

# Reward Prediction Error Computation in the Pedunculopontine Tegmental Nucleus Neurons

Yasushi Kobayashi and Ken-ichi Okada

*Osaka University, ATR Computational Neuroscience Laboratories & PRESTO JST  
Japan*

## 1. Introduction

Reinforcement learning (Houk et al., 1995; Sutton & Barto, 1998; Schultz, 2002) has been one of the central topics in a broad range of scientific fields for the last two decades. Understanding of reinforcement learning is expected to provide a systematic understanding of adaptive behaviors, including simple classical and operant conditioning of animals (Waelti et al., 2001; Richmond et al., 2003; Satoh et al., 2003; Graybiel, 2005; Samejima et al., 2005; Hikosaka et al., 2006) as well as all complex social and economical human behaviors that are designed to maximize benefits (Montague & Berns, 2002); and is also useful in machine learning and robotics (Tesauro, 1994).

Reinforcement learning, whether performed by living organisms or computational models, involves choosing a behavior that is expected to yield the maximal reward and then revising this prediction so as to minimize the reward prediction error (Schultz, 2002), which is the difference between the predicted and actual reward.

Recent neurophysiological studies have shown that midbrain dopamine neurons encode the reward prediction error signal (Schultz et al., 1997; Hollerman & Schultz, 1998; Waelti et al., 2001; Fiorillo et al., 2003; Nakahara et al., 2004; Morris et al., 2006) and that the striatum (Hollerman & Schultz, 1998; Hikosaka et al., 2006) and cerebral cortices (Watanabe, 1996; Lee & Seo, 2007) use this signal to perform reinforcement learning with dopamine-induced synaptic plasticity (Reynolds et al., 2001; Wickens et al., 2003). Thus, computing the reward prediction error is one of the most essential aspects of reinforcement learning, however the identity of the neural structures that provide the signals to the midbrain dopamine neurons and the mechanism by which the 'reward prediction error' is computed remain rather elusive. The pedunculopontine tegmental nucleus (PPTN) of the midbrain feeds strong excitatory inputs to dopamine neurons in the midbrain, and receives reward-related signals from various areas including the cerebral cortices and the striatum. We hypothesize that the PPTN is the key structure for computing the reward prediction error. To test this hypothesis, we recorded the activity of PPTN neurons in monkeys performing a saccade task for a juice reward (Kobayashi et al., 2002; Okada et al., 2009).

In the most recent study (Okada et al., 2009), we used multiple analytical approaches, including receiver operating characteristic (ROC) analysis (Lusted, 1978), mutual information (Werner & Mountcastle, 1963; Schreiner et al., 1978; Kitazawa et al., 1998), and correlation analyses to examine neuronal responses in the PPTN neurons in monkeys performing saccade tasks, during which the magnitudes of rewards were predicted in

normal and reversed fashions. All analyses consistently indicated the existence of two neuronal groups, one signalling the expected reward magnitude predicted from the visual stimulus and the other signalling the magnitude of the actual reward, both necessary and sufficient pieces of information for computing the reward prediction error. The reward prediction error may be directly computed by subtracting the signals encoded by the two PPTN neuronal groups, or alternatively, by adding the time derivatives of the reward prediction signals to the actual reward signals, as originally hypothesized by the temporal difference reinforcement learning model. Thus, we concluded that the PPTN is indeed a key structure for computing of the reward prediction error.

## 2. Background

### 2.1 Classical view of the PPTN

The cholinergic system is one of the most important modulatory neurotransmitter systems in the brain, and controls neuronal activity that depends on selective attention. Anatomical and physiological evidence supports the idea of a 'cholinergic component' of conscious awareness (Perry et al., 1999). The PPTN in the brainstem contains both cholinergic (Mesulam et al., 1983) and non-cholinergic neurons (Jones & Beaudet, 1987; Clements & Grant, 1990; Spann & Grofova, 1992; Ford et al., 1995; Takakusaki et al., 1996; Wang & Morales, 2009), but is one of the major sources of cholinergic projections in the brainstem (Mesulam et al., 1983). The PPTN is thought to be the central part of the reticular activating system (Garcia-Rill, 1991), which provides background excitation for several sensory and motor systems essential for automatic control of movement (Takakusaki et al., 2004), perception and cognitive processes (Steckler et al., 1994). It has long been known that the PPTN is a crucial element in the regulation of the rhythms in the cortex (Steriade et al., 1990) that are associated with wakefulness and rapid eye movement sleep (Leonard & Llinas, 1994).

Anatomically, the PPTN has reciprocal connections with the basal ganglia: the subthalamic nucleus, the globus pallidus, and the substantia nigra (Edley & Graybiel, 1983; Lavoie & Parent, 1994), and more recently, was argued to form a part of the basal ganglia (Mena-Segovia et al., 2004). Further, the PPTN also has reciprocal connections with catecholaminergic systems in the brainstem: the locus coeruleus (noradrenergic) (Garcia-Rill, 1991; Garcia-Rill et al., 1995) and the dorsal raphe nucleus (serotonergic) (Steininger et al., 1992; Honda & Semba, 1994; Kayama & Koyama, 2003). This basal ganglia-PPTN-catecholaminergic complex was proposed to play an important role in gating movement, controlling several forms of attentional behavior (Garcia-Rill, 1991) and the reinforcement process (Doya, 2002). Despite these abundant anatomical findings, however, the functional importance of the PPTN is not yet fully understood.

### 2.2 The possible role of the PPTN in reinforcement process

Several of lesion and drug administration studies on rodents indicate that the PPTN is involved in various reinforcement processes (Bechara & van der Kooy, 1989; Kippin & van der Kooy, 2003; Alderson et al., 2006; Winn, 2006; Wilson et al., 2009). According to a physiological study in operantly conditioned cats, the PPTN relays either a reward or a salient event signal (Dormont et al., 1998). Anatomically, the PPTN receives reward input from the lateral hypothalamus (Semba & Fibiger, 1992) and the limbic cortex (Chiba et al., 2001). Conversely, the PPTN abundantly projects to midbrain dopamine neurons of the

substantia nigra pars compacta and ventral tegmental area (Beckstead et al., 1979; Jackson & Crossman, 1983; Beninato & Spencer, 1987; Charara et al., 1996), which encode a reward prediction error signal for reinforcement learning (Schultz, 1998).

The PPTN is one of the strongest excitatory input sources for the dopamine neurons (Matsumura, 2005). These afferent PPTN neurons release glutamate and acetylcholine to target neurons, make glutamatergic and cholinergic synaptic connections with dopamine neurons in the midbrain (Scarnati et al., 1986; Futami et al., 1995; Takakusaki et al., 1996). In the rat, electrical stimulation of the PPTN induces a time-locked burst in dopamine neurons (Lokwan et al., 1999; Floresco et al., 2003), and chemical or electrical stimulation of the PPTN increases dopamine release in the striatum (Chapman et al., 1997; Forster & Blaha, 2003; Miller & Blaha, 2004). Other electrophysiological experiments have shown that acetylcholine acts through both nicotinic and muscarinic receptors to depolarize dopamine neurons and to alter their firing pattern (Calabresi et al., 1989; Lacey et al., 1990; Gronier & Rasmussen, 1998; Sorenson et al., 1998). Thus, PPTN activity and acetylcholine provided by the PPTN can facilitate the burst firing in dopamine neuron (Mena-Segovia et al., 2008) and appear to do so via muscarinic (Kitai et al., 1999; Scroggs et al., 2001) and nicotinic (Grenhoff et al., 1986; Pidoplichko et al., 1997; Sorenson et al., 1998; Yamashita & Isa, 2003) acetylcholine receptor activation. In addition, some of the effects induced by PPTN stimulation can be blocked by administration of the muscarinic acetylcholine receptor agonist carbachol into the PPTN (Chapman et al., 1997). This finding is consistent with the fact that cholinergic neurons in the PPTN express the inhibitory muscarinic autoreceptors (Yeomans, 1995) and suggests that activation of these receptors inhibits cholinergic inputs to the dopamine neurons (Tzavara et al., 2004; Chen et al., 2006).

Furthermore, midbrain dopamine neurons are dysfunctional following excitotoxic lesioning of the PPTN (Blaha & Winn, 1993). A number of studies have found impairments in learning following excitotoxic lesions of the PPTN (Fujimoto et al., 1989; Fujimoto et al., 1992; Steckler et al., 1994; Inglis et al., 2000; Alderson et al., 2002). Thus, abundant anatomical, electrophysiological and pharmacological studies of slice and whole animal preparations indicate that the PPTN receives signals from the reward related structures including the cerebral cortices and the striatum (Winn et al., 1997) and provides strong excitatory inputs to the dopamine neurons (Clements & Grant, 1990; Blaha & Winn, 1993; Futami et al., 1995; Oakman et al., 1995; Blaha et al., 1996; Conde et al., 1998; Dormont et al., 1998; Mena-Segovia et al., 2004; Pan & Hyland, 2005; Mena-Segovia et al., 2008). Interestingly, the dopamine/ acetylcholine interaction seems to be mutual (Scarnati et al., 1987); dopamine neurons in the substantia nigra pars compacta project back to PPTN neurons, affecting their excitability. Even though the dopaminergic input to the PPTN is low compared with the massive cholinergic innervation of the dopamine neurons (Semba & Fibiger, 1992; Grofova & Zhou, 1998; Ichinohe et al., 2000), dopamine released within the PPTN may play an important part in controlling its activity (Steiniger & Kretschmer, 2003).

Therefore, it is plausible that the PPTN provides important information for computing reward prediction error by the dopamine neurons. Recent studies (Matsumura et al., 1997; Pan & Hyland, 2005) reported that the PPTN encodes sensory or motor rather than reward information of task events. However, using a visually guided saccade task requiring the animal to shift its gaze from a fixation to a saccade target, we demonstrated the existence of two groups of neurons within the PPTN, one whose responses to presentation of the fixation target to initiate the task were correlated with the success and failure of individual task trials, and another that was responsive to the reward delivery (Kobayashi et al., 2002).

We hypothesized that the task performance-related neurons signal the reward prediction and the reward delivery-related neurons signal the reward outcome. This hypothesis was tested in monkeys by studying the activity of PPTN neurons during visually guided saccade tasks that were rewarded with different amounts of juice that were cued by the shape of the fixation target (Okada et al., 2009).

### 3. Responses of PPTN neurons to different reward magnitude

#### 3.1 Effect of reward prediction on behavior and neuronal activity of PPTN

In this study, Japanese monkeys were trained on a visually guided saccade task that required them to maintain fixation on a central fixation target, and to make a horizontal saccade to a peripheral saccade target that was presented after the disappearance of the fixation target (Fig. 1A). Correct trials were rewarded randomly with either one or three

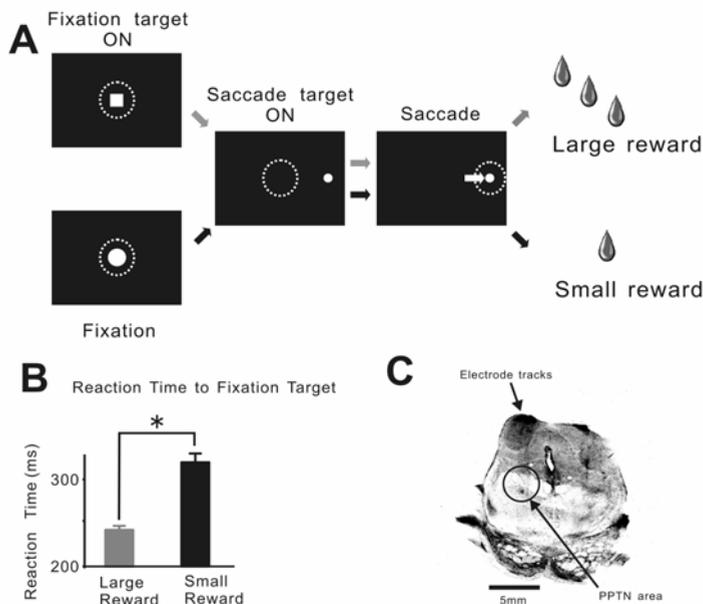


Fig. 1. Two-valued reward, visually guided saccade task.

**A.** Schematic of screen views for the two-valued visually guided saccade task. A fixation target (*square or circle*) was presented for 400-1000 ms. A saccade target was presented to the left or the right of the fixation target (eccentricity, 10°) 300-500ms after fixation target offset. The monkey was required to maintain fixation on the fixation target during the entire time it was presented, and to then make a saccade to the saccade target within 500 ms after saccade target onset. They were rewarded for successful trials with either one or three drop of juice in a quasirandom fashion.

**B.** Mean reaction times on the the fixation target. Error bars = *SEM*, \* indicates  $p < 0.001$  (Student's t-test).

**C.** Photomicrograph of a coronal section through midbrain of one monkey showing electrode tracks and the lesion (within the circle) marking the recording site in the PPTN.

Figures were modified from our recent paper (Okada et al., 2009).

drops of juice; the amount (large or small) being cued at the outset by the shape of the initial central fixation target (square or circle, respectively).

The behavior of the monkeys was influenced by the reward value expectation, the percentage of successful trial being significantly higher for large rewards than for small ones. In the unsuccessful trials, three types of errors occurred: monkeys failed to fixate on the fixation target (fixation error), they failed to maintain fixation until the appearance of the saccade target (fixation hold error), and they failed to make a saccade towards the saccade target (saccade error). The reaction time to fixate on the the fixation target was significantly shorter in the successful than in the unsuccessful ones. There was also a systematic difference in the reaction time within the successful trials: those associated with large rewards were significantly shorter than those for small rewards (Fig. 1B).

One hundred fifty-three PPTN neurons (see, recording sites in Fig. 1C) exhibited significant responses to one or more task events. Of these, 30 neurons exhibited increased firing around the time of the onset of the fixation target, with significant dependency on the magnitude of the predicted reward (*fixation target* neurons), and 15 neurons exhibited increased firing only around the time of the reward delivery with significant dependency on the reward magnitude of the current reward (*reward delivery* neurons).

Figures 2A, B show raster displays and spike density functions for a representative fixation target neuron. This neuron showed elevated firing throughout the trial that was greater when the cued reward was large: compare the red raster lines and traces (large reward) with the green (small rewards). The population plot for the 30 fixation target neurons (Fig. 2C) indicates that the differences in responses to the large and small reward cues generally began to emerge about 100 ms after the cue was presented (the 1<sup>st</sup> dotted line), even though there were non-differential responses before the onset of the fixation target/ cue, presumably in anticipation of its appearance. Note that the differential responses extended throughout the working memory period following offset of the fixation target/ cue and lasted until and even after reward delivery (3<sup>rd</sup> dotted line), almost unaffected by other task events, such as the onset of the peripheral saccade target (black bars in Fig. 2A, 2<sup>nd</sup> dotted line in Fig. 2C) and the saccade to the saccade target (inverted triangles in Fig. 2A).

In contrast, reward delivery neurons were almost unresponsive jst before the reward was delivered, when they discharged transiently, reaching a peak discharge rate shortly after reward delivery and then rapidly declining back to baseline (Figs. 2E, F). In trial with larger rewards, the discharge rate of the transient response reached a higher peak at a slightly later time and took a few hundred milliseconds longer to decay back to baseline than did that during small reward trials.

The clear suggestion here is that the differential dependencies of the fixation target and reward delivery neurons in encode the magnitudes of the predicted and current rewards, respectively. Further, we analyzed the precision for neuronal activity to encode the reward magnitude in two ways; 1) by ROC analysis for discrimination between the small and large rewards; 2) by mutual information analysis to estimate the information contained in the spike discharges with respect to the magnitude of the reward (Werner & Mountcastle, 1963; Schreiner et al., 1978; Kitazawa et al., 1998) where. These two analyses were conducted using a sliding time window of 200 ms moved in 1 ms steps.

First, the reliability with which the activity of individual neurons encoded large or small reward was estimated by deriving an ROC value (cumulative probability of the ROC curve) that measures the accuracy by which an ideal observer could correctly distinguish between large and small reward from the neuronal signal:

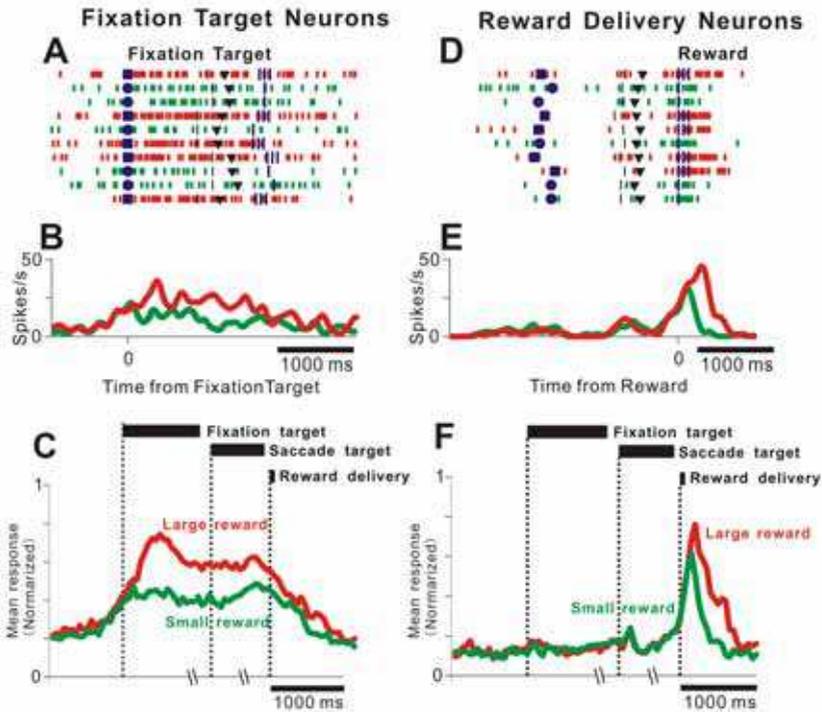


Fig. 2. Responses of the fixation target and the reward delivery neurons to task events. **A, B**, a rastergram and peri-task event spike density function for activities of a representative fixation target neuron over 10 successive trials, aligned to the onset of the fixation target. Red and green rasters (**A**) and traces (**B**) indicate large and small reward trials, respectively. In (**A**) blue squares and circles indicate fixation target onset, black bars onset of the saccade target, blue triangles saccade onset and the blue lines the times at which large (three bars) and small (one bar) rewards were delivered. **C**, The population spike density function for the 30 fixation target neurons. Responses are aligned to fixation target and saccade target onsets and the moment of reward delivery (vertical dotted lines). Large and small reward trials are indicated once again with red and green, respectively, as above, and thick horizontal bars above indicate the durations of the respective events. **D-F**, a similar rastergram (**D**) and response histogram (**E**) for a representative reward delivery neuron and the population response histograms (**F**) for the 15 reward delivery neurons. Formats are the same as in **A-C**. Figures were modified from our recent paper (Okada et al., 2009).

$$ROC = \int_0^1 P(Q) dQ \tag{1}$$

$$P(x) = \int_0^x p(x) dx \tag{2}$$

$$Q(x) = \int_0^x q(x) dx \tag{3}$$

$x$  denotes the neuronal activity sampled through the moving window.  $p(x)$  and  $q(x)$  denote the probability distributions for a ideal observer to answer whether the reward is large or small, respectively;  $P(x)$  and  $Q(x)$  denote the cumulative probability of these functions.  $P(Q)$  represents an ROC curve, and the ROC value is the area under the ROC curve evaluated as  $\int_0^1 P(Q)dQ$ , and  $Q$  is the cumulative probability function for small reward trials that was taken as the reference distribution.

In principle, ROC analysis evaluates the reliability with which an ideal observer can tell whether the reward is large or small from the noisy signal in terms of statistical significance of the signal difference between the two rewards in comparison with the baseline noise. Therefore, an ROC value = 0.5 and > 0.56 imply that the answer is 50 and 95 % correct, respectively.

Second, the information capacity for the PPTN neuronal ensemble to signal reward magnitude during the three task periods was estimated via mutual information analysis where:

$$I_{\text{reward}} = [- (L/N) \log_2(L/N) - (S/N) \log_2(S/N) - (High/N) \log_2(High/N) - (Low/N) \log_2(Low/N)] - [- \sum_{i=1}^2 (l_i/N) \log_2(l_i/N) - \sum_{i=1}^2 (s_i/N) \log_2(s_i/N)] \tag{4}$$

$L$ ,  $S$  and  $N$  denote numbers of large and small reward and total trials respectively.  $High$  and  $Low$  denote the numbers of trials where the neuronal response was larger and smaller than the median response for all trials, respectively. Therefore  $l_1$  and  $l_2$  and  $s_1$  and  $s_2$  represent large and small reward trials where the neuronal response was larger and smaller than the median response, respectively. Mutual information plots for individual neurons evaluate the information capacity for the neurons to express the reward magnitude in terms of a response correlation with the reward magnitude, and cumulative plots evaluate that for the ensemble neurons for an ideal case where the individual neuronal responses are perfectly independent.

Therefore, these two analyses estimate different aspects of neuronal signal precision, although they are related. Our ROC methods estimate the signal significance in comparison with the baseline noise, and the mutual information analysis evaluates the signal precision in terms of signal correlation with the reward magnitude.

We conducted an ROC analysis on the 45 fixation target and reward delivery neurons to estimate how reliably the discharges of the individual neurons indicated whether the reward was large or small. ROC values for the fixation target neurons (top 30 rows in Fig. 3A) started out near the chance level (ROC value = 0.5) and generally first acquired significance (ROC value > 0.56) during the fixation target/ cue period. Most fixation target neurons continued to show significant ROC values through the working memory periods after the fixation target/ cue disappeared, albeit with some substantial fluctuations, and more than half of them remained above the chance level even after reward delivery. The ROC values of the reward delivery neurons (the bottom rows in Fig. 3A), on the other hand, did not rise above chance level until after reward delivery, and then only transiently. Thus, the ROC analysis reinforced the idea that the fixation target neurons convey information about the magnitude of the predicted reward during the cue and working memory periods as well as up to and beyond the time of reward delivery and the reward delivery neurons convey information about the magnitude of the current reward only after it has been delivered. The free reward paradigm experiment also supports this view (Okada et al., 2009).

We obtained further support for this view by computing the mutual information (Kitazawa et al., 1998) in the responses about the magnitude of the reward (large or small). The cumulative plots of the mutual information conveyed by the individual fixation target neurons (cyan traces in Fig. 3B) indicated that the information grew rapidly during the fixation target/ cue period, peaked roughly as the fixation target/ cue disappeared, and then declined thereafter during the working memory period, but did not reach baseline in most neurons until after the reward was delivered, as did the ROC values. The mutual information conveyed by the individual reward delivery neurons (black traces in Fig. 3B) generally did not rise above the null level until after reward delivery, when it showed an abrupt substantial increase often lasting more than half a second.

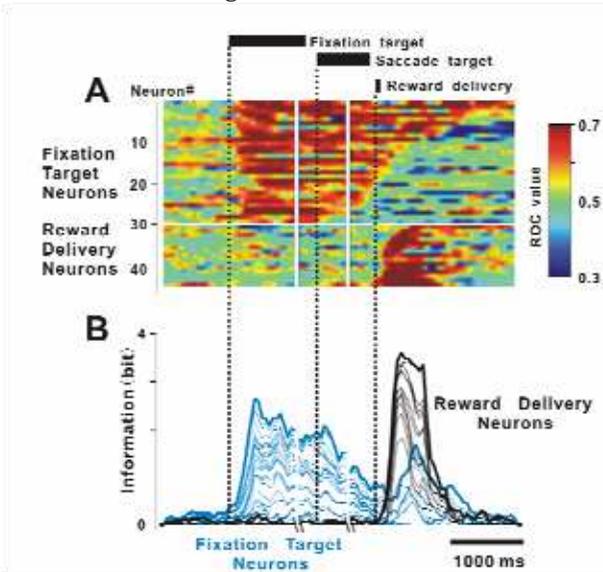


Fig. 3. Receiver operating characteristic (ROC) and mutual information analyses of responses in the fixation target and reward delivery neurons.

**A**, Pseudo color plots of the instantaneous ROC values (sliding time window, 200 ms) for large and small rewards indicated by activities in each of the 30 fixation target and 15 reward delivery neurons. The plots are aligned to fixation target onset and saccade target onset, and reward delivery (dotted lines) and ordered according to the neuronal ROC value after reward delivery. A white horizontal line indicates the border between the fixation target and reward delivery neurons. **B**, Cumulative plots of mutual information about reward amounts encoded by the 30 fixation target (cyan traces) and 15 reward delivery (black traces) neurons. A thick horizontal white bar indicates the duration of the respective neuronal type. Time axes in **A** and **B** are broken to align the responses to the onset of fixation, saccade target and reward delivery. Figures were modified from our recent paper (Okada et al., 2009).

Further insights were obtained by recording the activities of fixation target and reward delivery neurons in a *context reversal paradigm*, in which the meaning of the fixation target/ cue was suddenly reversed while recording from a given neuron so that squares and circles indicated large and small rewards, respectively, in the first 10 trials and the opposite

in the next 10 trials. The responses of the fixation target neurons during both the fixation target/ cue period (Fig. 4A) and the subsequent working memory period (the maintenance period of reward prediction) (Fig. 4B) clearly reflected the context reversal with a delay of one trial, the net result being that by the second trial after the context reversal the cue predicting the larger reward was again associated with the higher discharge rate (i.e., one-trial learning). In contrast, the responses of the reward delivery neurons did not change after the reversal, so that larger rewards were still associated with larger neuronal responses even on the first trial after the context reversal (Fig. 4C).

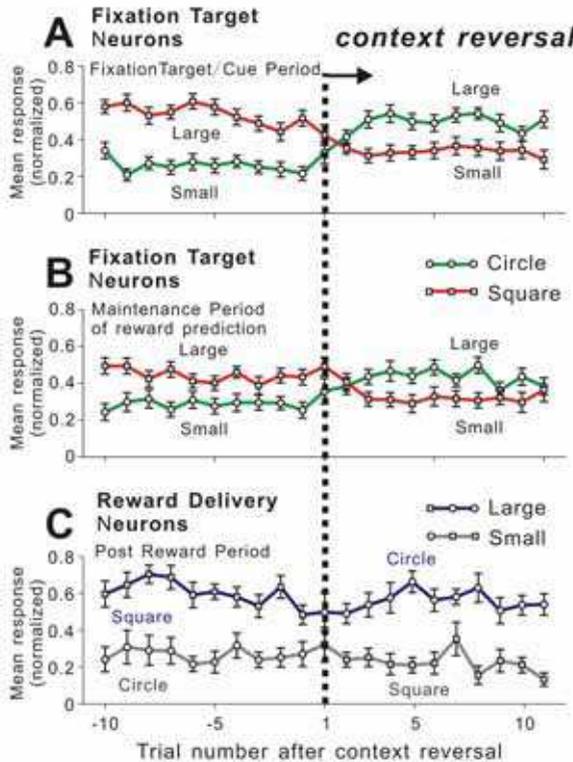


Fig. 4. Effects of context reversal on the responses of fixation target and reward delivery neurons. **A**, Responses of the fixation target neurons to fixation target (squares and circles) presentation (mean responses of 200-600 ms after fixation target on, fixation target/ cue period) before and after reversal of fixation target context (from squares and circles for large and small rewards in the initial 10 trials to squares and circles for small and large rewards in the last 10 trials). **B**, Similar to **A**, but for responses after fixation target offset (maintenance period of reward prediction, 200-600 ms after fixation target off). **C**, Similar to **A** and **B** but for the responses of the reward delivery neurons to reward delivery (200-600 ms after reward delivery, post reward delivery period) to large and small rewards. Responses were estimated as the average firing frequency normalized for the peak responses of the individual neurons. Error bars indicate standard error of mean. Figures were modified from our recent paper (Okada et al., 2009).

### 3.2 Correlation of fixation target response with behavioral performance

In a previous study on visually guided saccade task with single-value rewards we demonstrated that PPTN neuronal responses were stronger during trials that were successfully completed (Kobayashi et al., 2002). Therefore we questioned whether the reward prediction signaled by the fixation target neuronal response might also be related to the motivation of the monkey to perform the task. We tested this in the two-value reward task by studying the correlation of the fixation target responses of the the reward magnitude-dependent fixation target neurons with the task performance. Figure 5 shows the comparison of representative and ensemble fixation target neuronal responses to large and small rewards across fixation error, fixation-hold error, and successful trials. This representative neuron (Fig. 5A) showed no significant increase in its activity during the entire period of the fixation error trials, during which the animal failed to fixate on the target. Conversely, in the fixation hold error trials during which the animal did initially fixate on the target but failed to maintain the fixation, the activity increased during the pre-cue period (onset, -100 ms from fixation target presentation) and declined roughly at the time of the fixation break (200 ms, cf. the upward arrow in the cyan eye movement trace of Fig. 5A). The pre-cue period response during these trials was reward magnitude-independent in that the responses during the large and small reward trials were nearly equal, while the response was magnitude-dependent during the cue period, being larger for large reward trials than for small reward ones (cf. the red and green spike density traces before the dotted line with those after the line in Fig. 5A). In the successful trials, the fixation target period responses also consisted of a reward magnitude-independent component during the pre-cue period that matched that for the fixation hold error trials (cf. cyan spike density trace with the red and green spike density traces of Fig. 5A). A late reward magnitude-dependent component that emerged during the fixation target/ cue period, was much stronger than that in the fixation hold error trials and was sustained across the maintenance period until the postreward delivery period. The ensemble response for the fixation target neurons also showed a similar tendency as that of the representative neuron (Fig. 5B). The pre-cue period response was virtually absent in the fixation error trials, but there were significant pre-cue period responses in the fixation hold error and the successful trials. The magnitude-dependent response in the fixation hold error trials was small and transient, while that in the successful trials was much larger and was sustained until the post-reward delivery period.

The fact that the reward magnitude-independent pre-cue period response was absent in the fixation error trials and commonly present in both the fixation hold error and the successful trials indicates that it may reflect the monkey's motivation to fixate on the fixation target in anticipation of its presentation. Although the task intervals were quasi-randomized, monkeys appeared to be able to anticipate the onset of the fixation target and to be motivated to fixate on the target in both the fixation hold error and the successful trials prior to fixation target onset, but were probably not motivated to do so in the fixation error trials.

In addition, these findings indicate that the activities of the 52 reward magnitude-independent neurons also signal the early component of the motivational drive to fixate on the fixation target in an almost equal fashion as that of the reward magnitude-dependent fixation target neurons (Fig. 5C).

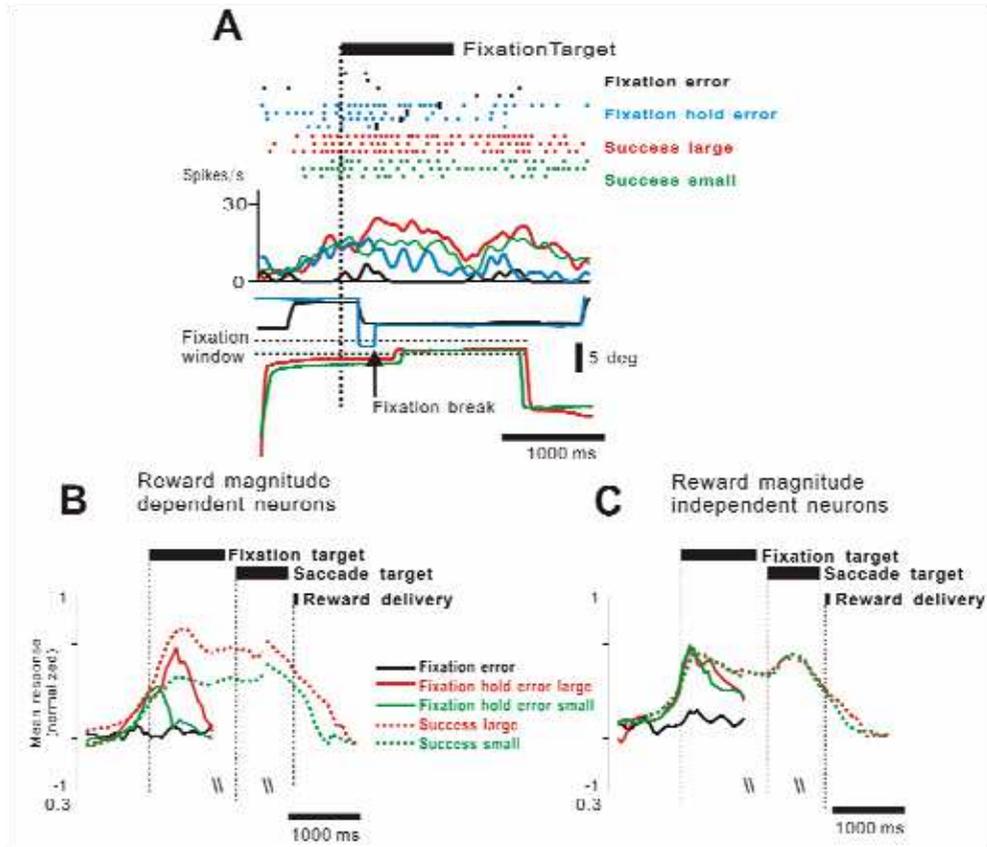


Fig. 5. Fixation target neuronal responses in unsuccessful and successful trials of the two-valued visually guided saccade task. *A*, (top and middle) Rastergram and spike density functions of a representative reward magnitude-dependent fixation target neuronal response over five success, fixation hold error and fixation error trials with normal cue (bottom) eye positions in a single representative case of each of the four trial categories. Upward arrow indicates the time of the fixation break. Two horizontal dotted lines indicate the fixation window within which the monkey was required to maintain eye position. *B*, Population spike density function of 30 reward magnitude-dependent fixation target neurons averaged for fixation error (black solid trace), fixation hold error (solid red and solid green traces for trial with large and small reward cues), and successful trials (dotted red and dotted green traces for trial with large and small reward cues), aligned to fixation target onset, saccade target onset and reward delivery. The spike density is the population average normalized for the peaks of the mean individual neuronal responses. *C*, Population spike density functions of 52 reward magnitude-independent neurons following the same format as (*B*). Figures were modified from our recent paper (Okada et al., 2009).

## 4. Computation of reward prediction error in dopamine neurons with input from the PPTN

### 4.1 PPTN neuronal activity for predicted and actual reward

We previously demonstrated that PPTN activity in the fixation period of a simple visually-guided saccade task predicted task outcome (Kobayashi et al., 2002). In the two-valued reward visually guided saccade task just described, we revealed new functional aspects of PPTN activity. The temporal profiles of the activities of fixation target and reward delivery neurons in this task indicated that these functional neuronal classes may encode the predicted and actual reward magnitudes, respectively. ROC analysis of the magnitude-dependent fixation target and reward delivery neuronal responses in our task revealed that most fixation target and reward delivery neurons reliably signaled whether reward is large or small. Mutual information analysis further showed that fixation target and reward delivery neurons signaled reward magnitude with high precision (maximum information capacities of 2.6 and 3.5 bits, corresponding to 0.04 and 0.25 bits/ neuron), comparable to those reported for the sensory (0.2 bits/ neuron (Gochin et al., 1994)) and motor systems (0.05 bits/ neuron (Kitazawa et al., 1998)). The high information capacities of fixation target and reward delivery neurons imply that they are potentially capable of differentiating 6 and 11 levels of reward magnitude, respectively. Mutual information analysis also showed that fixation target neurons conveyed information about predicted reward magnitude throughout the cue and maintenance periods, with no significant attenuation until the reward delivery neurons signaled actual reward magnitude.

Finally, the fixation target neurons responded to changes in the cue-reward contingency within two trials, rapidly revising their prediction of reward magnitude following changes in cue shape. These results are consistent with a role of fixation target neurons in reward prediction error computation in reinforcement learning. Conversely, the responses of the reward delivery neurons were based on the magnitude of the rewards delivered, regardless of cue shape. These results are consistent with reward delivery neurons signalling the magnitude of the delivered reward.

### 4.2 PPTN neuronal activity for motivation

Consistent with previous lesion (Conde et al., 1998) and recording studies (Kobayashi et al., 2002), PPTN (the fixation target) neurons may also signal motivation to perform a given task, the monkey's reaction times to fixate on the fixation and saccade targets were significantly correlated with their subsequent successful/ unsuccessful completion of the task, and the responses of the fixation target neurons were significantly smaller in the fixation error trials than in the successful trials.

We also found a significant number of reward magnitude-independent fixation target neurons whose responses were significantly correlated with the successful/ unsuccessful completion of the task (Fig. 5C). The functional implication of the reward magnitude-independent fixation target neurons remains unclear, but they may represent the timestamp of the reward expectation (Pan & Hyland, 2005). The neurons responsive to reward delivery also included reward magnitude-dependent and -independent groups; however, none of these reward delivery neurons showed a response correlation with the successful/ unsuccessful completion of the task, which is consistent with the view that they monitor the time and magnitude of the actual task reward. Finally, the responses of the reward magnitude-dependent fixation target and reward delivery neurons did not signal

only the reward magnitude but also the timestamps of reward expectation like those found in the reward magnitude-independent fixation target and reward delivery neurons; this was reflected in the anticipatory responses preceding the onset of the fixation target and reward delivery.

### 4.3 Possible source of the fixation target response

We demonstrated neuronal activity within the PPTN in response to the appearance of a fixation target that was predictive of the animal's performance on the task (Fig. 5) (Kobayashi et al., 2002). This activity appeared to be related to motivation level, reward prediction and conditioned sensory responses. This last association is consistent with a previous study in cats showing that neuronal activity in the PPTN was elicited during classical conditioning tasks in response to the conditioned stimulus (Dormont et al., 1998). Our result further suggests that the salience of the conditioned stimulus in the particular task (i.e. fixation target onset in the visually guided saccade task) was influenced by the monkey's motivation for performing the task. Thus, PPTN neurons may comprise a substrate, whose role is to transform a sensory cue into a behavioral action. If this hypothesis is correct, it is quite reasonable to expect the response of a given neuron to a cue in a cue-reward association task are modulated by the magnitude of the expected reward.

From where does the PPTN receive this motivational or reward prediction signal? The fixation target neurons may receive the signals of reward prediction from the orbitofrontal cortex (Tremblay & Schultz, 1999; Hikosaka & Watanabe, 2000; Roesch & Olson, 2004; Simmons & Richmond, 2008), prefrontal cortex (Kitazawa et al., 1998; Leon & Shadlen, 1999; Roesch & Olson, 2003; Kennerley & Wallis, 2009; Luk & Wallis, 2009), cingulate cortex (Cornwall et al., 1990), striatum (Mena-Segovia et al., 2004; Hikosaka et al., 2006; Winn, 2006) or hippocampus (Yang & Mogenson, 1987).

We propose that the signals travel via 1) the ventral striatum-ventral pallidum pathway, which receives input mainly from the limbic cortex (Yang & Mogenson, 1987; Schultz et al., 1992; Brown et al., 1999), 2) the amygdala and the subthalamic nucleus (Sesma & Fibiger, 1992), and 3) the cerebral cortices. Recently, Matsumura has emphasized the functional role of cortical input to the PPTN in the integration mechanism of limbic-motor control (Matsumura, 2005).

The dopamine neurons respond to expected, unexpected and salient sensory events with short latency, but little is known about the sensory systems underlying this response (Ljungberg et al., 1992). Studies of rats, cats and primates indicate that neurons in the superior colliculus, relaying visual information, make direct synaptic contacts with dopamine neurons in the substantia nigra (Comoli et al., 2003; McHaffie et al., 2006; May et al., 2009). In addition to the inputs of the substantia nigra via the superior colliculus, the dopamine neurons are also innervated by neurons, as described above. Furthermore, as the PPTN also receives input from the superior colliculus (Huerta & Harting, 1982; Redgrave et al., 1987; May & Porter, 1992). We propose that the PPTN may also relay visual information to dopamine neurons. We showed that PPTN neurons exhibited responses to the fixation target (a salient visual stimulus) that varied with subsequent performance of the task (Fig. 5). The responses of some of these neurons occurred with short latency (about 100ms), similar to the reported latency of dopamine neurons to the cue signal (50-120 ms, (Mirenowicz & Schultz, 1994; Schultz, 1998)). There have been only a few studies examining visual responses of PPTN neurons. Pan & Hyland (2005), reported visual responses of PPTN

neurons in rats, which had a mean response latency to the onset of a light stimulus of 70 ms, but they observed no variable visual responses for reward prediction (Pan & Hyland, 2005). In contrast to these results, a population of our recorded PPTN neurons in primates responded differentially to a visual stimulus with dependent on motivational state. Our results may be closer to another study of PPTN neurons in cats, whose conditioned cue responses occurred with a short latency (Dormont et al., 1998). Further studies are needed to examine the effect of reward prediction on the short latency response to salient stimulus in the PPTN (Stewart & Dommett, 2006).

Interestingly, similar to the cholinergic structure PPTN, the noradrenergic locus coeruleus has been implicated in responses to both salient and motivational sensory events. Locus coeruleus neurons were phasically activated prior to behavioral responses on both correct and incorrect trials, but were not activated by stimuli that failed to elicit lever responses or by lever movements outside the task (Clayton et al., 2004). In contrast to the locus coeruleus neurons, we observed a sustained, tonic activity in the PPTN during the task. Recent pharmacological studies suggest that another monoaminergic neurotransmitter, serotonin, is also involved in reward processing. Nakamura and colleagues showed that serotonergic neurons in the dorsal raphe nucleus were tonically modulated by the size of expected reward with either a large- or small-reward preference, and after reward delivery, they were tonically modulated by the size of the received reward (Nakamura et al., 2008). Thus, dorsal raphe nucleus neurons also encode the expected and received reward value, albeit, in a different pattern than the PPTN neurons. There are reciprocal mutual, inhibitory interactions between PPTN, locus coeruleus, and dorsal raphe nucleus neurons (Koyama & Kayama, 1993). Thus, we should compare the reward-related activities of neurons in these area while controlling arousal, motivation, and learning.

#### **4.4 Possible primary reward signal in the PPTN**

In the PPTN, we observed transient reward responses for free reward and reward during the two-valued reward task (Okada et al., 2009). The reward delivery neurons may receive the actual reward signals from the lateral hypothalamus (Rolls et al., 1980; Fukuda et al., 1986; Nakamura & Ono, 1986). This pathway directly excites the PPTN (Semba & Fibiger, 1992), which responds with a brief burst and then accommodates or habituates (Takakusaki et al., 1997; Dormont et al., 1998). This brief burst, in turn, directly excites the midbrain dopamine neurons via cholinergic and glutamatergic projections (Conde, 1992) and thereby causes a phasic burst in dopamine neurons projecting to the striatum (Gerfen, 1992) for actual reward. We plan to examine whether the response properties of the PPTN fulfill the necessary features of a primary reward signal (i.e., whether the activity is related to reward occurrence, to value coding, and shows no adaptation under a fully learned condition).

#### **4.5 Computation of reward prediction error signal in dopamine neurons**

As described above, dopamine neurons have unique firing patterns related to the predicted volume and actual times of reward (Hollerman & Schultz, 1998; Schultz, 1998). Computational models (Houk et al., 1995; Montague et al., 1996; Schultz et al., 1997; Berns & Sejnowski, 1998; Suri & Schultz, 1998; Contreras-Vidal & Schultz, 1999) of dopamine firing have noted similarities between the response patterns of dopamine neurons and well-known learning algorithms, especially temporal difference reinforcement learning algorithms (Montague et al., 1996; Schultz et al., 1997; Suri & Schultz, 1998). The temporal difference

model uses fast-sustained excitatory reward prediction and delayed slow-sustained inhibitory pulse signals in dopamine neurons, a sustained tonic reward prediction pulse originating from the striatum is temporally differentiated to produce an onset burst followed by an offset suppression. In the model the neurons in the striatum (the striosome) provide a significant source of GABAergic inhibition to dopamine neurons (Gerfen, 1992), and the fast excitatory, reward-predicting signals are derived via a double inhibition mechanism to dopamine neurons (matrisome-pallidum-dopamine neuron pathway (Houk et al., 1995)). Thus, the polysynaptic double inhibition pathway and monosynaptic direct inhibition may provide temporal differentiation of reward prediction in dopamine neurons. However, the model may not be realistic, because it is assumed that (1) the polysynaptic, net excitatory signal is faster than the direct monosynaptic inhibitory signal, and (2) the double inhibition pathway is required to strongly excite burst activity in dopamine neurons in response to a conditioned cue. A significant difference between the model we will propose, derived from the present findings, and the previous model is the source of excitation for dopamine neurons (Contreras-Vidal & Schultz, 1999). We propose that the excitatory PPTN neurons may send both a tonic reward prediction signal and a transient current reward signal to dopamine neurons.

Interestingly, the predictive and actual reward responses of the fixation target and reward delivery neurons follow comparable time courses to those supposed for the value function and the actual reward signals, respectively, in the temporal difference model of reinforcement learning (Houk et al., 1995; Schultz et al., 1997; Doya, 2000; Suri, 2002; Laurent, 2008). Therefore, the reward prediction error may be computed in the dopamine neurons from the fixation target and reward delivery signals, using the temporal difference algorithm, (Doya, 2000).

It is known from the classical conditioning paradigm of reinforcement learning that dopamine neurons show transient excitatory responses to cue presentation but not to reward delivery, and inhibitory responses to reward omission at the expected reward delivery time (Brown et al., 1999; Contreras-Vidal & Schultz, 1999; Doya, 2000; Fiorillo et al., 2008). The fixation target neuronal response that slowly rises at fixation target/ cue presentation may be conveyed to the dopamine neurons, transformed by temporal differentiation of the temporal difference mechanism as transient excitatory (Lokwan et al., 1999) and inhibitory signals timed at fixation target presentation and reward delivery, respectively, and summed with the actual reward signals of the reward delivery neurons, for computation of reward prediction errors.

The excitatory transients impinge on the dopamine neurons in the absence of neuronal reward delivery signals, producing a sharp cue response, while upon reward delivery, the inhibitory transients are summed with the excitatory actual reward signals for computation of the reward prediction error, producing no response when the reward prediction matches with the actual one (Tobler et al., 2003; Fiorillo et al., 2008).

In our recent study, the fixation target responses in the PPTN do not primarily explain this inhibitory omission response of the dopamine neurons, as the responses of the majority of the fixation target neurons were shutdown at the actual, rather than the expected, reward delivery timing in the temporal reward omission experiments (Okada et al., 2009). Therefore, they would feed the inhibitory transients to the dopamine neurons through the temporal difference mechanism, at the time of the actual rather than the expected reward. However, a minority of fixation target neurons, whose responses were terminated at the time of the expected reward delivery (Okada et al., 2009), could convey the inhibitory

transients to the dopamine neurons, producing the inhibitory omission response. It is possible that the former and latter fixation target neurons, whose responses were shutdown at the times of the actual and expected rewards, respectively, represent the value functions  $V(t)$  and  $V(t+1)$  for the current and predicted task events (Houk et al., 1995; Sutton & Barto, 1998; Doya, 2000). Furthermore GABAergic axon terminals originating from the PPTN were observed in the midbrain (Charara et al., 1996), these inhibitory connections may inhibit dopamine neurons and generate the inhibitory reward omission response. Alternatively, the inhibitory reward signals may be sent to the dopamine neurons from other neuronal structures such as the striosome (Brown et al., 1999; Contreras-Vidal & Schultz, 1999), ventral pallidum (Wu et al., 1996), habenula (Matsumoto & Hikosaka, 2007) and rostromedial tegmental nucleus (Jhou et al., 2009).

Finally, we present our hypothesis of how the PPTN drives dopamine neurons to compute the reward prediction error signal (Fig. 6). Our recent observations support the view that the fixation target and reward delivery neurons signal the predicted and actual reward magnitude, respectively. The prolonged response of the fixation target neurons indicates that they may maintain the signals of the predicted reward from the time of cue presentation until the reward delivery neurons signal the actual reward magnitude.

This study revealed that the strong excitatory inputs exerted by the PPTN on midbrain dopamine neurons (Mena-Segovia et al., 2004; Pan & Hyland, 2005; Winn, 2006) convey the memory of the predicted reward and the signals of the actual reward, two essential elements needed for computing the reward prediction error. The high information capacities of the fixation target and reward delivery neurons to signal the reward magnitude may help the dopamine neurons to accurately compute the reward prediction error and to efficiently execute reinforcement learning.

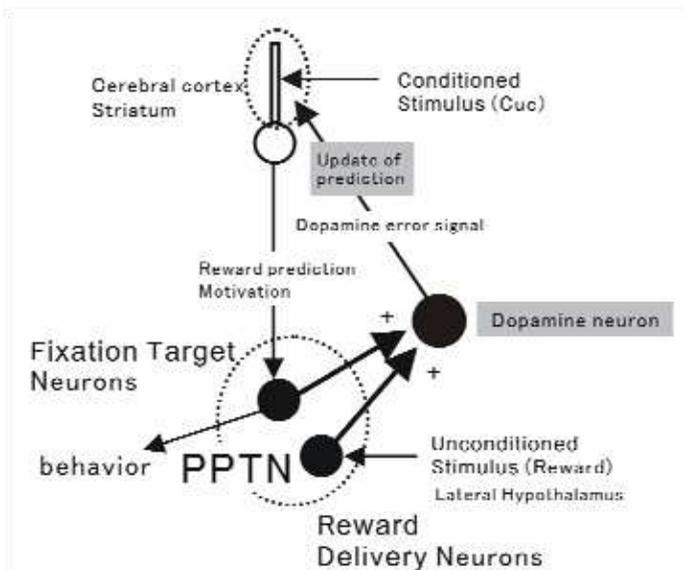


Fig. 6. Possible PPTN neuronal circuit for exciting dopamine neurons in reinforcement learning.

Computation of the reward prediction error requires a temporal memory of the predicted reward (established at cue onset and sustained until reward delivery) and a comparison of the actual reward with the predicted one. The reward predictive structures (cerebral cortex and striatum) may learn the cue-reward magnitude contingency during the training and task periods as a synaptic memory and recall that memory as the signals of the predicted reward magnitude at the time of cue presentation. These signals would then be transferred to the fixation target neurons and stored as working memory (Compte, 2006) of the reward prediction until the time of reward delivery. Thus, the PPTN is an important center, providing information of both reward prediction and actual reward to dopamine neurons. Moreover, our study addresses the broader science of memory: we demonstrated that the memory of the task reward is recalled as neuronal activity signaling the predicted reward magnitude, which is then compared with neuronal activity signaling the actual reward magnitude. To our knowledge, the mechanism whereby past memories, engrained in synaptic efficacy, are decoded into dynamic neural activity for comparison with the current neuronal activity, remains totally unexplored, in spite of the fact that the inverse process of encoding the firing rate of current neural events into synaptic efficacy has been extensively studied by plasticity researchers. Thus, our study is the first demonstration that structural memories of past experience are decoded into dynamic neural activity and compared with that for the present experience. And moreover, that the PPTN is the site where both signals are simultaneously represented.

## 5. References

- Alderson, H.L.; Latimer, M.P. & Winn, P. (2006) Intravenous self-administration of nicotine is altered by lesions of the posterior, but not anterior, pedunculopontine tegmental nucleus. *Eur J Neurosci* 23: pp. 2169-2175.
- Alderson, H.L.; Brown, V.J.; Latimer, M.P.; Brasted, P.J.; Robertson, A.H. & Winn, P. (2002) The effect of excitotoxic lesions of the pedunculopontine tegmental nucleus on performance of a progressive ratio schedule of reinforcement. *Neuroscience* 112: pp. 417-425.
- Bechara, A. & van der Kooy, D. (1989) The tegmental pedunculopontine nucleus: a brain-stem output of the limbic system critical for the conditioned place preferences produced by morphine and amphetamine. *J Neurosci* 9: pp. 3400-3409.
- Beckstead, R.M.; Domesick, V.B. & Nauta, W.J. (1979) Efferent connections of the substantia nigra and ventral tegmental area in the rat. *Brain Res* 175: pp. 191-217.
- Beninato, M. & Spencer, R.F. (1987) A cholinergic projection to the rat substantia nigra from the pedunculopontine tegmental nucleus. *Brain Res* 412: pp. 169-174.
- Berns, G.S. & Sejnowski, T.J. (1998) A computational model of how the basal ganglia produce sequences. *J Cogn Neurosci* 10: pp. 108-121.
- Blaha, C.D. & Winn, P. (1993) Modulation of dopamine efflux in the striatum following cholinergic stimulation of the substantia nigra in intact and pedunculopontine tegmental nucleus-lesioned rats. *J Neurosci* 13: pp. 1035-1044.
- Blaha, C.D.; Allen, L.F.; Das, S.; Inglis, W.L.; Latimer, M.P.; Vincent, S.R. & Winn, P. (1996) Modulation of dopamine efflux in the nucleus accumbens after cholinergic stimulation of the ventral tegmental area in intact, pedunculopontine tegmental nucleus-lesioned, and laterodorsal tegmental nucleus-lesioned rats. *J Neurosci* 16: pp. 714-722.

- Brown, J; Bullock, D. & Grossberg, S. (1999) How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *JNeurosci* 19: pp. 10502-10511.
- Calabresi, P.; Lacey, M.G. & North, R.A. (1989) Nicotinic excitation of rat ventral tegmental neurons in vitro studied by intracellular recording. *Br JPharmacol* 98: pp. 135-140.
- Chapman, C.A.; Yeomans, J.S.; Blaha, C.D. & Blackburn, J.R. (1997) Increased striatal dopamine efflux follows scopolamine administered systemically or to the tegmental pedunculopontine nucleus. *Neuroscience* 76: pp. 177-186.
- Charara, A.; Smith, Y. & Parent, A. (1996) Glutamatergic inputs from the pedunculopontine nucleus to midbrain dopaminergic neurons in primates: Phaseolus vulgaris-leucoagglutinin anterograde labeling combined with postembedding glutamate and GABA immunohistochemistry. *JComp Neurol* 364: pp. 254-266.
- Chen, J; Nakamura, M.; Kawamura, T.; Takahashi, T. & Nakahara, D. (2006) Roles of pedunculopontine tegmental cholinergic receptors in brain stimulation reward in the rat. *Psychopharmacology (Berl)* 184: pp. 514-522.
- Chiba, T.; Kayahara, T. & Nakano, K. (2001) Efferent projections of infralimbic and prelimbic areas of the medial prefrontal cortex in the Japanese monkey, *Macaca fuscata*. *Brain Res* 888: pp. 83-101.
- Clayton, E.C.; Rajkowski, J; Cohen, J.D. & Aston-Jones, G. (2004) Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. *J Neurosci* 24: pp. 9914-9920.
- Clements, J.R. & Grant, S. (1990) Glutamate-like immunoreactivity in neurons of the laterodorsal tegmental and pedunculopontine nuclei in the rat. *Neurosci Lett* 120: pp. 70-73.
- Comoli, E.; Coizet, V.; Boyes, J; Bolam, J.P.; Canteras, N.S.; Quirk, R.H.; Overton, P.G. & Redgrave, P. (2003) A direct projection from superior colliculus to substantia nigra for detecting salient visual events. *Nat Neurosci* 6: pp. 974-980.
- Compte, A. (2006) Computational and in vitro studies of persistent activity: edging towards cellular and synaptic mechanisms of working memory. *Neuroscience* 139: pp. 135-151.
- Conde, H. (1992) Organization and physiology of the substantia nigra. *Exp Brain Res* 88: pp. 233-248.
- Conde, H.; Dormont, J.F. & Farin, D. (1998) The role of the pedunculopontine tegmental nucleus in relation to conditioned motor performance in the cat. II. Effects of reversible inactivation by intracerebral microinjections. *Exp Brain Res* 121: pp. 411-418.
- Contreras-Vidal, J.L. & Schultz, W. (1999) A predictive reinforcement model of dopamine neurons for learning approach behavior. *JComput Neurosci* 6: pp. 191-214.
- Cornwall, J; Cooper, J.D. & Phillipson, O.T. (1990) Afferent and efferent connections of the laterodorsal tegmental nucleus in the rat. *Brain Res Bull* 25: pp. 271-284.
- Dormont, J.F.; Conde, H. & Farin, D. (1998) The role of the pedunculopontine tegmental nucleus in relation to conditioned motor performance in the cat. I. Context-dependent and reinforcement-related single unit activity. *Exp Brain Res* 121: pp. 401-410.
- Doya, K. (2000) Reinforcement learning in continuous time and space. *Neural Comput* 12: pp. 219-245.
- Doya, K. (2002) Metalearning and neuromodulation. *Neural Netw* 15: pp. 495-506.

- Edley, S.M. & Graybiel, A.M. (1983) The afferent and efferent connections of the feline nucleus tegmenti pedunculopontinus, pars compacta. *J Comp Neurol* 217: pp. 187-215.
- Fiorillo, C.D.; Tobler, P.N. & Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299: pp. 1898-1902.
- Fiorillo, C.D.; Newsome, W.T. & Schultz, W. (2008) The temporal precision of reward prediction in dopamine neurons. *Nat Neurosci* 11: pp. 966-973.
- Floresco, S.B.; West, A.R.; Ash, B.; Moore, H. & Grace, A.A. (2003) Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat Neurosci* 6: pp. 968-973.
- Ford, B.; Holmes, C.J.; Mainville, L. & Jones, B.E. (1995) GABAergic neurons in the rat pontomesencephalic tegmentum: codistribution with cholinergic and other tegmental neurons projecting to the posterior lateral hypothalamus. *J Comp Neurol* 363: pp. 177-196.
- Forster, G.L. & Blaha, C.D. (2003) Pedunculopontine tegmental stimulation evokes striatal dopamine efflux by activation of acetylcholine and glutamate receptors in the midbrain and pons of the rat. *Eur J Neurosci* 17: pp. 751-762.
- Fujimoto, K.; Ikeguchi, K. & Yoshida, M. (1992) Impaired acquisition, preserved retention and retrieval of avoidance behavior after destruction of pedunculopontine nucleus areas in the rat. *Neurosci Res* 13: pp. 43-51.
- Fujimoto, K.; Yoshida, M.; Ikeguchi, K. & Nijima, K. (1989) Impairment of active avoidance produced after destruction of pedunculopontine nucleus areas in the rat. *Neurosci Res* 6: pp. 321-328.
- Fukuda, M.; Ono, T.; Nishino, H. & Nakamura, K. (1986) Neuronal responses in monkey lateral hypothalamus during operant feeding behavior. *Brain Res Bull* 17: pp. 879-883.
- Futami, T.; Takakusaki, K. & Kitai, S.T. (1995) Glutamatergic and cholinergic inputs from the pedunculopontine tegmental nucleus to dopamine neurons in the substantia nigra pars compacta. *Neurosci Res* 21: pp. 331-342.
- Garcia-Rill, E. (1991) The pedunculopontine nucleus. *Prog Neurobiol* 36: pp. 363-389.
- Garcia-Rill, E.; Biedermann, J.A.; Chambers, T.; Skinner, R.D.; Mrak, R.E.; Husain, M. & Karson, C.N. (1995) Mesopontine neurons in schizophrenia. *Neuroscience* 66: pp. 321-335.
- Gerfen, C.R. (1992) The neostriatal mosaic: multiple levels of compartmental organization. *J Neural Transm Suppl* 36: pp. 43-59.
- Gochin, P.M.; Colombo, M.; Dorfman, G.A.; Gerstein, G.L. & Gross, C.G. (1994) Neural ensemble coding in inferior temporal cortex. *J Neurophysiol* 71: pp. 2325-2337.
- Graybiel, A.M. (2005) The basal ganglia: learning new tricks and loving it. *Curr Opin Neurobiol* 15: pp. 638-644.
- Grenhoff, J.; Aston-Jones, G. & Svensson, T.H. (1986) Nicotinic effects on the firing pattern of midbrain dopamine neurons. *Acta Physiol Scand* 128: pp. 351-358.
- Grofova, I. & Zhou, M. (1998) Nigral innervation of cholinergic and glutamatergic cells in the rat mesopontine tegmentum: light and electron microscopic anterograde tracing and immunohistochemical studies. *J Comp Neurol* 395: pp. 359-379.
- Gronier, B. & Rasmussen, K. (1998) Activation of midbrain presumed dopaminergic neurones by muscarinic cholinergic receptors: an in vivo electrophysiological study in the rat. *Br J Pharmacol* 124: pp. 455-464.
- Hikosaka, K. & Watanabe, M. (2000) Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb Cortex* 10: pp. 263-271.

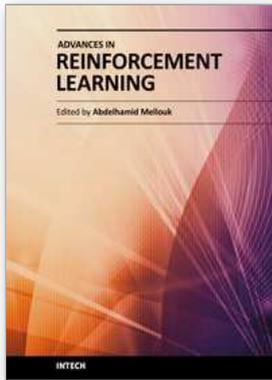
- Hikosaka, O.; Nakamura, K. & Nakahara, H. (2006) Basal ganglia orient eyes to reward. *J Neurophysiol* 95: pp. 567-584.
- Hollerman, J.R. & Schultz, W. (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1: pp. 304-309.
- Honda, T. & Semba, K. (1994) Serotonergic synaptic input to cholinergic neurons in the rat mesopontine tegmentum. *Brain Res* 647: pp. 299-306.
- Houk, J.C.; Adams, J.L. & Barto, A.G. (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of Information Processing in the Basal Ganglia*, pp 249-270. New York: The MIT Press.
- Huerta, M.F. & Harting, J.K. (1982) The projection from the nucleus of the posterior commissure to the superior colliculus of the cat: patch-like endings within the intermediate and deep grey layers. *Brain Res* 238: pp. 426-432.
- Ichinohe, N.; Teng, B. & Kitai, S.T. (2000) Morphological study of the tegmental pedunculopontine nucleus, substantia nigra and subthalamic nucleus, and their interconnections in rat organotypic culture. *Anat Embryol (Berl)* 201: pp. 435-453.
- Inglis, W.L.; Olmstead, M.C. & Robbins, T.W. (2000) Pedunculopontine tegmental nucleus lesions impair stimulus-reward learning in autoshaping and conditioned reinforcement paradigms. *Behav Neurosci* 114: pp. 285-294.
- Jackson, A. & Crossman, A.R. (1983) Nucleus tegmenti pedunculopontinus: efferent connections with special reference to the basal ganglia, studied in the rat by anterograde and retrograde transport of horseradish peroxidase. *Neuroscience* 10: pp. 725-765.
- Jhou, T.C.; Fields, H.L.; Baxter, M.G.; Saper, C.B. & Holland, P.C. (2009) The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* 61: pp. 786-800.
- Jones, B.E. & Beaudet, A. (1987) Distribution of acetylcholine and catecholamine neurons in the cat brainstem: a choline acetyltransferase and tyrosine hydroxylase immunohistochemical study. *J Comp Neurol* 261: pp. 15-32.
- Kayama, Y. & Koyama, Y. (2003) Control of sleep and wakefulness by brainstem monoaminergic and cholinergic neurons. *Acta Neurochir Suppl* 87: pp. 3-6.
- Kennerley, S.W. & Wallis, J.D. (2009) Reward-dependent modulation of working memory in lateral prefrontal cortex. *J Neurosci* 29: pp. 3259-3270.
- Kippin, T.E. & van der Kooy, D. (2003) Excitotoxic lesions of the tegmental pedunculopontine nucleus impair copulation in naive male rats and block the rewarding effects of copulation in experienced male rats. *Eur J Neurosci* 18: pp. 2581-2591.
- Kitai, S.T.; Shepard, P.D.; Callaway, J.C. & Scroggs, R. (1999) Afferent modulation of dopamine neuron firing patterns. *Curr Opin Neurobiol* 9: pp. 690-697.
- Kitazawa, S.; Kimura, T. & Yin, P.B. (1998) Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature* 392: pp. 494-497.
- Kobayashi, Y.; Inoue, Y.; Yamamoto, M.; Isa, T. & Aizawa, H. (2002) Contribution of pedunculopontine tegmental nucleus neurons to performance of visually guided saccade tasks in monkeys. *J Neurophysiol* 88: pp. 715-731.
- Koyama, Y. & Kayama, Y. (1993) Mutual interactions among cholinergic, noradrenergic and serotonergic neurons studied by iontophoresis of these transmitters in rat brainstem nuclei. *Neuroscience* 55: pp. 1117-1126.
- Lacey, M.G.; Calabresi, P. & North, R.A. (1990) Muscarine depolarizes rat substantia nigra zona compacta and ventral tegmental neurons in vitro through M1-like receptors. *J Pharmacol Exp Ther* 253: pp. 395-400.

- Laurent, P.A. (2008) The emergence of saliency and novelty responses from Reinforcement Learning principles. *Neural Netw* 21: pp. 1493-1499.
- Lavoie, B. & Parent, A. (1994) Pedunculopontine nucleus in the squirrel monkey: distribution of cholinergic and monoaminergic neurons in the mesopontine tegmentum with evidence for the presence of glutamate in cholinergic neurons. *J Comp Neurol* 344: pp. 190-209.
- Lee, D. & Seo, H. (2007) Mechanisms of reinforcement learning and decision making in the primate dorsolateral prefrontal cortex. *Ann N Y Acad Sci* 1104: pp. 108-122.
- Leon, M.I. & Shadlen, M.N. (1999) Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24: pp. 415-425.
- Leonard, C.S. & Llinas, R. (1994) Serotonergic and cholinergic inhibition of mesopontine cholinergic neurons controlling REM sleep: an in vitro electrophysiological study. *Neuroscience* 59: pp. 309-330.
- Ljungberg, T.; Apicella, P. & Schultz, W. (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *JNeurophysiol* 67: pp. 145-163.
- Lokwan, S.J.; Overton, P.G.; Berry, M.S. & Clark, D. (1999) Stimulation of the pedunculopontine tegmental nucleus in the rat produces burst firing in A9 dopaminergic neurons. *Neuroscience* 92: pp. 245-254.
- Luk, C.H. & Wallis, J.D. (2009) Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *JNeurosci* 29: pp. 7526-7539.
- Lusted, L.B. (1978) General problems in medical decision making with comments on ROC analysis. *Semin Nucl Med* 8: pp. 299-306.
- Matsumoto, M. & Hikosaka, O. (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447: pp. 1111-1115.
- Matsumura, M. (2005) The pedunculopontine tegmental nucleus and experimental parkinsonism. A review. *JNeurol* 252 Suppl 4: pp. IV5-IV12.
- Matsumura, M.; Watanabe, K. & Ohye, C. (1997) Single-unit activity in the primate nucleus tegmenti pedunculopontinus related to voluntary arm movement. *Neurosci Res* 28: pp. 155-165.
- May, P.J. & Porter, J.D. (1992) The laminar distribution of macaque tectobulbar and tectospinal neurons. *Vis Neurosci* 8: pp. 257-276.
- May, P.J.; McHaffie, J.G.; Stanford, T.R.; Jang, H.; Costello, M.G.; Coizet, V.; Hayes, L.M.; Haber, S.N. & Redgrave, P. (2009) Tectonigral projections in the primate: a pathway for pre-attentive sensory input to midbrain dopaminergic neurons. *Eur J Neurosci* 29: pp. 575-587.
- McHaffie, J.G.; Jang, H.; May, P.J.; Coizet, V.; Overton, P.G.; Stein, B.E. & Redgrave, P. (2006) A direct projection from superior colliculus to substantia nigra pars compacta in the cat. *Neuroscience* 138: pp. 221-234.
- Mena-Segovia, J.; Bolam, J.P. & Magill, P.J. (2004) Pedunculopontine nucleus and basal ganglia: distant relatives or part of the same family? *Trends Neurosci* 27: pp. 585-588.
- Mena-Segovia, J.; Winn, P. & Bolam, J.P. (2008) Cholinergic modulation of midbrain dopaminergic systems. *Brain Res Rev* 58: pp. 265-271.
- Mesulam, M.M.; Mufson, E.J.; Wainer, B.H. & Levey, A.I. (1983) Central cholinergic pathways in the rat: an overview based on an alternative nomenclature (Ch1-Ch6). *Neuroscience* 10: pp. 1185-1201.
- Miller, A.D. & Blaha, C.D. (2004) Nigrostriatal dopamine release modulated by mesopontine muscarinic receptors. *Neuroreport* 15: pp. 1805-1808.

- Mirenowicz, J & Schultz, W. (1994) Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72: pp. 1024-1027.
- Montague, P.R. & Berns, G.S. (2002) Neural economics and the biological substrates of valuation. *Neuron* 36: pp. 265-284.
- Montague, P.R.; Dayan, P. & Sejnowski, T.J. (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16: pp. 1936-1947.
- Morris, G.; Nevet, A.; Arkadir, D.; Vaadia, E. & Bergman, H. (2006) Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci* 9: pp. 1057-1063.
- Nakahara, H.; Itoh, H.; Kawagoe, R.; Takikawa, Y. & Hikosaka, O. (2004) Dopamine neurons can represent context-dependent prediction error. *Neuron* 41: pp. 269-280.
- Nakamura, K. & Ono, T. (1986) Lateral hypothalamus neuron involvement in integration of natural and artificial rewards and cue signals. *J Neurophysiol* 55: pp. 163-181.
- Nakamura, K.; Matsumoto, M. & Hikosaka, O. (2008) Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *J Neurosci* 28: pp. 5331-5343.
- Oakman, S.A.; Faris, P.L.; Kerr, P.E.; Cozzari, C. & Hartman, B.K. (1995) Distribution of pontomesencephalic cholinergic neurons projecting to substantia nigra differs significantly from those projecting to ventral tegmental area. *J Neurosci* 15: pp. 5859-5869.
- Okada, K.; Toyama, K.; Inoue, Y.; Isa, T. & Kobayashi, Y. (2009) Different pedunclopontine tegmental neurons signal predicted and actual task rewards. *J Neurosci* 29: pp. 4858-4870.
- Pan, W.X. & Hyland, B.I. (2005) Pedunclopontine tegmental nucleus controls conditioned responses of midbrain dopamine neurons in behaving rats. *J Neurosci* 25: pp. 4725-4732.
- Perry, E.; Walker, M.; Grace, J & Perry, R. (1999) Acetylcholine in mind: a neurotransmitter correlate of consciousness? *Trends Neurosci* 22: pp. 273-280.
- Pidoplichko, V.I.; DeBiasi, M.; Williams, J.T. & Dani, J.A. (1997) Nicotine activates and desensitizes midbrain dopamine neurons. *Nature* 390: pp. 401-404.
- Redgrave, P.; Mitchell, I.J & Dean, P. (1987) Descending projections from the superior colliculus in rat: a study using orthograde transport of wheatgerm-agglutinin conjugated horseradish peroxidase. *Exp Brain Res* 68: pp. 147-167.
- Reynolds, J.N.; Hyland, B.I. & Wickens, J.R. (2001) A cellular mechanism of reward-related learning. *Nature* 413: pp. 67-70.
- Richmond, B.J.; Liu, Z. & Shidara, M. (2003) Neuroscience. Predicting future rewards. *Science* 301: pp. 179-180.
- Roesch, M.R. & Olson, C.R. (2003) Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J Neurophysiol* 90: pp. 1766-1789.
- Roesch, M.R. & Olson, C.R. (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304: pp. 307-310.
- Rolls, E.T.; Burton, M.J & Mora, F. (1980) Neurophysiological analysis of brain-stimulation reward in the monkey. *Brain Res* 194: pp. 339-357.
- Samejima, K.; Ueda, Y.; Doya, K. & Kimura, M. (2005) Representation of action-specific reward values in the striatum. *Science* 310: pp. 1337-1340.
- Satoh, T.; Nakai, S.; Sato, T. & Kimura, M. (2003) Correlated coding of motivation and outcome of decision by dopamine neurons. *J Neurosci* 23: pp. 9913-9923.
- Scarnati, E.; Proia, A.; Campana, E. & Pacitti, C. (1986) A microiontophoretic study on the nature of the putative synaptic neurotransmitter involved in the pedunclopontine-

- substantia nigra pars compacta excitatory pathway of the rat. *Exp Brain Res* 62: pp. 470-478.
- Scarnati, E.; Proia, A.; Di Loreto, S. & Pacitti, C. (1987) The reciprocal electrophysiological influence between the nucleus tegmenti pedunculopontinus and the substantia nigra in normal and decorticated rats. *Brain Res* 423: pp. 116-124.
- Schreiner, R.C.; Essick, G.K. & Whitsel, B.L. (1978) Variability in somatosensory cortical neuron discharge: effects on capacity to signal different stimulus conditions using a mean rate code. *JNeurophysiol* 41: pp. 338-349.
- Schultz, W. (1998) Predictive reward signal of dopamine neurons. *JNeurophysiol* 80: pp. 1-27.
- Schultz, W. (2002) Getting formal with dopamine and reward. *Neuron* 36: pp. 241-263.
- Schultz, W.; Dayan, P. & Montague, P.R. (1997) A neural substrate of prediction and reward. *Science* 275: pp. 1593-1599.
- Schultz, W.; Apicella, P.; Scarnati, E. & Ljungberg, T. (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. *JNeurosci* 12: pp. 4595-4610.
- Scroggs, R.S.; Cardenas, C.G.; Whittaker, J.A. & Kitai, S.T. (2001) Muscarine reduces calcium-dependent electrical activity in substantia nigra dopaminergic neurons. *J Neurophysiol* 86: pp. 2966-2972.
- Semba, K. & Fibiger, H.C. (1992) Afferent connections of the laterodorsal and the pedunculopontine tegmental nuclei in the rat: a retro- and antero-grade transport and immunohistochemical study. *JComp Neurol* 323: pp. 387-410.
- Simmons, J.M. & Richmond, B.J. (2008) Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. *Cereb Cortex* 18: pp. 93-103.
- Sorenson, E.M.; Shiroyama, T. & Kitai, S.T. (1998) Postsynaptic nicotinic receptors on dopaminergic neurons in the substantia nigra pars compacta of the rat. *Neuroscience* 87: pp. 659-673.
- Spann, B.M. & Grofova, I. (1992) Cholinergic and non-cholinergic neurons in the rat pedunculopontine tegmental nucleus. *Anat Embryol (Berl)* 186: pp. 215-227.
- Steckler, T.; Inglis, W.; Winn, P. & Sahgal, A. (1994) The pedunculopontine tegmental nucleus: a role in cognitive processes? *Brain Res Brain Res Rev* 19: pp. 298-318.
- Steiniger, B. & Kretschmer, B.D. (2003) Glutamate and GABA modulate dopamine in the pedunculopontine tegmental nucleus. *Exp Brain Res* 149: pp. 422-430.
- Steininger, T.L.; Rye, D.B. & Wainer, B.H. (1992) Afferent projections to the cholinergic pedunculopontine tegmental nucleus and adjacent midbrain extrapyramidal area in the albino rat. I. Retrograde tracing studies. *JComp Neurol* 321: pp. 515-543.
- Steriade, M.; Datta, S.; Pare, D.; Oakson, G. & Curro Dossi, R. (1990) Neuronal activities in brain-stem cholinergic nuclei related to tonic activation processes in thalamocortical systems. *JNeurosci* 10: pp. 2541-2559.
- Stewart, R.D. & Dommett, E.J. (2006) Subcortical control of dopamine neurons: the good, the bad and the unexpected. *Brain Res Bull* 71: pp. 1-3.
- Suri, R.E. (2002) TD models of reward predictive responses in dopamine neurons. *Neural Netw* 15: pp. 523-533.
- Suri, R.E. & Schultz, W. (1998) Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Exp Brain Res* 121: pp. 350-354.
- Sutton, R.S. & Barto, A.G. (1998) *Reinforcement Learning*. New York: The MIT Press.
- Takakusaki, K.; Shiroyama, T. & Kitai, S.T. (1997) Two types of cholinergic neurons in the rat tegmental pedunculopontine nucleus: electrophysiological and morphological characterization. *Neuroscience* 79: pp. 1089-1109.

- Takakusaki, K.; Shiroyama, T.; Yamamoto, T. & Kitai, S.T. (1996) Cholinergic and noncholinergic tegmental pedunculopontine projection neurons in rats revealed by intracellular labeling. *JComp Neurol* 371: pp. 345-361.
- Takakusaki, K.; Oohinata-Sugimoto, J.; Saitoh, K. & Habaguchi, T. (2004) Role of basal ganglia-brainstem systems in the control of postural muscle tone and locomotion. *Progress in Brain Research* 143: pp. 231-237.
- Tesauro, G. (1994) TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Computation* 6: pp. 215-219.
- Tobler, P.N.; Dickinson, A. & Schultz, W. (2003) Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *JNeurosci* 23: pp. 10402-10410.
- Tremblay, L. & Schultz, W. (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398: pp. 704-708.
- Tzavara, E.T.; Bymaster, F.P.; Davis, R.J.; Wade, M.R.; Perry, K.W.; Wess, J.; McKinzie, D.L.; Felder, C. & Nomikos, G.G. (2004) M4 muscarinic receptors regulate the dynamics of cholinergic and dopaminergic neurotransmission: relevance to the pathophysiology and treatment of related CNS pathologies. *FASEB J* 18: pp. 1410-1412.
- Waelti, P.; Dickinson, A. & Schultz, W. (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412: pp. 43-48.
- Wang, H.L. & Morales, M. (2009) Pedunculopontine and laterodorsal tegmental nuclei contain distinct populations of cholinergic, glutamatergic and GABAergic neurons in the rat. *Eur JNeurosci* 29: pp. 340-358.
- Watanabe, M. (1996) Reward expectancy in primate prefrontal neurons. *Nature* 382: pp. 629-632.
- Werner, G. & Mountcastle, V.B. (1963) The Variability of Central Neural Activity in a Sensory System, and Its Implications for the Central Reflection of Sensory Events. *J Neurophysiol* 26: pp. 958-977.
- Wickens, J.R.; Reynolds, J.N. & Hyland, B.I. (2003) Neural mechanisms of reward-related motor learning. *Curr Opin Neurobiol* 13: pp. 685-690.
- Wilson, D.I.; Maclaren, D.A. & Winn, P. (2009) Bar pressing for food: differential consequences of lesions to the anterior versus posterior pedunculopontine. *Eur J Neurosci* 30: pp. 504-513.
- Winn, P. (2006) How best to consider the structure and function of the pedunculopontine tegmental nucleus: evidence from animal studies. *JNeurol Sci* 248: pp. 234-250.
- Winn, P.; Brown, V.J & Inglis, W.L. (1997) On the relationships between the striatum and the pedunculopontine tegmental nucleus. *Crit Rev Neurobiol* 11: pp. 241-261.
- Wu, M.; Hrycyshyn, A.W. & Brudzynski, S.M. (1996) Subpallidal outputs to the nucleus accumbens and the ventral tegmental area: anatomical and electrophysiological studies. *Brain Res* 740: pp. 151-161.
- Yamashita, T. & Isa, T. (2003) Fulfenamic acid sensitive, Ca(2+)-dependent inward current induced by nicotinic acetylcholine receptors in dopamine neurons. *Neurosci Res* 46: pp. 463-473.
- Yang, C.R. & Mogenson, G.J. (1987) Hippocampal signal transmission to the pedunculopontine nucleus and its regulation by dopamine D2 receptors in the nucleus accumbens: an electrophysiological and behavioural study. *Neuroscience* 23: pp. 1041-1055.
- Yeomans, J.S. (1995) Role of tegmental cholinergic neurons in dopaminergic activation, antimuscarinic psychosis and schizophrenia. *Neuropsychopharmacology* 12: pp. 3-16.



## **Advances in Reinforcement Learning**

Edited by Prof. Abdelhamid Mellouk

ISBN 978-953-307-369-9

Hard cover, 470 pages

**Publisher** InTech

**Published online** 14, January, 2011

**Published in print edition** January, 2011

Reinforcement Learning (RL) is a very dynamic area in terms of theory and application. This book brings together many different aspects of the current research on several fields associated to RL which has been growing rapidly, producing a wide variety of learning algorithms for different applications. Based on 24 Chapters, it covers a very broad variety of topics in RL and their application in autonomous systems. A set of chapters in this book provide a general overview of RL while other chapters focus mostly on the applications of RL paradigms: Game Theory, Multi-Agent Theory, Robotic, Networking Technologies, Vehicular Navigation, Medicine and Industrial Logistic.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yasushi Kobayashi and Ken-ichi Okada (2011). Reward Prediction Error Computation in the Pedunculopontine Tegmental Nucleus Neurons, *Advances in Reinforcement Learning*, Prof. Abdelhamid Mellouk (Ed.), ISBN: 978-953-307-369-9, InTech, Available from: <http://www.intechopen.com/books/advances-in-reinforcement-learning/reward-prediction-error-computation-in-the-pedunculopontine-tegmental-nucleus-neurons>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.