

Howard C. Cromwell · Oum K. Hassani
Wolfram Schultz

Relative reward processing in primate striatum

Abstract Rewards are often not only valued according to their physical characteristics but also relative to other available rewards. The striatum (caudate nucleus, putamen, ventral striatum including nucleus accumbens) is involved in the organization of movement and the processing of reward information. We studied the activity of single striatal neurons in macaques that were presented with different combinations of two rewards. We found in nearly half of the investigated neurons that the processing for one reward shifted, relative to the other rewards that were available in a given trial block. The relative reward processing concerned all forms of striatal activity related to reward-predicting visual stimuli, arm movements and reception of rewards. The observed changes may provide a neural basis for the known shifts in valuation of rewarding outcomes relative to known references.

Keywords Neurophysiology · Behavior · Caudate · Putamen · Ventral striatum

Introduction

Although rewards are objects with well defined physical properties, their motivational value can vary (Tinklepaugh 1928; Crespi 1942; Bevan 1968). For example, animals may choose a usually preferred reward less frequently, or work less energetically for it, when other, even more preferred rewards have been encountered (Reynolds 1961; Black 1968; Dunham

1968; Cox 1975; Flaherty 1996). The reward contrast may occur even when a more valued reward is observed with another animal (Brosnan and de Waal 2003). Thus rewards may influence each other, and the value of a reward can depend on other available rewards.

Besides playing a major role in movement processes, the striatum is also a key reward structure of the brain. Lesions in different parts of the striatum impair reward-seeking behavior, adaptation to changed rewards, and valuation of one reward relative to others (Flaherty 1996; Salinas and White 1998; Bowman and Brown 1998; Leszczuk and Flaherty 2000). Single striatal neurons show sustained activations during the expectation of reward (Hollerman et al. 1998; Kawagoe et al. 1998) and discriminate between different rewards or reward magnitudes (Bowman et al. 1996; Hassani et al. 2001; Cromwell and Schultz 2003). However, it is unknown whether individual striatal neurons process rewards according to their physical characteristics or relative to other rewards.

This work extends our previous studies on simple reward discrimination in the anterior striatum (Hassani et al. 2001; Cromwell and Schultz 2003) and examines how the neural processing of one reward may depend on other rewards. We used a spatial delayed response task in which specific, reward-predicting instructions signaled to the animal which of two rewards would be delivered for correct performance in each trial. Different reward combinations served to study the same reward in relation to different alternatives.

Methods

We employed the same three monkeys and similar methods as in two previous studies (Hassani et al. 2001; Cromwell and Schultz 2003) (A, B: *Macaca fascicularis*, female 3.4 kg; male 2.8 kg; C: *Macaca mulatta*, male 6.0 kg). All experiments conformed to Swiss Animal Protection Law and the NIH guiding principles.

W. Schultz (✉)
Department of Anatomy, University of Cambridge,
Downing Street, Cambridge, CB2 3DY, UK
E-mail: ws234@cam.ac.uk

H. C. Cromwell · O. K. Hassani
Institute of Physiology, University of Fribourg,
1700 Fribourg, Switzerland

http://doc.rero.ch

Behavioral task

A visual instruction (13°×13°) appeared for 1 s on a 13-inch computer monitor at a left or right position to indicate a target lever below the stimulus for the operant arm movement (Fig. 1a). The shape and color of the instruction predicted which of two rewards would be delivered for correct target acquisition. At 3.5–4.5 s after instruction onset a uniform twin trigger stimulus appeared on the monitor. The moderately fluid-deprived animal released the resting key, touched the lever previously indicated by the instruction, and received the reward predicted by the instruction at 2.0 s after lever touch through a spout at its mouth. Error trials were unrewarded and repeated. Animals received no further conditioning for associating instructions with rewards.

The left and right targets and the two rewards varied pseudorandomly between trials, with a maximum of three identical consecutive trials. We presented 7–15 trials for each of two targets and two rewards (total of

28–60 trials per block), a trial lasted 7–9 s, and intertrial intervals were 9–11 s.

Behavioral reward preferences were assessed in blocks of choice trials in which the two instructions were presented simultaneously. Each instruction pair contained one picture associated with a preferred, or larger, reward and one with a nonpreferred, or smaller, reward. Stimuli inducing visual preferences were discarded. The animal chose the reward after the trigger stimulus by touching the lever previously indicated by the instruction associated with the desired reward. Neural data from choice trials are not presented in this report.

We used different juice rewards with the same quantity (0.18 mL; animals A, C) or one juice reward with different quantities (0.12, 0.18, 0.24 mL; animals A, B). Rewards were selected in choice trials for consistent and satiety-insensitive preference. Rewards were blackcurrant or raspberry (most preferred), orange (intermediately preferred), and grenadine or lemon juices (least preferred; animal A), or raspberry (most preferred),

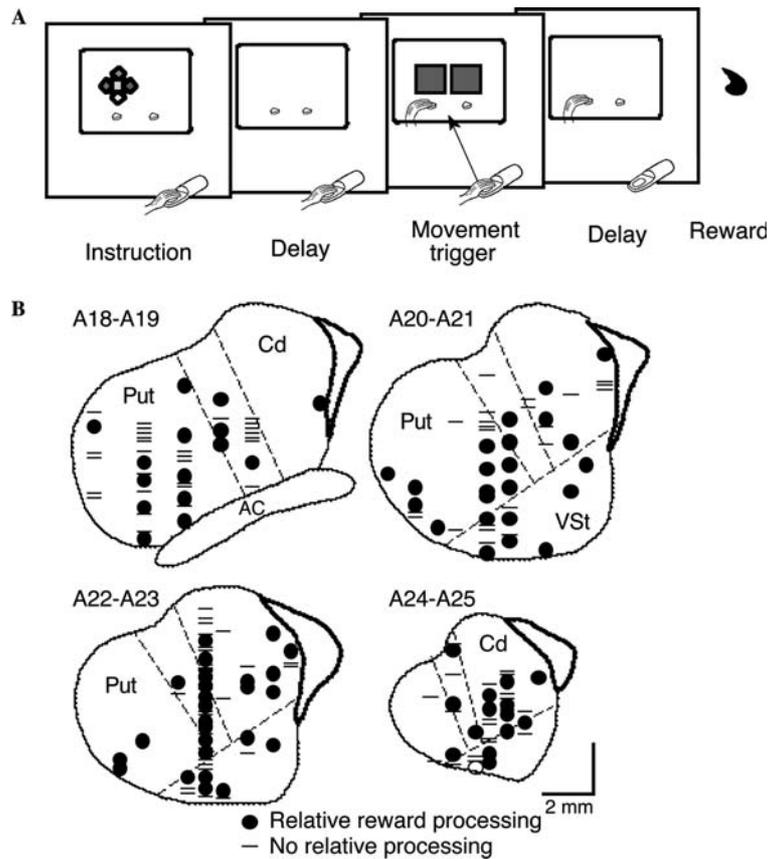


Fig. 1 **A** Spatial delayed response task. The position of the colored instruction picture indicated the left or right movement target. The visual aspect of the instruction predicted the kind or amount of liquid reward. Following a brief delay after the instruction, two identical red squares appeared and triggered the arm movement from the resting key to the left or right target lever previously indicated by the instruction. Correct performance was rewarded after a brief delay with a drop of liquid. Two different kinds or amounts of reward, each predicted by a different instruction picture, alternated pseudorandomly in a given block of trials, as did

the left and right movement targets. **B** Positions of striatal neurons showing relative reward processing (*dots*). Data were plotted superimposed from the three monkeys on coronal histological sections at approximate antero-posterior levels (A18–25 denote mm rostral to earbars). *Interrupted lines* denote approximate boundaries of ventral striatum (Haber et al. 1995) and internal capsule. Neurons located in the internal capsule were recorded from cell bodies, not fibers, as inferred from their long impulses (>0.8 ms). (*Put* putamen, *Cd* caudate, *VSt* ventral striatum, *AC* anterior commissure)

blackcurrant (intermediately preferred), and orange or lemon juices (least preferred; animal C). For studying reward quantity we used raspberry (animal A), blackcurrant (animals A and B) and grenadine juices (animal B).

Blocks of choice trials assessed reward preferences at least once on each neural recording day. Animals showed stable reward preferences over several weeks in all reward comparisons. They chose the same, preferred kind of reward in $92.8 \pm 1.1\%$ of trials (mean \pm SE of the mean, SEM; 93 choice blocks of 28–60 trials/block, range 56–100%). When reward magnitude was varied, animals selected the larger reward in $>95\%$ of trials.

Data acquisition and evaluation

After behavioral training, the activity of single, slowly discharging neurons in the left (animals A, B, C) and right striatum (animal A) was recorded extracellularly with moveable microelectrodes using standard electrophysiological techniques while monitoring licking movements. Tonicly active striatal neurons (TAN) were not studied. Recording sites were reconstructed from small electrolytic lesions on 50 μ m thick, cresyl violet-stained, coronal brain sections.

We evaluated relative reward processing in single neurons in three steps. First, we assessed task-related activations for a given reward by comparing impulse rates in single trials during postinstruction (≤ 1 s duration), pretrigger, posttrigger, prereward or postreward periods against 1–2 s preinstruction control periods ($P < 0.01$, 1-tailed Wilcoxon test). We rejected $\sim 5\%$ of neurons with nonstationary activity within or across identical trial blocks ($P < 0.01$; 2-tailed Mann–Whitney U test, M–W). Second, we assessed reward discrimination within a trial block by comparing task-related activity in identical periods between two rewards (M–W; $P < 0.01$). Third, we assessed relative reward processing across trial blocks by comparing activations for the same reward between different combinations with other rewards ($P < 0.01$; M–W).

Results

We investigated each neuron with three different rewards or reward magnitudes (A, B and C) but delivered only two of them in each trial block (A–B, B–C or A–C) and only one reward on each trial. The block design allowed the animal to predict after a few trials the particular combination of two rewards that would pseudorandomly alternate between trials. Animals discriminated between rewards by showing shorter arm movement reaction times (276–465 ms), longer anticipatory licking durations (850–2,900 ms) and fewer behavioral errors (3–14%) for preferred versus less preferred kinds or magnitudes of reward (250–2,980 trials per comparison, three animals; $P < 0.0005$ – 0.025

in 10 of 16 comparisons; 1-tailed t test or Wilcoxon test).

Basic relative reward processing

To test relative reward processing, we compared neural activity for the same reward between two different reward combinations. We investigated 185 striatal neurons, which were a subset of 984 slowly discharging striatal neurons tested for simple reward discrimination (Hassani et al. 2001; Cromwell and Schultz 2003). Of the 185 neurons, 79 (43%) showed significantly different, task-related activity for the same reward between two different reward combinations ($P < 0.01$; M–W). Behavioral reaction times, reward-anticipating licking durations and error rates varied for the same reward when tested in two different reward combinations, although the differences reached significance only in 12 of the 79 pairs of trial blocks (overall $P > 0.25$ in all three measures). Relative reward processing occurred in 27 neurons showing higher activity for preferred rewards, 26 neurons showing higher activity for nonpreferred rewards, eight neurons showing highest activity for a particular reward object or magnitude, and 18 neurons that failed to discriminate between any of the rewards used. The 79 neurons were recorded from monkeys A ($n = 33$ neurons), B ($n = 4$) and C ($n = 42$) in the “associative” striatum rostral to the anterior commissure (caudate, $n = 20$; putamen, $n = 40$; ventral striatum, $n = 19$) (Fig. 1b).

An example of relative reward processing is shown in Fig. 2a. The neuron responded more to grenadine than orange or blackcurrant juice. However, the response to the identical grenadine juice was higher when blackcurrant juice was the alternative reward (bottom), compared to orange as alternative (top) ($P < 0.01$; M–W). Similar differences were seen in the 79 neurons in any of the five task periods (postinstruction, pretrigger, posttrigger, prereward or postreward; Table 1). They amounted to an overall median increase in activation of 185% for the more effective above the less effective reward combination ($n = 79$; Fig. 2c). Linear regression analysis across individual neurons failed to reveal relationships between differences in neural activation and reaction times ($R^2 < 0.1$), including premovement and movement epochs ($R^2 = 0.064$), suggesting that simple movement differences were not responsible for the observed neural changes. Thus some striatal neurons do not process reward information in a fixed manner but relative to other available rewards.

Shift of reward processing

We investigated further the nature of relative reward processing in 20 of the 79 neurons that discriminated between all rewards and labeled the rewards according to the animal’s preferences as $A < B < C$. We presented the same,

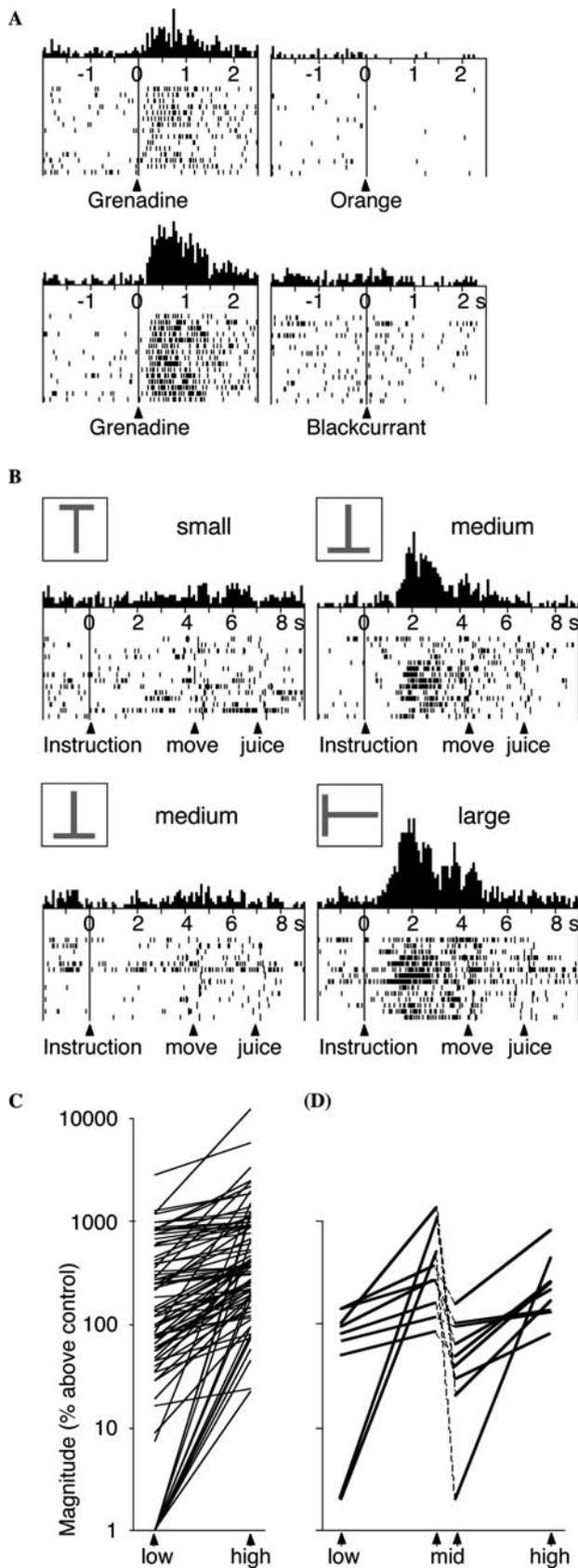


Fig. 2 Relative reward processing in striatal neurons. **A** Change of response to identical grenadine juice depending on the alternative reward. This putamen neuron discriminated between grenadine and orange (*top*) and between grenadine and blackcurrant juices (*bottom*). However, the grenadine response was lower with orange juice as alternative (*top*), compared with blackcurrant as alternative in a different trial block (*bottom*). During this neural recording, the animal preferred blackcurrant > orange > grenadine juice. The neuron showed higher activity for less preferred rewards. Prereward background activity was higher in blackcurrant compared with orange trials but was similar in grenadine trials across trial blocks, arguing against a change in general neural excitability. Chronology of trials is from top to bottom in each raster. Perivient time histograms of neural impulses are shown above rasters in which each dot denotes the time of a neural impulse relative to reward onset. Each line shows one trial. Trials alternated pseudorandomly between two rewards and were separated for analysis. Top and bottom trials were run in separate blocks. **B** Shift of reward coding in a caudate neuron (variation of reward magnitude). Activation with medium reward depended on the magnitude of the other reward within the same block (*top vs. bottom*). Thus, neural activity was higher for whatever reward was relatively larger. Chronology of trials is from top to bottom in each raster. Reward was blackcurrant juice. Reward amount was indicated at trial onset by the instruction (Ts at different orientations). **C** Quantitation of relative reward processing in the 79 neurons. *Low* and *high* refer to the lower and higher activations for a given reward tested in two trial blocks against different kinds or amounts of reward, irrespective of activation increases or decreases with more preferred reward. Activation magnitudes are expressed in % above preinstruction control activity and comprise all types of task-related activation (postinstruction, pretrigger, posttrigger, prereward, postreward). Medians were 137% for “low” and 391% for “high” activations (increase of 185%). **D** Shift of reward processing according to currently available rewards in nine neurons showing increases with rewards of increasing preference (reward kind or magnitude). *Interrupted lines* connect magnitudes of activation for the same intermediately preferred reward (*mid*) tested in combination with reward of “low” (*left*) or “high” preference (*right*). (Note that “low” and “high” refer to behavioral preference in **D** and to neural activation in **C**)

intermediately preferred reward **B** in two trial blocks in which it was either the relatively more (**A–B** block) or less preferred reward (**B–C** block). Behavioral reaction times with reward **B** were significantly shorter in **A–B** compared to **B–C** blocks, indicating a behavioral contrast, although the differences reached significance only during 10 of the 20 neural recordings ($P < 0.1$; M–W). The neuron of Fig. 2b showed stronger instruction-trigger delay activity for the same intermediate reward **B** in **A–B** (*top*) compared to **B–C** blocks (*bottom*). Within both blocks, activations were strongest with whatever was the larger of two rewards, irrespective of actual reward size. Of the 20 neurons, 12 showed similar shifts in any of the five task periods (median 268% activation increase with reward **B**; $P < 0.01$; M–W). The change occurred irrespective of kind ($n = 7$ neurons) or amount of reward ($n = 5$), and irrespective of activation increases ($n = 9$; Fig. 2d) or decreases ($n = 3$) with the more preferred reward. Thus neural reward coding in some striatal neurons can shift depending on the currently available rewards.

Table 1 Numbers of striatal neurons with relative reward processing

Task relationship	Reward kind	Reward amount	Total neurons with relative processing	Total task-related neurons tested
Response to instruction	4/7	0/3	4 (40%)	10
Activation during instruction-movement delay	19/52	8 ^a /15	27 (40%)	67
Activation following trigger	13/23	0/3	13 (50%)	26
Activation preceding reward	8/13	2/8	10 (48%)	21
Response to reward	22/48	3/13	25 (41%)	61
Sum	66/143	13/42	79 (43%)	185

A total of 185 task-related neurons was tested for relative reward processing. The 185 neurons showed one of the five forms of task relationship indicated on the left. For reasons of simplicity, only the strongest activation was analyzed when individual neurons showed multiple activations, irrespective of relative reward coding. The two left numeric columns contain the numbers of neurons showing relative reward processing/numbers of task-related neurons tested

Measurement periods for statistical analysis were ≤ 1.0 s after the instruction, ≤ 3.5 s during the instruction-movement delay, ≤ 1.0 s following the trigger stimulus and during the movement, ≤ 2.0 s preceding the reward, ≤ 1.0 s after the reward

^aOf these eight neurons, three showed significantly different activation for movements to the left versus right, the remaining five had nonspatial task-related activations

Discussion

These data show that some neurons in the anterior striatum were not only influenced by the reward in the current trial but also by rewards presented in other trials within the same block. Whereas our previous studies concerned the neural discrimination of rewards (Hassani et al. 2001; Cromwell and Schultz 2003), the current results suggest rapid shifts in reward processing through which neurons would adapt the dynamic range of responding to the currently available rewards. This mechanism would enable optimum reward discrimination at the instant at which an animal makes a decision concerning rewards.

Although the neural changes concerned all forms of task-related activation, they were only accompanied by subtle behavioral effects. Behavioral changes were insignificant when testing the basic relative processing and reached significance in only half the tests for shifted reward processing. These comparisons suggest that the relative reward processing may represent a basic neural mechanism operating even without overt behavioral effects.

The shift in reward processing may reflect an adaptation that could be expressed, in the most simple form, as a change in the linear input-output function of $y = a + b(x)$ to $y = a + b(x - p)$, with p as offset of processing induced by prediction about the current reward distribution (a is a constant, b is the slope) (Fig. 2d). Thus reward sensitivity may adapt rapidly to changes in the distribution of rewards, although the details of the adaptation remain to be investigated.

Similar relative, rather than absolute, processing of reward seems to occur also in other brain structures. Orbitofrontal neurons adapt their discrimination to the current rewards (Tremblay and Schultz 1999), although they rarely show the goal-related (Dickinson and Balentine 1994) movement preparatory delay activity of striatal neurons (Hollerman et al. 1998) (Fig. 2b). Dorsolateral prefrontal neurons show activations in unre-

warded trials that depend on the available reward in rewarded trials (Watanabe et al. 2002). Posterior parietal neurons show rapidly adapting, constant responses to preferred rewards irrespective of absolute value in a game task in which animals by definition chose outcomes with the highest utility (Dorris and Glimcher 2004). Taken together, relative reward processing involving response adaptation seems to constitute a rather general reward mechanism.

In a general sense, relative reward processing in the striatum may contribute to the shift in reward valuation relative to variable references in a number of behavioral situations conceptualized in learning theory (Flaherty 1996) and microeconomic decision theory (Kahneman and Tversky 1984). The reduction of reward contrast in instrumental responses after ventral striatal lesions in rats (Leszczuk and Flaherty 2000) may suggest that the relative reward effect observed in this work provides a neural mechanism for valuation of predictable outcomes relative to other available outcomes and contributes to the efficacy of goal-directed behavior.

Acknowledgements We thank B. Aebischer, J. Corpataux, A. Gaillard, B. Morandi and F. Tinguely for expert technical assistance. The study was supported by the Wellcome Trust, Swiss NSF, EU Human Capital and Mobility and Biomed 2 programs via the Swiss Office of Education and Science, and US NSF (HCC).

References

- Bevan W (1968) The contextual basis of behavior. *Am Psychol* 23:701–714
- Black RW (1968) Shifts in magnitude of reward and contrast effects in instrumental and selective learning: a reinterpretation. *Psychol Rev* 75:114–126
- Bowman EM, Brown VJ (1998) Effects of excitotoxic lesions of the rat ventral striatum on the perception of reward cost. *Exp Brain Res* 123:439–448
- Bowman EM, Aigner TG, Richmond BJ (1996) Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J Neurophysiol* 75:1061–1073
- Brosnan SF, de Waal FBM (2003) Monkeys reject unequal pay. *Nature* 425:297–299

- Cox WM (1975) A review of recent incentive contrast studies involving discrete-trial procedures. *Psychol Rec* 25:373–393
- Crespi LP (1942) Quantitative variations of incentive and performance in the white rat. *Am J Psychol* 55:467–517
- Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol* 89:2823–2838
- Dickinson A, Balleine B (1994) Motivational control of goal-directed action. *Anim Learn Behav* 22:1–18
- Dorris MC, Glimcher PW (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44:365–378
- Dunham PJ (1968) Contrasted conditions of reinforcement: a selective critique. *Psychol Bull* 69:295–315
- Flaherty CF (1996) *Incentive relativity*. Cambridge University Press, Cambridge
- Haber S, Kunishio K, Mizobuchi M, Lynd-Balta E (1995) The orbital and medial prefrontal circuit through the primate basal ganglia. *J Neurosci* 15:4851–4867
- Hassani OK, Cromwell HC, Schultz W (2001) Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J Neurophysiol* 85:2477–2489
- Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related activity in primate striatum. *J Neurophysiol* 80:947–963
- Kahneman D, Tversky A (1984) Choices, values, and frames. *Am Psychol* 4:341–350
- Kawagoe R, Takikawa Y, Hikosaka O (1998) Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* 1:411–416
- Leszczuk MH, Flaherty CF (2000) Lesions of nucleus accumbens reduce instrumental but not consummatory negative contrast in rats. *Behav Brain Res* 116:61–79
- Reynolds GS (1961) Behavioral contrast. *J Exp Anal Behav* 4:57–71
- Salinas J, White NM (1998) Contributions of the hippocampus, amygdala and dorsal striatum to the response elicited by reward reduction. *Behav Neurosci* 112:812–826
- Tinklepaugh OL (1928) An experimental study of representation factors in monkeys. *J Comp Psychol* 8:197–236
- Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708
- Watanabe M, Hikosaka K, Sakagami M, Shirakawa SI (2002) Coding and monitoring of behavioral context in the primate prefrontal cortex. *J Neurosci* 22:2391–2400