



Reward prediction error signals by reticular formation neurons

Corey B. Puryear and Sheri J.Y. Mizumori

Learn. Mem. 2008 15: 895-898

Access the most recent version at doi:[10.1101/lm.1072808](https://doi.org/10.1101/lm.1072808)

References

This article cites 25 articles, 8 of which can be accessed free at:
<http://learnmem.cshlp.org/content/15/12/895.full.html#ref-list-1>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

To subscribe to *Learning & Memory* go to:
<http://learnmem.cshlp.org/subscriptions>

Brief Communication

Reward prediction error signals by reticular formation neurons

Corey B. Puryear¹ and Sheri J.Y. Mizumori²

University of Washington, Department of Psychology, Seattle, Washington 98195, USA

As a key part of the brain's reward system, midbrain dopamine neurons are thought to generate signals that reflect errors in the prediction of reward. However, recent evidence suggests that "upstream" brain areas may make important contributions to the generation of prediction error signals. To address this issue, we recorded neural activity in midbrain reticular formation (MRNm) while rats performed a spatial working memory task. A large proportion of these neurons displayed positive and negative reward prediction error-related activity that was strikingly similar to that observed in dopamine neurons. Therefore, MRNm may be a source of reward prediction error signals.

The capacity of an organism to respond appropriately to environmental stimuli depends on the ability to detect changes in the outcome of its behavior. The mesocorticolimbic dopamine system is thought to be central to this function (Wise 2004; Fields et al. 2007). Dopamine neurons in the ventral tegmental area (VTA) and substantia nigra *pars compacta* (SNc) increase activity relative to the presentation of cues that predict rewards and rewards of greater value than expected, and decrease activity relative to rewards of less value than predicted (Nakahara et al. 2004; Bayer and Glimcher 2005; Pan et al. 2005; Tobler et al. 2005). This activity is thought to be involved in a computation about errors in the prediction of reward (Schultz and Dickinson 2000) that can be used to correct behavior. A central issue relevant to the behavioral and computational interpretation of dopamine signals is whether prediction error signals are generated by dopamine neurons, *per se*, or by cells in "upstream" brain areas.

Recent data suggest that brain areas afferent to dopamine neurons generate, or participate in, reward prediction error computations. Lateral habenula, which has been shown to inhibit the activity of VTA and SNc dopamine neurons (Herkenham and Nauta 1979; Christoph et al. 1986), has recently been identified as a potential source of reward prediction error signals (Matsumoto and Hikosaka 2007). A similar finding has been demonstrated in the pedunculopontine tegmental nucleus (PPTg), which is an important regulator of dopamine neuron activity (Floresco et al. 2003). PPTg neural responses varied according to whether or not the animal received expected rewards (Kobayashi and Okada 2007).

As part of a larger study investigating the role of VTA in context-dependent spatial working memory (C.B. Puryear, M.J. Kim, and S.J.Y. Mizumori, in prep.), we recorded the activity of neurons in the magnocellular region of the midbrain reticular formation (MRNm according to the method of Swanson [2003]) at the level of the diencephalon. The reticular formation is thought to be important for modulating arousal and vigilance levels necessary for attending to and acting upon salient stimuli (Pragay et al. 1978; Mesulam 1981), and it has recently been shown that this portion of the reticular formation provides glutamatergic input to VTA (Geisler et al. 2007). Thus, MRNm is in a prime position to modulate the activity of VTA dopamine neurons when the outcome of behavior does not meet expectations and therefore may be a source of reward prediction error signals.

Accordingly, we investigated whether MRNm neurons exhibited reward related activity, and whether this activity was related to the ability to predict acquisition of reward.

Four male Long-Evans rats (4–6 mo old from Simonson Laboratory, Gilroy, CA) were housed individually in Plexiglas cages in a temperature- and humidity-controlled environment (12:12 h light:dark). Rats were provided with food and water *ad libitum* for 5 d prior to being handled daily and reduced to 85% of *ad libitum* feeding weights. Animal care and use was conducted according to University of Washington's Institutional Animal Care and Use Committee guidelines.

Rats were habituated to the testing environment and trained to perform a differential reward, win-shift spatial working memory task using radial maze procedures reported previously (Pratt and Mizumori 2001; C.B. Puryear, M.J. Kim, and S.J.Y. Mizumori, in prep.). Briefly, prior to the start of each trial, the end of each of the eight maze arms was baited with either a large (five drops) or small (one drop) amount of chocolate milk on alternating arms. Maze arms containing large or small amounts of reward were counterbalanced across rats and held constant throughout training. Trials started with a sample phase by presenting four maze arms (two large and two small reward arms; individually and randomly selected) to the rat. Immediately after presentation of the fourth arm, a test phase began by making all maze arms accessible so the rat could collect the remaining rewards. The trial ended once all arms were visited. The rat was then confined to the center of the maze for a 2-min intertrial interval. Arm re-entries were counted as errors. Once the rat was able to perform at 15 trials in less than 1 h for seven consecutive days, recording electrodes were surgically implanted.

Details concerning the construction of recording electrodes and microdrives and surgical procedures can be found in previous works (McNaughton et al. 1983; Puryear et al. 2006; C.B. Puryear, M.J. Kim, and S.J.Y. Mizumori, in prep.). Briefly, rats were chronically implanted with either eight stereotrodes (four/hemisphere) or four tetrodes (two/hemisphere) made from 25- μ m lacquer-coated tungsten wire (California Fine Wire), centered on the following coordinates relative to bregma: -5.25 mm posterior, 0.7 mm lateral, 6 mm ventral. One week of free feeding was allowed for rats to recover from surgery before recording experiments began.

Recordings were performed as described previously (Puryear et al. 2006; C.B. Puryear, M.J. Kim, and S.J.Y. Mizumori, in prep.). If no clear spontaneous neural activity was encountered, electrodes were lowered in ~ 25 - μ m increments (up to 175 μ m/d) until unambiguous, isolatable units were observed. Single units were isolated from multiunit records using standard cluster-

¹Present address: Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

²Corresponding author.

E-mail mizumori@u.washington.edu; fax (206) 685-3157.

Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.1072808>.

cutting software (MClust; A.D. Redish, University of Wisconsin). A template-matching algorithm was also used to facilitate separation of unique spike waveforms. We only included cells that had a high signal-to-noise ratio (>3:1), exhibited stable clusters throughout the recording session, and had clear refractory periods in the interspike interval histograms following cluster cutting.

The final position of each stereotrode was marked by passing a 25- μ A current through each recording wire for 25 sec while rats were under 5% isofluorane anesthesia. Rats were then given an overdose of sodium pentobarbital and transcardially perfused (0.9% buffered saline, followed by 10% formalin). Electrodes were retracted, and the brain was removed and allowed to sink in 30% sucrose-formalin. Coronal sections (40 μ m) were sliced with a cryostat and stained with cresyl violet. Recording locations were verified by comparing depth measurements and reconstructions of the electrode tracks. Only cells determined to be located in MRNm (Swanson 2003) were considered for analysis.

Rats were well trained on the spatial working memory task, committing 0.86 ± 0.2 (mean \pm SEM) errors per trial during the first five trials (baseline trials). Importantly, rats demonstrated the ability to discriminate large and small reward locations. There was a significant negative correlation between the first four test phase arm choices (i.e., first, second, third, and fourth arm choice) and the probability that the arm chosen contained a large reward (Fig. 1B; Spearman's $\rho = -0.65$, $P < 0.001$), indicating that rats reliably visited large reward arms before small reward arms during the test phase of each trial.

A total of 18 cells localized to MRNm were recorded while rats performed the task. Of these, one cell was omitted from analysis due to a very low average firing rate (-0.2 spikes/sec), yielding 17 cells included in the following analyses. Figure 1A

depicts the distribution of cells localized to MRNm. These cells exhibited a range of average firing rates, spike durations (defined as the time from the start to the end of the action potential) (Fig. 1C), and firing patterns (for representative interspike interval and autocorrelation histograms, see Fig. 1D).

Reward-related neural activity was obtained by placing rewards in small metal cups mounted to the end of each maze arm and connected to the recording equipment (custom designed by Neuralynx, Inc.), which served as "lick detectors." An event marker was automatically inserted into the data stream when the rat licked the cup, providing an instantaneous measurement of the time the rat first obtained reward.

In order to determine whether MRNm neurons exhibited significant reward-related activity, peri-event time histograms (PETHs) were constructed (50 msec bins, ± 2.5 sec around each reward event). A cell was considered to have a significant excitatory reward response if it passed the following two criteria: (1) The cell had a peak firing rate within ± 150 msec of reward acquisition and (2) the peak rate was >150% of its average firing rate for the block of trials. These criteria were applied to PETHs collapsed across reward amounts and separately for large and small reward events. Overall, 47% (eight of 17) of MRNm neurons were found to exhibit significant excitatory responses upon acquisition of reward (Fig. 2D). Of these cells, most (88%, seven of eight) were found to fire relative to acquisition of only large rewards (e.g., Fig. 2A–C), while the remaining neuron fired relative to acquisition of both reward amounts. No cells were found to fire preferentially to acquisition of only small rewards. Aspects of animals' movement (e.g., velocity) were not found to be a major contributor to the firing patterns of MRNm neurons during performance of the spatial working memory task (data not shown). Therefore, it appears that MRNm unit activity is predominantly biased to represent higher reward values.

In order to determine whether MRNm reward-related activity was associated with reward prediction, we tested unit responses to unexpected alterations of reward outcome or elimination of visuospatial information important for reward prediction. To do this, we allowed the rat to perform a second block of five trials with either the locations of large and small rewards switched (reward location switch condition), with two rewards (one large and one small, randomly selected) omitted from the study phase of each trial (reward omission condition), or with the maze room lights extinguished (darkness condition). Importantly, each of these manipulations created situations in which reward prediction errors likely occurred. Overall, these three testing conditions created the following situations, respectively: a mismatch between the locations of large and small rewards, a decreased probability of obtaining a reward, and a situation in which rats are not able to discriminate between arms associated with large and small amounts of reward (C.B. Puryear, M.J. Kim, and S.J.Y. Mizumori, in prep.). Therefore, positive prediction errors could occur when the animal received a large amount of reward on an arm previously associated with a small amount,

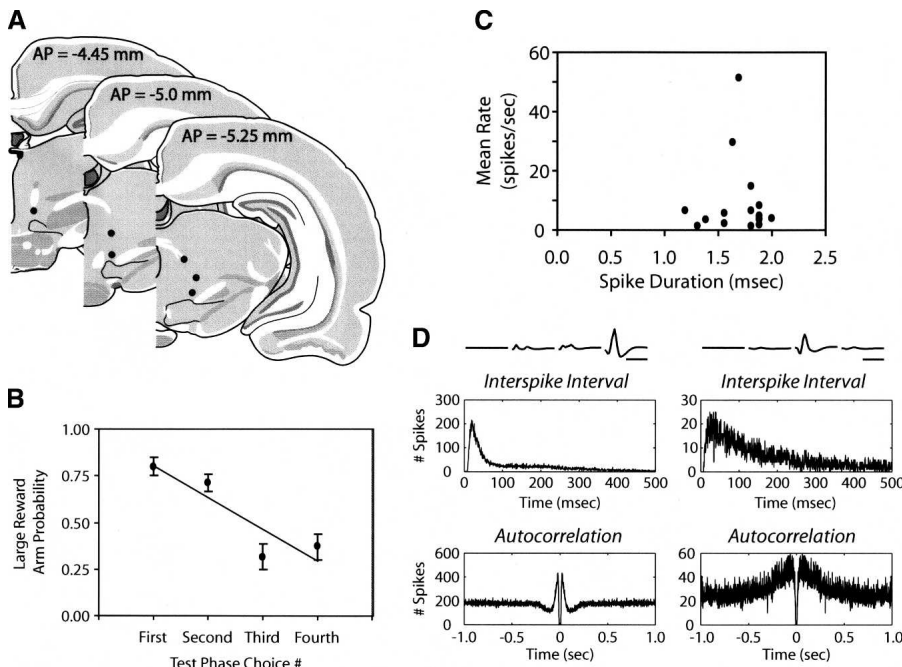


Figure 1. Histology and basic firing properties of MRNm neurons. (A) Distribution of cells localized to MRNm. Each dot may represent the location of more than one neuron. Coronal slices adapted from Swanson (2003) (reprinted with permission from Academic Press ©2003). (B) Rats displayed preference for arms that contained large amounts of reward. Plotted is the average probability of choosing a large reward arm during the first four arm choices of the test phase of each trial. Error bars represent SEM. (C) Distribution of average firing rates and spike duration of MRNm neurons. Most cells fired less than 10 spikes/sec and exhibited waveform durations between 1.5 and 2.0 msec. (D) Examples of two MRNm neurons. *Top row* shows their average waveform on each wire of the tetrode. Scale bar = 1 msec. *Middle and bottom rows* depict their interspike interval and autocorrelation histograms, respectively.

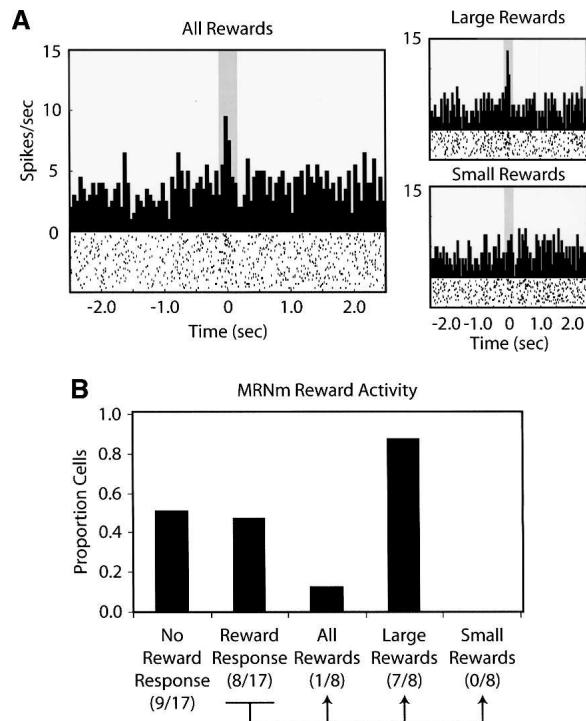


Figure 2. Reward-related activity of MRNm neurons. (A) Peri-event time histograms of one cell that exhibited a short-latency, excitatory response upon acquisition of rewards ($t_{(0)}$, bin width = 50 msec). *Left* histogram shows only a modest excitatory response when considering all rewards together. However, *top* and *bottom right* histograms show that the reward-related firing occurred upon acquisition of large and not small rewards. Gray-shaded areas indicate time periods analyzed for significant increases in firing rate. (B) Population summary of the proportion of MRNm neurons that demonstrated significant reward-related activity.

when the rat retrieved rewards after visiting arms in which reward had been omitted, and when the rat obtained a large reward in darkness. Negative reward prediction errors could occur when the animal received a small amount reward on a maze arm previously associated with a large amount, when the rat visited an arm that did not contain a reward, and when the rat obtained a small reward in darkness.

Eight cells with significant responses to large rewards were recorded during these tests (two during reward location switch, four during reward omission conditions, and two during darkness conditions). In order to determine whether the reward manipulations affected the reward-related activity of these cells, a reward activity value (RA) was first calculated, which was the average firing rate in the ± 150 msec around the time of acquisition of rewards, expressed as a percentage in change relative to the cell's average firing rate for each block of trials. These values were normalized to the maximum RA value observed, yielding a normalized RA value for the first and second block of trials (RA_{n1} and RA_{n2} , respectively). These calculations were made for large and small rewards separately, and for non-rewarded arms in the reward omission condition.

We then created scatterplots of RA_n 's for each block of trials. If reward-related activity was independent of the expectation of the reward received, RA_{n1} and RA_{n2} should be similar in each block of trials. As can be seen in Figure 3A, the reward-related activity was consistently higher when rats received more reward than expected in the second block of trials. Conversely, Figure 3B clearly shows that neural activity was consistently suppressed when rats received less reward than expected. These differences

in reward-related firing were quantified by calculating the distance of each data point to the diagonal (i.e., the reward activity change index, or RACI):

$$RACI = \frac{\sqrt{2(RA_{n1} - RA_{n2})^2}}{2}$$

Directionality of the change in reward activity was taken into account in order to discern between increases and decreases in firing rate. A one-sample t -test ($\alpha = 0.05$) indicated that average RACI values were significantly increased when rats received more reward than expected ($t_{(7)} = -2.73$, $P < 0.03$) and were significantly decreased when rats received less reward than expected ($t_{(7)} = -4.88$, $P < 0.001$). These results are consistent with positive and negative reward prediction error signals, respectively. An example of an MRNm neuron that exhibited both positive and negative prediction error-related activity in the reward location switch condition is depicted in Figure 3D.

We demonstrate here that a large proportion of MRNm neurons may be involved in computations about reward acquisition. Similar to dopamine neurons (Tobler et al. 2005), the majority of reward-related MRNm neurons preferentially fired relative to acquisition of large amounts of reward. To our knowledge, this is the first demonstration of discriminative reward responses of reticular formation neurons and highlights a novel role for the reticular formation in reward value representations. Furthermore, these data suggest that MRNm neurons represent similar reward prediction error signals as dopamine neurons (Nakahara et al. 2004; Bayer and Glimcher 2005; Pan et al. 2005; Tobler et al. 2005). It is important to note that in this initial sample, there was remarkable overall consistency and reliability of the positive and negative reward prediction error signals by MRNm neurons. This is similar to the homogeneity of dopamine neuron responses, suggesting that reward prediction may be a major function of the overall population of MRNm reward-related neurons. Nevertheless, further parametric studies are necessary to determine whether MRNm neural activity conforms to the same basic firing profiles that have been well-described for dopamine neurons (i.e., predictive cues and reward probabilities).

The reticular formation has traditionally been thought to be important for initiating general arousal states. This is in part due to initial reports of changes in unit activity during transitions from sleep to wakefulness (Huttenlocher 1961; Kasamatsu 1970; Manohar et al. 1972). In addition, a more specific role for the reticular formation in attention has been described in primates performing visual discrimination tasks (Pragay et al. 1978; Fabre et al. 1983). This is consistent with reports of sensory neglect following reticular formation lesions (Watson et al. 1974). Together, these foundational data suggest that reticular formation may function to enhance the overall level of arousal and vigilance necessary for attending to and acting upon salient stimuli (Mesulam 1981). Accordingly, changes in reward-related MRNm neuronal activity could provide an important signal indicating that the contingencies of recently executed behaviors have changed.

The striking similarity of the reward prediction error signals of MRNm neurons reported here suggests that MRNm, along with brain regions such as lateral habenula (Matsumoto and Hikosaka 2007), pedunclopontine nucleus (PPTg) (Kobayashi and Okada 2007), and dorsal raphe nucleus (Nakamura et al. 2008) may contribute to the generation of reward prediction error signals observed in dopamine neurons. Furthermore, these data suggest the possibility that such signals are a general property of a large network of midbrain structures. Given that the projections from the reticular formation to VTA are glutamatergic (Geisler et al. 2007), it is possible that the changes in reward-related activity

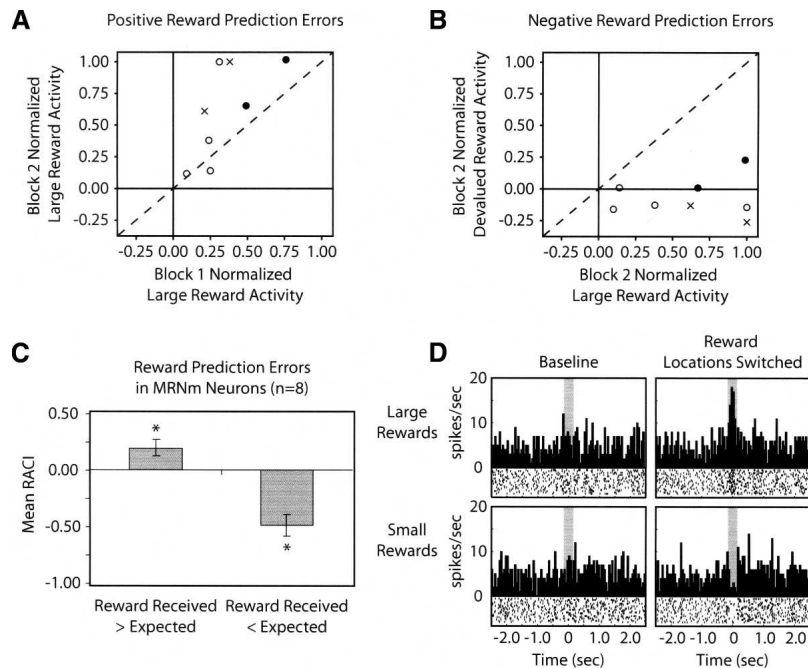


Figure 3. Reward prediction errors in MRNm neurons. (A) Plotted is each neuron's normalized large reward activity (RA_N , defined in text) for each block of trials. The reward activity during block 2 (y-axis) represents activity at times when more reward than expected was obtained. Note that reward-related activity during these times is consistently more robust than during times in which the rat received the expected reward (block 1), indicating that positive reward prediction errors occurred. (B) Plotted are RA_N values for rewarded and devalued arms in block 2 (x- and y-axes, respectively). Devalued arms include arms associated with a large amount of reward but baited with a small amount of reward, arms in which reward was omitted, and arms containing small rewards visited in darkness. Note that reward-related activity on devalued arms is consistently suppressed, indicating that negative reward prediction errors occurred. (Symbols in A,B: ● indicates cell recorded in darkness condition; ○, cell recorded in reward omission condition; and ×, cell recorded in reward location switch condition.) (C) Average changes in reward activity (RACI, defined in text) for times in which the rat obtained more and less reward than expected. Asterisks indicate significant differences ($P < 0.05$). Error bars indicate SEM. (D) An example of one neuron that did not respond to acquisition of rewards during the first block of trials. When the locations of large and small rewards were switched, however, the cell developed an excitatory response to acquisition of large rewards on arms previously associated with small amounts of reward (positive reward prediction error). Furthermore, the firing of the cell was inhibited upon acquisition of small rewards on arms previously associated with large amounts of reward (negative reward prediction error).

of MRNm neurons, in concert with PPTg, could provide an excitatory component of the reward prediction error signal. In combination with inhibitory inputs from lateral habenula, this may then selectively activate dopamine neurons to initiate the coordinated selection of appropriate behaviors in response to changes in reward outcome (Humphries et al. 2007).

Acknowledgment

This research was supported by NIMH Grant MH 58755 to S.J.Y.M.

References

- Bayer, H.M. and Glimcher, P.W. 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**: 129–141.
- Christoph, G.R., Leonzio, R.J., and Wilcox, K.S. 1986. Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *J. Neurosci.* **6**: 613–619.
- Fabre, M., Rolls, E.T., Ashton, J.P., and Williams, G. 1983. Activity of neurons in the ventral tegmental region of the behaving monkey. *Behav. Brain Res.* **9**: 213–235.
- Fields, H.L., Hjelmstad, G.O., Margolis, E.B., and Nicola, S.M. 2007. Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. *Annu. Rev. Neurosci.* **30**: 289–316.

- Floresco, S.B., West, A.R., Ash, B., Moore, H., and Grace, A.A. 2003. Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat. Neurosci.* **6**: 968–973.
- Geisler, S., Derst, C., Veh, R.W., and Zahm, D.S. 2007. Glutamatergic afferents of the ventral tegmental area in the rat. *J. Neurosci.* **27**: 5730–5743.
- Herkenham, M. and Nauta, W.J. 1979. Efferent connections of the habenular nuclei in the rat. *J. Comp. Neurol.* **187**: 19–47.
- Humphries, M.D., Gurney, K., and Prescott, T.J. 2007. Is there a brainstem substrate for action selection? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**: 1627–1639.
- Huttenlocher, P.R. 1961. Evoked and spontaneous activity in single units of medial brain stem during natural sleep and waking. *J. Neurophysiol.* **24**: 451–468.
- Kasamatsu, T. 1970. Maintained and evoked unit activity in the mesencephalic reticular formation of the freely behaving cat. *Exp. Neurol.* **28**: 450–470.
- Kobayashi, Y. and Okada, K. 2007. Reward prediction error computation in the pedunculopontine tegmental nucleus neurons. *Ann. N. Y. Acad. Sci.* **1104**: 310–323.
- Manohar, S., Noda, H., and Adey, W.R. 1972. Behavior of mesencephalic reticular neurons in sleep and wakefulness. *Exp. Neurol.* **34**: 140–157.
- Matsumoto, M. and Hikosaka, O. 2007. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**: 1111–1115.
- McNaughton, B.L., O'Keefe, J., and Barnes, C.A. 1983. The stereotrode: A new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. *J. Neurosci. Methods* **8**: 391–397.
- Mesulam, M.M. 1981. A cortical network for directed attention and unilateral neglect. *Ann. Neurol.* **10**: 309–325.
- Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., and Hikosaka, O. 2004. Dopamine neurons can represent context-dependent prediction error. *Neuron* **41**: 269–280.
- Nakamura, K., Matsumoto, M., and Hikosaka, O. 2008. Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *J. Neurosci.* **28**: 5331–5343.
- Pan, W.X., Schmidt, R., Wickens, J.R., and Hyland, B.I. 2005. Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. *J. Neurosci.* **25**: 6235–6242.
- Pragay, E.B., Mirsky, A.F., Ray, C.L., Turner, D.F., and Mirsky, C.V. 1978. Neuronal activity in the brain stem reticular formation during performance of a "go-no go" visual attention task in the monkey. *Exp. Neurol.* **60**: 83–95.
- Pratt, W.E. and Mizumori, S.J. 2001. Neurons in rat medial prefrontal cortex show anticipatory rate changes to predictable differential rewards in a spatial memory task. *Behav. Brain Res.* **123**: 165–183.
- Puryear, C.B., King, M., and Mizumori, S.J. 2006. Specific changes in hippocampal spatial codes predict spatial working memory performance. *Behav. Brain Res.* **169**: 168–175.
- Schultz, W. and Dickinson, A. 2000. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* **23**: 473–500.
- Swanson, L.W. 2003. *Brain maps: Structure of the rat brain*, 3rd ed. Academic Press, San Diego, CA.
- Tobler, P.N., Fiorillo, C.D., and Schultz, W. 2005. Adaptive coding of reward value by dopamine neurons. *Science* **307**: 1642–1645.
- Watson, R.T., Heilman, K.M., Miller, B.D., and King, F.A. 1974. Neglect after mesencephalic reticular formation lesions. *Neurology* **24**: 294–298.
- Wise, R.A. 2004. Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**: 483–494.

Received May 18, 2008; accepted in revised form October 7, 2008.