c0004

PART IV

UNDERSTANDING VALUATION – LEARNING VALUATION

p0010

CHAPTER

21

c00021

Midbrain Dopamine Neurons: a Retina of the Reward System?

Wolfram Schultz

OUTLINE

s0010

INTRODUCTION

p0010 We can define rewards as objects or events that generate approach and consummatory behavior, produce learning of such behavior, represent positive outcomes of economic decisions, and may engage positive emotions. Rewards are crucial for individual and gene survival and support elementary processes, such as drinking, eating, and reproduction. This behavioral definition attributes reward function also to certain non-alimentary and non-sexual entities, including money, technical artifacts, esthetic stimulus attributes, and mental events. Rewards engage agents in such diverse behaviors as foraging, and trading on stock markets.

p0020 The functions of primary sensory systems are based on the physics and chemistry of specific stimuli impinging on their receptors. By contrast, rewards are often polysensory and do not engage specialized receptors. Reward functions are not defined primarily by the physics and chemistry of their objects, but by the behavioral reactions they induce. In order to use rewards for specific behavioral reactions, the brain needs to extract the reward information from heterogeneous mechanical, electromagnetic, chemical, and other events impinging on our sensory organs.

p0030 In the absence of dedicated reward receptors, it would be helpful for the brain to have an explicit neuronal signal that identifies a reward object and contains information about its motivational or economic value, irrespective of other sensory properties. Such a reward signal could provide an input to neuronal processes underlying approach behavior, learning, and economic decision-making, just as, for example,

GLIMCHER 978-0-12-374176-9 00021

CH021.indd 321 5/14/2008 5:24:27 PM

visual responses in retinal ganglion cells identify specific electromagnetic events and deliver this information to brain centers involved in visual perception and sensorimotor reactions.

p0040 This chapter describes how the search for a "retina of the reward system" has revealed phasic signals in midbrain dopamine neurons that reflect information about the motivational or economic value of rewards regardless of other stimulus components and irrespective of specific behavioral reactions necessary to obtain the rewards.

s0020 CONCEPTS OF REWARD

p0050 In the absence of dedicated reward receptors, the behavioral definition of reward allows us to use behavioral theories for investigating the neural processing of reward. In analogy to David Marr's suggestion for the visual system (Marr, 1982), behavioral theories help us to define and understand the problem to be solved (approach behavior, learning, and decision-making), and neurobiology helps us to understand how the problem is being solved (neural reward mechanisms).

s0030 Reward Value, Utility and Risk

p0060 Rewards are positive outcomes that have specific magnitudes and occur with specific probabilities. Agents make choices between options whose motivational or economic values are determined by the magnitudes and probabilities of outcomes, all other outcome parameters being equal.

p0070 The term "mean (of reward)" seems more appropriate than "expectation (of reward)". Monkeys may not have a notion of mathematical expectation, and instead rely on experienced population or sample means. In addition the term "expectation" is used to denote a psychological process. The term "risk" refers to a form of uncertainty in which the probability distributions are known, whereas "ambiguity" indicates uncertainty with incomplete knowledge of probabilities and is often referred to simply as "uncertainty." Risk in economics refers to the chances of losing and winning, rather than the narrower common sense meaning of the potential for loss.

p0080 Standard expected utility theory assumes that decisionmakers assign subjective utilities to objective outcome values, as measured objectively in choice preferences (Bernoulli, 1738). Preference is expressed as the probability of choosing one option over all others. Utility usually increases non-linearly as a function of value, and utility functions flatten gradually when utility increases progressively less with higher outcome values. Then, the losses from probabilistically occurring rewards appear larger than the gains, the mathematical expectation of the probability distribution of utilities (expected utility) drops with increasing variance, and the decision-maker shows risk aversion. By contrast, assigning progressively more utility to higher outcome values favors variable outcomes and produces risk-seeking. Thus, the concave or convex shape of utility functions determines how much risk-sensitive individuals lose or gain in expected utility with risky outcomes. Only individuals with linear utility functions are risk-neutral. For more details, Chapter 3 of this volume.

Predictions s0040

Evolutionary pressure demands the energy-efficient processing of information. One potential solution is to store predictions about future events at higher brain centers, process the differences between the newly incoming information and the predictions (prediction errors) in lower brain centers, and update the predictions in higher brain centers by forwarding only the less information-containing prediction errors rather than the full information about the state of the external world (Rao and Ballard, 1999). In this way, higher brain centers have access to full information about the external world for perceptions, decisions, and behavioral reactions at a much lower cost. This fundamental role of predictions would lead to two processes, namely learning, as defined by changes in behavior derived from updated predictions, and informed decision-making through advance information about the available choice options (as opposed to guesses, when outcomes are unknown). As choice options are best described by probability distributions rather than singular values, reward predictions also comprise distributions, and specify the key parameters of mean or expected value, expected utility, and the corresponding SDs or variances. p0090

Prediction Errors and Learning s0050

"Prediction error" refers to the difference between the actual outcome and its prediction. Animal learning theory and efficient temporal difference (TD) reinforcement models postulate a crucial role for outcome prediction errors in Pavlovian and operant conditioning (Rescorla and Wagner, 1972; Sutton and Barto, 1981). Current views attribute Pavlovian learning to any acquisition of predictions, irrespective of vegetative p0100

GLIMCHER 978-0-12-374176-9 00021

CH021.indd 322 5/14/2008 5:24:27 PM

reactions or striated muscle contractions, as long as the outcome is not conditional on a behavioral reaction. The crucial role of prediction errors is based on Kamin's blocking effect (1969), which demonstrates that learning and extinction advance only to the extent to which a reinforcer is better or worse than predicted, and slow as the prediction approaches the reinforcer.

s0060 DOPAMINE REWARD-RELATED RESPONSES

s0070 Pure Reward Signals

p0110 Pure reward neurons would carry a signal with abstract information about the objective economic reward value or utility that is useful for eliciting the known behavioral effects of rewards (approach behavior, learning, decision-making, and positive emotion). Such a pure reward signal should not be specific for the sensory properties of the reward, and thus not distinguish between individual reward objects; nor should it reflect the specifics of the behavioral reaction necessary to obtain the reward. However, some reward-related neurons do carry such additional information – for example, action-value neurons that combine specific movement information with specific reward information.

s0080 Dopamine Response to Reward

p0120 The majority of midbrain dopamine neurons (75–80%) show rather homogeneous, phasic activations with latencies of <100 ms and durations of <200 ms to temporally unpredicted food and liquid rewards. The remaining neurons are unresponsive to all stimuli tested. The responses increase monotonically with reward magnitude, such as liquid volume (Tobler *et al*. 2005), apparently coding reward value or utility. Although further tests are necessary, the dopamine reward responses seem to constitute pure reward signals.

s0090 *Prediction-error Coding*

p0130 The dopamine response to the reward appears to code a prediction error, such that a reward that is better than predicted elicits an activation (positive prediction error), a fully predicted reward draws no response, and a reward that is worse than predicted induces a depression (negative error). Thus, the dopamine response implements fully the crucial term of the Rescorla-Wagner learning model.

The error response varies quantitatively with the mean reward, irrespective of specific combinations of reward magnitude and probability (Fiorillo *et al*., 2003; Satoh *et al*., 2003; Morris *et al*., 2004; Nakahara *et al*., 2004; Bayer and Glimcher, 2005; Tobler *et al*., 2005). The prediction-error response is sensitive to the time of the reward, as a delayed reward induces a depression at its original time and an activation at its new time. Whereas the error coding is evident for activations reflecting positive prediction errors, it is more difficult to assess for the negative-error related depressions with their narrow dynamic range unless the full period of depression is considered (Bayer *et al*., 2007). p0140

Stringent Tests for Prediction Error s0100

The coding of reward prediction errors should be tested in stringent paradigms developed by animal learning theory. In the blocking test (Kamin, 1969), a stimulus that is paired with a fully predicted reward does not become a valid reward predictor. The absence of a reward following the blocked stimulus does not produce a prediction error, or a response in dopamine neurons, even after extensive stimulus–reward pairings (Waelti *et al*., 2001). By contrast, the delivery of a reward after a blocked stimulus constitutes a positive prediction error and, accordingly, elicits a dopamine activation. p0150

In the conditioned inhibition paradigm, a test stimulus is presented simultaneously with an established reward-predicting stimulus, but no reward is given after the compound, making the test stimulus a predictor for the absence of reward. Reward omission after such a conditioned inhibitor does not produce a negative prediction error or a depression in dopamine neurons (Tobler *et al*., 2003). By contrast, delivery of a reward after the inhibitor produces a strong positive prediction error and, accordingly, a strong dopamine activation. p0160

The results from these formal paradigms confirm that dopamine neurons indeed show bidirectional coding of reward-prediction errors. p0170

Adaptive Coding s0110

The dopamine response does not simply reflect the prediction error of the mean reward, but also adapts to the predicted reward distribution within 2 seconds. When visual stimuli predict specific binary distributions of equiprobable reward magnitudes ($P = 0.5$ each), delivery of reward with the larger of the two magnitudes within each distribution elicits the same dopamine activation, despite 10-fold differences in p0180 reward magnitude and prediction error (Tobler *et al*., 2005). Experimental separation of mean and standard deviation (SD) suggests that the dopamine response codes the prediction error divided by, or normalized to, the SD, reflecting how much the delivered reward differs from its prediction in units of SD rather than mean. Theoretical considerations suggest that error teaching signals that are scaled by risk rather than mean can mediate stability in learning that is resistant to the predicted risk of outcomes (Preuschoff and Bossaerts, 2007; see also Chapter 23 of this volume).

p0190 Adaptive shifts in responses to predicted outcome distributions also occur in the orbitofrontal cortex and striatum of monkeys (Tremblay and Schultz, 1999; Cromwell *et al*., 2005; Hosokawa et al., 2007) and humans (Breiter *et al*., 2001; Nieuwenhuis *et al*., 2005). However, shifts are not observed in all orbitofrontal reward neurons, nor do they occur when there is insufficient time for adaptation (Padoa-Schioppa and Assad, 2008). Future experiments might search for a potential common mechanism across reward structures that may match the probability distributions of neuronal responses to the probability distributions of rewards.

s0120 Dopamine Response to Reward-predicting Stimuli

p0200 Dopamine neurons are activated by reward-predicting visual, auditory, and somatosensory stimuli (for review, see Schultz, 1998). These responses occur irrespectively of the sensory modalities and spatial positions of stimuli, and of arm, mouth, and eye movements being executed. The responses increase with the means of predicted reward probability distributions, without distinguishing between magnitude and probability (Tobler *et al*., 2005). Response magnitude increases with decreasing reaction time, and thus may incorporate the animal's motivation (Satoh *et al*., 2003). In choices between different reward values or delays, the responses to the presentation of choice options reflect the animal's future chosen reward (Morris *et al*., 2006) or the highest reward across options (Roesch *et al*., 2007).

p0210 During the course of learning, the dopamine response to the reward decreases gradually, and a response to the reward-predicting stimulus develops (Mirenowicz and Schultz, unpublished; Takikawa *et al*., 2004; Pan *et al*., 2005). The acquisition of conditioned responding is sensitive to blocking, indicating that prediction errors play a role in the learning of dopamine neurons. The response transfer to reward-predicting stimuli complies with the principal characteristics of teaching signals of efficient reinforcement models (Sutton and Barto, 1981). The response shift does not involve the back-propagation of prediction errors of earlier TD models (Montague *et al*., 1996), and is reproduced in recent TD implementations (Suri and Schultz, 1999; Pan *et al*., 2005).

p0220 Physically intense stimuli with attention-inducing and possibly rewarding functions induce activations in dopamine neurons. These responses are enhanced by stimulus novelty (Ljungberg *et al*., 1992; Horvitz *et al*., 1997), but disappear quickly with aversive conditioning (Mirenowicz and Schultz, unpublished). However, attention-inducing reward omission induces only depressions. Other strong attentional stimuli, such as punishers and conditioned inhibitors, induce predominantly depressions and only brief, non-differential, initial activation components, and rarely full neuronal activations (Mirenowicz and Schultz, 1996; Schultz *et al*., 1997; Tobler *et al*., 2003). Some of these non-reward-related activations may be due to response generalization. Responsiveness to non-reward-predicting stimuli increases in contexts with higher proportions of stimuli predicting reward as opposed to non-reward (Schultz and Romo, 1990; Waelti *et al*., 2001; Tobler *et al*., 2003).

p0230 Taken together, these characteristics may qualify the main dopamine activations following reward-predicting stimuli as pure reward signals, with the possible exception of an additional motivational component. The responses to non-reward-predicting stimuli also occur irrespectively of sensory modalities, including responses to intense stimuli, initial non-differential response components to reward-predicting stimuli, and generalized responses; their functions may be related to the positively motivating attributes of intense, novel, and reward-resembling stimuli, but do not seem to reflect general arousing or attentional functions irrespective of motivational valence.

Dopamine Concentrations s0130

p0240 Fast-scan cyclic voltammetry detects changes in dopamine concentration in dopamine projection areas downsteam of impulse activity of dopamine cell bodies, at time-courses somewhat comparable with electrophysiology. Dopamine concentrations increase after unpredicted food rewards, shift to reward-predicting stimuli during conditioning, and show a propensity for generalization to non-reward-predicting stimuli in ventral striatum (Day *et al*., 2007). These data replicate some of the essential reward responses seen with electrophysiology, and suggest that the impulse responses indeed lead to dopamine release.

s0140 Reward Signals in Other Brain Structures

p0250 Among rather heterogeneous neuronal populations, a fraction of neurons in orbitofrontal cortex and striatum are activated in relation to rewards and reward-predicting stimuli (Thorpe *et al.*, 1983; Ravel *et al.*, 2003). A subset of orbitofrontal reward neurons show graded responses with reward magnitude (Wallis and Miller, 2003), and some of them carry pure reward signals by reflecting the economic value of reward objects determined by choice preferences, without distinguishing between different rewards ("chosen value" neurons of Padoa-Schioppa and Assad, 2006). Most orbitofrontal and some striatal reward neurons do not discriminate between spatial positions or visual features of reward-related stimuli (Tremblay and Schultz, 1999; Hassani *et al.*, 2001).

p0260 However, other orbitofrontal and striatal reward neurons do not seem to carry pure reward signals, as they distinguish between different reward objects with same economic values ("offer value" neurons of Padoa-Schioppa and Assad, 2006) or are in addition sensitive to visual object features. Action-value neurons encode both reward and movement parameters, and are often viewed as parts of goal-directed mechanisms by which reward information labels specific movements with specific economic value, and thus may determine decisions between differently valued actions. Such neurons are found in the striatum (Hollerman *et al.*, 1998; Kawagoe *et al.*, 1998; Samejima *et al.*, 2005), and constitute the typical reward neurons found in prefrontal and parietal cortex (Watanabe, 1996; Platt and Glimcher, 1999; Sugrue *et al.*, 2004).

s0150 DOPAMINE REWARD RISK SIGNAL

p0270 If a reward signal reflects the mean reward scaled by the standard deviation (SD) of reward probability distributions, could there be a direct neuronal signal for risk as measured by SD? When varying reward probabilities from 0 to 1 while keeping magnitude constant, the mean increases monotonically with probability, whereas risk (SD, variance) and entropy show an inverted U function peaking at $P = 0.5$. More than one-third of dopamine neurons show a relatively slow, moderate activation that increases gradually in time between the reward-predicting stimulus and the reward, and varies monotonically with risk (Fiorillo *et al.*, 2003). The activation occurs in individual trials, and does not seem to constitute a prediction-error response propagating back from reward to the reward-predicting stimulus. The activation also increases monotonically with SD or variance when binary distributions of different equiprobable, non-zero reward magnitudes are used, entropy being constant at 1 bit. Thus, SD and variance appear to be viable measures for risk as coded by dopamine neurons. Risk-related activations have slower time-courses and lower peaks compared with the responses to reward-predicting stimuli and prediction errors at different behavioral periods.

A dopamine risk signal may have three functions. p0280
First, it might influence the scaling of responses by SD in dopamine and other reward neurons. Second, it might provide an input to brain structures dealing with the assessment of reward risk *per se*. Third, it might combine with a mean economic value signal to represent information about the expected utility in risk-sensitive individuals, according to the mean-variance concept in financial decision theory (Levy and Markowitz, 1958).

DOPAMINE CONFUSIONS s0160

Neurotransmitters can influence postsynaptic neu- p0290
rons through phasic signals transmitting time-specific information for specific behavioral processes, and through tonic, permissive actions, without much variation over time. Dopamine seems to have both actions (for review, see Schultz, 2007).

The phasic, informational signals of dopamine p0300
neurons relate to reward and risk, and are compatible with subsecond time-courses of rapid behavioral processes. They produce subsecond rises of dopamine concentrations well above 100 nM. These signals are suitable for assigning reward credit to stimuli and actions during learning by enabling synaptic modifications, and they may inform rapid neuronal decision mechanisms. At mostly slower timescales of seconds, a negative dopamine signal reports aversive events, and voltammetrically measured dopamine is released with movements. Microdialysis reveals moderate dopamine fluctuations over several minutes with punishment and stress, but these are much slower than in most learning and decision-making mechanisms.

In addition to its informational function, dopamine p0310
appears to play a permissive role devoid of specific, behavioral information that is required for a large variety of motor, cognitive, motivational, and learning processes to function properly. These processes become deficient when dopamine transmission is impaired by lesions or receptor antagonists, and recover partly after pharmacological stimulation of dopamine receptors without the phasic informational signals being restored. This role involves sustained,

GLIMCHER 978-0-12-374176-9 00021
CH021.indd 325 5/14/2008 5:24:27 PM

finely controlled, and only slowly varying extracellular dopamine concentrations in the order of 5–20 nM that exert a necessary, sustained influence on postsynaptic striatal and cortical receptors, membranes, and secondary messenger systems processing a large variety of specific, informational, behavior-related signals. The permissive role would allow the limited number of dopamine neurons to be involved in a large variety of functions residing in postsynaptic structures like striatum and cortex, without providing them with phasic input signals containing specific information.

p0320 Confusion will arise when confounding the different informational and permissive modes of dopamine action and trying to identify a single behavioral dopamine function to the exclusion of other functions, as we usually do for primary sensory systems. For example, strong dopamine reductions in the striatum lead to Parkinsonian movement deficits, but dopamine neurons do not encode movement parameters. Apparently, dopamine has a permissive effect on movement processes residing in the striatum without carrying its own movement-related information, and Parkinsonism may reflect a disorder of dopamine-dependent, movement-related activity in striatal neurons rather than a deficit in informational coding of dopamine neurons. Likewise, impaired inputs from dopamine neurons to the ventral striatum induce deficiencies in incentive motivational salience, but dopamine neurons do not carry a phasic incentive salience signal, as they are activated very little by punishers, which are motivationally very salient. Apparently, the deficits in incentive salience are due to a disfunctional ventral striatum following impairment of the permissive rather than informational dopamine action. Thus, dopamine is involved in movement and incentive salience, but this involvement reflects the permissive rather than the informational reward- and risk-related dopamine function. Postsynaptic mechanisms would be able to discriminate between the permissive and informational dopamine functions on the basis of their time-courses and instantaneous dopamine concentrations.

s0170 CONCLUSIONS

p0330 Just as the retina transfers information from electromagnetic waves to action potentials, dopamine neurons transfer information from polysensory stimuli to action potentials signaling rewards. The reward signal informs neural decision-making mechanisms about the abstract motivational or economic reward value, or possibly utility, as the key decision variable for economic choices. Neurons in other reward structures carry additional reward information, including reference to drive states and sensory details about individual reward objects. However, the economic reward value signal of dopamine neurons codes neither simply reward nor mean economic value. Rather, it encodes reward outcome as it deviates from prediction, thus fulfilling the basic requirement for a prediction-error teaching signal postulated by learning theory, and it is scaled in units of standard deviation rather than mean. A separate, slower dopamine signal informs directly about the degree of risk. Whereas the dopamine reward and risk signals are fast enough for immediate behavioral reactions, dopamine also operates with slower time-courses in a much wider spectrum of behavior, often in a permissive role without much temporal and behavioral information.

Acknowledgments s0180

Our work was supported by the Wellcome Trust, the Swiss National Science Foundation, the Human Frontiers Science Program, and several other grant and fellowship agencies. p0340

References

Bayer, H.M. and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Bayer, H.M., Lau, B., and Glimcher, P.W. (2007). Statistics of dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. Comentarii Academiae Scientiarum Imperialis Petropolitanae. *Papers Imp. Acad. Sci. St Petersburg* 5, 175–192, Translated as: (1954) Exposition of a new theory on the measurement of risk. *Econometrica* 22, 23–36.

Breiter, H.C., Aharon, I., Kahneman, D. *et al*. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Cromwell, H.C., Hassani, O.K., and Schultz, W. (2005). Relative reward processing in primate striatum. *Exp. Brain Res.* 162, 520–525.

Day, J.J., Roitman, M.F., Wightman, R.M., and Carelli, R.M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* 10, 1020–1028.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Hassani, O.K., Cromwell, H.C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.

Hollerman, J.R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* 80, 947–963.

Horvitz, J.C., Stewart, T., and Jacobs, B.L. (1997). Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* 759, 251–258.
Hosokawa, T., Kato, K., Inoue, M., and Mikami, A. (2007). Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. *Neurosci. Res.* 57, 434–445.
Kamin, L.J. (1969). Selective association and conditioning. In: N.J. Mackintosh and W.K. Honig (eds), *Fundamental Issues in Instrumental Learning*. Dalhousie, NB: Dalhousie University Press, pp. 42–64.
Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.
Levy, H. and Markowitz, H.M. (1979). Approximating expected utility by a function of mean and variance. *Am. Econ. Rev.* 69, 308–317.
Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.
Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.
Mirenowicz, J. and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379, 449–451.
Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.
Morris, G., Arkadir, D., Nevet, A. *et al.* (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43, 133–143.
Morris, G., Nevet, A., Arkadir, D. *et al*. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.
Nakahara, H., Itoh, H., Kawagoe, R. *et al*. (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269–280.
Nieuwenhuis, S., Heslenfeld, D.J., Alting van Geusau, N. *et al.* (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage* 25, 1302–1309.
Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
Padoa-Schioppa, C. and Assad, J.A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* 11, 95–102.
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.
Platt, M.L. and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.
Preuschoff, K. and Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Ann. NY Acad. Sci.* 1104, 135–146.
Rao, R.P.N. and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87.
Ravel, S., Legallet, E., and Apicella, P. (2003). Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. *J. Neurosci.* 23, 8489–8497.
Rescorla, R.A. and Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: A.H. Black and W.F. Prokasy (eds), *Classical Conditioning II: Current Research and Theory*. New York, NY: Appleton Century Crofts, pp. 64–99.
Roesch, M.R., Calu, D.J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.
Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.
Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* 23, 9913–9923.
Schultz, W. (2007). Multiple dopamine functions at different time courses. *Ann. Rev. Neurosci.* 30, 259–288.
Schultz, W. and Romo, R. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.* 63, 607–624.
Schultz, W., Dayan, P., and Montague, R.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
Suri, R. and Schultz, W. (1999). A neural network with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* 91, 871–890.
Sutton, R.S. and Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88, 135–170.
Takikawa, Y., Kawagoe, R., and Hikosaka, O. (2004). A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward maping. *J. Neurophysiol.* 92, 2520–2529.
Thorpe, S.J., Rolls, E.T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.
Tobler, P.N., Dickinson, A., and Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* 23, 10402–10410.
Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.
Tremblay, L. and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708.
Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48.
Wallis, J.D. and Miller, E.K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* 18, 2069–2081.
Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.