VLPFC (47)	−50	36	−2	56	4.4
Insula (13/22)	−50	6	4	70	4.1
PreMC (6/9)	68	8	28	164	4.8
PMC (4/6)	−50	−2	50	23	4.8
PSC (1/2/3)	−60	−18	42	144	6.7
PSC (1/2/40)	68	−22	22	53	4.3
SMA (6)	−4	−8	64	38	5.0
MTG (21)	−56	−28	−10	55	4.4
White matter	−24	−2	24	26	4.6

BA Brodmann area, *DLPFC* dorsolateral prefrontal cortex, *VLPFC* ventrolateral prefrontal cortex, *PreMC* premotor cortex, *PMC* primary motor cortex, *PSC* primary somatosensory cortex, *SMA* supplementary motor area, *MTG* middle temporal gyrus

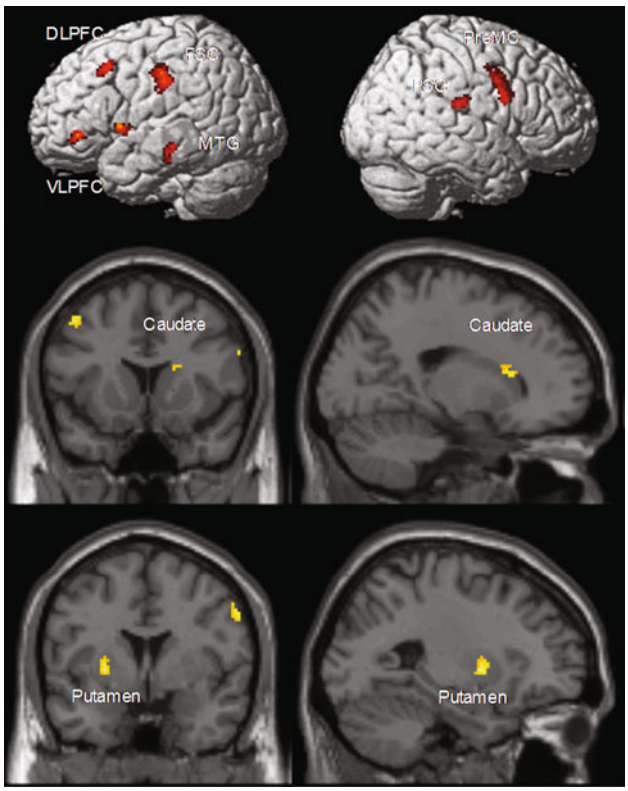

Fig. 3 Activation areas related to $V(t)$ in lateral view (*top*) and coronal/sagital slices (*bottom*). The threshold for the whole-brain one-sample t test was set uncorrected at $p < 0.001$ at the voxel level and voxels >20. *DLPFC* dorsolateral prefrontal cortex, *VLPFC* ventrolateral prefrontal cortex, *PreMC* premotor cortex, *PMC* primary motor cortex, *PSC* primary somatosensory cortex, *MTG* middle temporal gyrus

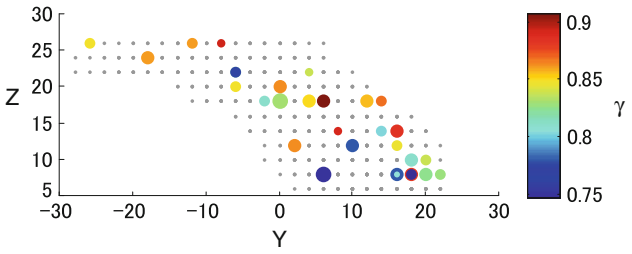

Fig. 4 Relationship between γ and the coordinate of the peak activation involved in $V(t)$ in the *right* caudate. The shape and extent of the caudate are shown by *black dots*, and the *dots* with *colors* show the coordinates of the peak activation involved in an individual's $V(t)$. The sizes of the *colored dots* correspond to the t value; larger size indicates higher t

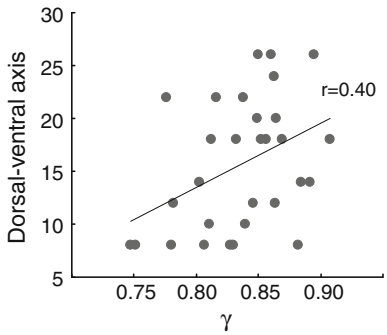

Fig. 5 Correlation between γ and the dorsal–ventral axis of the peak activation in the *right* caudate

prediction $V(t)$. Earlier studies have shown that the ventral striatum plays a key role in the appetitive phase of reward processing (Kalivas and Volkow 2005; Knutson et al. 2005). In humans, some studies have reported that anticipation-of-reward activated brain structures constitute a reward system located in the ventral striatum (Breiter et al. 2001; O'Doherty et al. 2002; O'Doherty et al. 2003; Bromberg-Martin and Hikosaka 2009). The ventral striatum seem to be associated with the future reward magnitude (McClure et al. 2004; Ballard and Knutson 2009), subjective value (Kable and Glimcher 2007; Pine et al. 2009), and preference for immediate over delayed rewards (Hariri et al. 2006). On the other hand, in a study that

involved having to learn which cues yielded a juice reward, participants' ventral striatal activation tracked reward-cue prediction in both passive and active tasks, while dorsal striatal activation only tracked reward cues in the active task (O'Doherty et al. 2004). McClure et al. (2004, 2007) suggested that separate neural systems evaluate and compare immediate and delayed rewards. They used a quasi-hyperbolic model (Phelps and Pollak 1968), sometimes referred to as beta-delta preference. The beta represents the value placed on immediate rewards relative to rewards received at any other point in time, and the delta is simply the discount rate in the standard exponential formula, which treats a given delay equivalently regardless of when it occurs. The betas are associated with limbic reward (including the ventral striatum and medial/orbital prefrontal cortex); in contrast, the deltas are associated with lateral prefrontal and parietal areas. Their evidence is basically consistent with the notion that separate cortico-striatum circuits are involved in immediate and delayed rewards (Doya 2002).

One of our previous studies involved a probabilistic task, during which participants learned to choose an option that afforded monetary gains and to avoid another that incurred losses. We found that a model of short-term reward prediction maximally correlated with activation in the more ventral striatum, whereas a model of long-term reward prediction correlated with activation of more dorsal striatum (Tanaka et al. 2004). We confirmed this gradient of striatum activation in another delayed reward choice task (Tanaka et al. 2007). In this previous study, the surrogate γ set (0.6–0.99) was applied for the calculation of $V(t)$ in all participants; in contrast, the current study used the individual γs computed on the basis of the behavioral data from each participant. We found that participants with larger γ had a peak activation in the more dorsal part of the right caudate, a result that may partially confirm the previous research. Our evidence was limited in the right caudate; one reason for this might be the narrower distribution of the individual γs when compared to the surrogate value.

In the current study, we found that the activity of the lateral prefrontal areas correlated with $V(t)$. Some imaging studies reported a contribution of the lateral prefrontal cortex to the temporal discounting (McClure et al. 2004, 2007; Wittmann et al. 2007; Pine et al. 2009; Xu et al. 2009). A repetitive transcranial magnetic stimulation (rTMS) study reported that disruption of function of the lateral prefrontal cortex increased choice of immediate rewards over larger delayed rewards (Figner et al. 2010). There is strong evidence that the prefrontal cortex is involved in impulse control (Horn et al. 2003; Cardinal et al. 2004; Huettel et al. 2006), and dysfunction of the prefrontal cortex has been implicated in impulse control disorders (Evenden 1999). These results indicate that the lateral prefrontal cortex is a crucial neural substrate for self-control processes in intertemporal choice.

There are some limitations of the current study, the first of which is caused by the slice thickness used in this study. We set the thickness at 6 mm, in order to cover the whole brain and maintain high time-resolution. This thickness might have compromised the validity of our results. The striatum that we focused on is a relatively small region. To validate our results, we need to duplicate our findings using a thinner slice thickness. Another limitation is the lack of a significant relationship between the prefrontal cortex (including insula) and individual γ. The prefrontal cortex is anatomically connected to the striatum, and these regions form a loop circuit (Cummings 1993). To investigate not only local γ-related activation, but also that of the networks, a study combining model-based analysis and functional connectivity analysis, such as graph theory, as well as independent component analysis, would be useful. Lastly, we must consider the effects of the effort of the motor response and decision-making. In this study, motor-related activity was analytically regressed out. In decision-making, the effective choice behavior in this task was not changing the choice, but maintaining the same response during a trial. Therefore, it was considered that the cost of decision-making was not different between the small and large reward trials. Because of these factors, we speculate that effort did not have a critical impact on the current result.

In conclusion, our model-based regression analysis using individual time scales in our delayed reward choice task confirms the contribution of the caudate to reward prediction and suggests that better reward prediction might be related to the more dorsal part of activation in the right caudate. As our result was limited to the right caudate, the evidence does not completely support the notion that different cortico-striatum circuits are involved in different time scales. The combination of the reward prediction task with event-related fMRI and manipulation of neurotransmitters (dopamine or serotonin) might further clarify the neural mechanisms of reward prediction at different time scales.

References

Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. Psychol Bull 82:463–496

Ashburner J (2007) A fast diffeomorphic image registration algorithm. Neuroimage 38:95–113. doi:10.1016/j.neuroimage.2007.07.007

Ballard K, Knutson B (2009) Dissociable neural representations of future reward magnitude and delay during temporal discounting. Neuroimage 45:143–150. doi:10.1016/j.neuroimage.2008.11.004

Botvinick MM, Huffstetler S, McGuire JT (2009) Effort discounting in human nucleus accumbens. Cogn Affect Behav Neurosci 9:16–27. doi:10.3758/CABN.9.1.16

Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001) Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron 30:619–639

Bromberg-Martin ES, Hikosaka O (2009) Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron 63:119–126. doi:10.1016/j.neuron.2009.06.009

Cardinal RN, Winstanley CA, Robbins TW, Everitt BJ (2004) Limbic corticostriatal systems and delayed reinforcement. Ann NY Acad Sci 1021:33–50. doi:10.1196/annals.1308.004

Cummings JL (1993) Frontal-subcortical circuits and human behavior. Arch Neurol 50:873–880

Doya K (2002) Metalearning and neuromodulation. Neural Netw 15:495–506

Evenden J (1999) Impulsivity: a discussion of clinical and experimental findings. J Psychopharmacol 13:180–192

Figner B, Knoch D, Johnson EJ, Krosch AR, Lisanby SH, Fehr E, Weber EU (2010) Lateral prefrontal cortex and self-control in intertemporal choice. Nat Neurosci 13:538–539. doi:10.1038/nn.2516

Green L, Myerson J, McFadden E (1997) Rate of temporal discounting decreases with amount of reward. Mem Cognit 25:715–723

Gregorios-Pippas L, Tobler PN, Schultz W (2009) Short-term temporal discounting of reward value in human ventral striatum. J Neurophysiol 101:1507–1523. doi:10.1152/jn.90730.2008

Hariri AR, Brown SM, Williamson DE, Flory JD, de Wit H, Manuck SB (2006) Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. J Neurosci 26:13213–13217. doi:10.1523/JNEUROSCI.3446-06.2006

Horn NR, Dolan M, Elliott R, Deakin JF, Woodruff PW (2003) Response inhibition and impulsivity: an fMRI study. Neuropsychologia 41:1959–1966

Huettel SA, Stowe CJ, Gordon EM, Warner BT, Platt ML (2006) Neural signatures of economic preferences for risk and ambiguity. Neuron 49:765–775. doi:10.1016/j.neuron.2006.01.024

Johnson MW, Bickel WK (2002) Within-subject comparison of real and hypothetical money rewards in delay discounting. J Exp Anal Behav 77:129–146. doi:10.1901/jeab.2002.77-129

Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10:1625–1633. doi:10.1038/nn2007

Kalivas PW, Volkow ND (2005) The neural basis of addiction: a pathology of motivation and choice. Am J Psychiatry 162:1403–1413. doi:10.1176/appi.ajp.162.8.1403

Killeen PR (2009) An additive-utility model of delay discounting. Psychol Rev 116:602–619. doi:10.1037/a0016414

Knutson B, Cooper JC (2005) Functional magnetic resonance imaging of reward prediction. Curr Opin Neurol 18:411–417

Knutson B, Taylor J, Kaufman M, Peterson R, Glover G (2005) Distributed neural representation of expected value. J Neurosci 25:4806–4812. doi:10.1523/JNEUROSCI.0642-05.2005

Koopmans TC (1960) Stationary ordinal utility and impatience. Econometrica 28:287–309

Laibson D (1997) Golden eggs and hyperbolic discounting. Q J Econ 112:443–477

Lancaster K (1963) An axiomatic theory of consumer time preference. Int Econ Rev 4:221–231

Madden GJ, Bickel WK, Jacobs EA (1999) Discounting of delayed rewards in opioid-dependent outpatients: exponential or hyperbolic discounting functions? Exp Clin Psychopharmacol 7:284–293

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19:1233–1239

Marco-Pallares J, Mohammadi B, Samii A, Munte TF (2010) Brain activations reflect individual discount rates in intertemporal choice. Brain Res 1320:123–129. doi:10.1016/j.brainres.2010.01.025

Mazur JE (2001) Hyperbolic value addition and general models of animal choice. Psychol Rev 108:96–112

Mazur JE (2007) Rats' choices between one and two delayed reinforcers. Learn Behav 35:169–176

Mazur JE, Biondi DR (2009) Delay-amount tradeoffs in choices by pigeons and rats: hyperbolic versus exponential discounting. J Exp Anal Behav 91:197–211. doi:10.1901/jeab.2009.91-197

McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. Science 306:503–507. doi:10.1126/science.1100907

McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD (2007) Time discounting for primary rewards. J Neurosci 27:5796–5804. doi:10.1523/JNEUROSCI.4246-06.2007

McKerchar TL, Green L, Myerson J, Pickford TS, Hill JC, Stout SC (2009) A comparison of four models of delay discounting in humans. Behav Process 81:256–259. doi:10.1016/j.beproc.2008.12.017

Myerson J, Green L (1995) Discounting of delayed rewards: models of individual choice. J Exp Anal Behav 64:263–276

O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. Neuron 33:815–826

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. Neuron 38:329–337

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304:452–454. doi:10.1126/science.1094285

Phelps ES, Pollak RA (1968) On second-best national saving and game-equilibrium growth. Rev Econ Stud 35:185–199

Pine A, Seymour B, Roiser JP, Bossaerts P, Friston KJ, Curran HV, Dolan RJ (2009) Encoding of marginal utility across time in the human brain. J Neurosci 29:9575–9581. doi:10.1523/JNEUROSCI.1126-09.2009

Reynolds B, Schiffbauer R (2004) Measuring state changes in human delay discounting: an experiential discounting task. Behav Process 67:343–356. doi:10.1016/j.beproc.2004.06.003

Samuelson PA (1937) A note on measurement of utility. Rev Econ Stud 4:155–161

Schweighofer N, Shishida K, Han CE, Okamoto Y, Tanaka SC, Yamawaki S, Doya K (2006) Humans can adopt optimal discounting strategy under real-time constraints. PLoS Comput Biol 2:e152. doi:10.1371/journal.pcbi.0020152

Schweighofer N, Tanaka SC, Doya K (2007) Serotonin and the evaluation of future rewards: theory, experiments, and possible neural mechanisms. Ann N Y Acad Sci 1104:289–300. doi:10.1196/annals.1390.011

Schweighofer N, Bertin M, Shishida K, Okamoto Y, Tanaka SC, Yamawaki S, Doya K (2008) Low-serotonin levels increase delayed reward discounting in humans. J Neurosci 28:4528–4532. doi:10.1523/JNEUROSCI.4982-07.2008

Sopher B, Sheth A (2005) A deeper look at hyperbolic discounting. Theory Decis 60:219–255

Sutton RS, Barto AG (1998) Reinforcement learning. MIT press, Cambridge

Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially

recruits cortico-basal ganglia loops. Nat Neurosci 7:887–893. doi:10.1038/nn1279

Tanaka SC, Schweighofer N, Asahi S, Shishida K, Okamoto Y, Yamawaki S, Doya K (2007) Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. PLoS One 2:e1333. doi:10.1371/journal.pone.0001333

Wittmann M, Paulus MP (2008) Decision making, impulsivity and time perception. Trends Cogn Sci 12:7–12. doi:10.1016/j.tics.2007.10.004

Wittmann M, Leland DS, Paulus MP (2007) Time and decision making: differential contribution of the posterior insular cortex and the striatum during a delay discounting task. Exp Brain Res 179:643–653. doi:10.1007/s00221-006-0822-y

Xu L, Liang ZY, Wang K, Li S, Jiang T (2009) Neural mechanism of intertemporal choice: from discounting future gains to future losses. Brain Res 1261:65–74. doi:10.1016/j.brainres.2008.12.061