Neuron
Perspective

Potential Vulnerabilities of Neuronal Reward, Risk, and Decision Mechanisms to Addictive Drugs

Wolfram Schultz[1,*]
[1]Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK
*Correspondence: ws234@cam.ac.uk
DOI 10.1016/j.neuron.2011.02.014

How do addictive drugs hijack the brain's reward system? This review speculates how normal, physiological reward processes may be affected by addictive drugs. Addictive drugs affect acute responses and plasticity in dopamine neurons and postsynaptic structures. These effects reduce reward discrimination, increase the effects of reward prediction error signals, and enhance neuronal responses to reward-predicting stimuli, which may contribute to compulsion. Addictive drugs steepen neuronal temporal reward discounting and create temporal myopia that impairs the control of drug taking. Tonically enhanced dopamine levels may disturb working memory mechanisms necessary for assessing background rewards and thus may generate inaccurate neuronal reward predictions. Drug-induced working memory deficits may impair neuronal risk signaling, promote risky behaviors, and facilitate preaddictive drug use. Malfunctioning adaptive reward coding may lead to overvaluation of drug rewards. Many of these malfunctions may result in inadequate neuronal decision mechanisms and lead to choices biased toward drug rewards.

Introduction

Drugs of addiction have two principal, closely related functions. First, they constitute rewards, as they induce learning, approach behavior, emotions, and positive feelings, just as natural rewards do. Second, they modify the brain's physiological reward system. However, their influence on the brain is not constrained by physiological receptors and many regulatory mechanisms engaged by natural rewards. Thus addiction constitutes primarily a disorder of the reward system.

The most straightforward influence of addictive drugs on the brain occurs on the dopamine system. Addictive drugs change the phasic characteristics of dopamine activity in reward signaling and the tonic function of dopamine levels in permitting and facilitating a large variety of motor and cognitive functions (see Schultz, 2007, for review). However, the dopamine system does not function in isolation; therefore addictive drugs influence all major reward systems including the striatum, orbitofrontal cortex, and amygdala. Addictive drugs are likely to disturb many of their functions, which may contribute to the drugs' behavioral effects and the addiction process itself.

This review describes a number of normal, physiological reward processes and speculates how they may be affected by addictive drugs. This is not an attempt to explain drug addiction, nor does it describe its pathophysiology, which is complex and beyond the present scope. Rather, the review assesses to which extent drugs may modify well-characterized, normal, physiological, neuronal reward processing. The description remains largely at the systems neuroscience level of reward function without going into details of cellular and synaptic functions. Although many effects of addictive drugs on the dopamine system are known, there is limited firm understanding about drug influence on neuronal processing of reward and cognition. We will use the existing data to make hypotheses about crucial dysfunctions of and beyond the dopamine system that might explain some of the behavioral alterations and possibly shed some light on the addiction process itself. The presented phenomena are primarily neurobiological and should be sufficiently general to apply to many current concepts of behavioral reward functions including wanting versus liking, habits versus goal-directed behavior, and hedonia versus decision-making that are elaborated in current psychological and economic addiction theories. As the behavioral neurophysiology of reward processing is usually restricted to learning, approach behavior, and decision-making, other behavioral components of drug addiction such as urges, cravings, and withdrawal will not be addressed.

Actions of Addictive Drugs on Dopamine Neurotransmission

The dopamine system constitutes the primary target of addictive drugs (Wise and Bozarth, 1987; Wise, 2002). The drugs affect all stages of phasic and tonic dopamine processes, from the generation of action potentials in dopamine cell bodies to the effects of dopamine on postsynaptic neurons. While focusing on striatal dopamine, similar mechanisms apply to cortical and amygdalar dopamine innervation.

Plasticity of Dopamine Neurons

Major drugs of addiction such as cocaine, amphetamine, morphine, heroine, nicotine, and ethanol act on glutamatergic synapses on midbrain dopamine neurons and lead to NMDA-dependent, AMPA-mediated long-term potentiation in dopamine neurons (Figure 1A) (Ungless et al., 2001; Saal et al., 2003). Thus excitatory influences on these neurons become enhanced (Figure 1B), in particular NMDA-dependent burst firing (Zweifel et al., 2009). Amphetamine leads also to reduction of long-term depression in dopamine neurons (Figure 1C) (Jones et al., 2000). Thus subthreshold fluctuations of excitatory

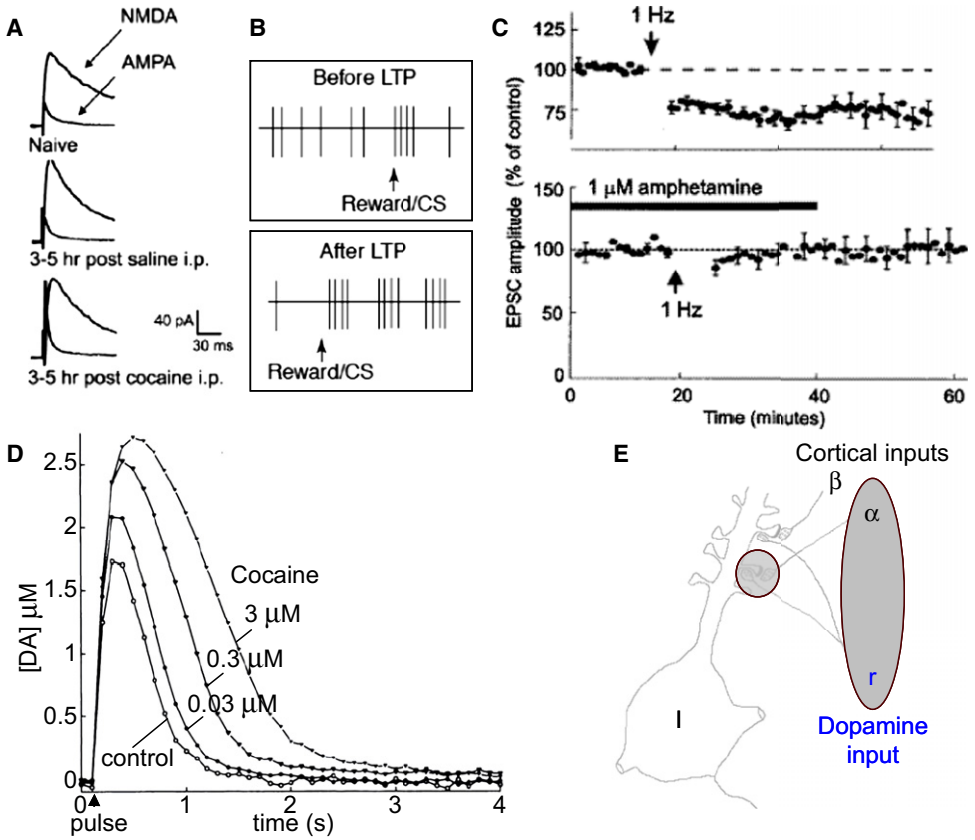

Figure 1. Actions of Addictive Drugs on Dopamine Processes

(A) Long-term potentiation in ventral tegmental dopamine neurons in vitro induced by cocaine. Note the increase in AMPA excitatory postsynaptic current (EPSC) following systemic cocaine (bottom). From Argilli et al. (2008), with permission by Society for Neuroscience.

(B) Scheme of increased burst responses of midbrain dopamine neurons following cocaine-induced long-term potentiation. From Jones and Bonci (2005), with permission by Elsevier Ltd.

(C) Blockade of long-term depression in ventral tegmental dopamine neurons in vitro by bath application of amphetamine. From Jones et al. (2000), with permission by Society for Neuroscience.

(D) Dose-dependent enhancement by cocaine of dopamine efflux induced by single electrical pulse in striatum slices. From Jones et al. (1995), with permission by the American Society for Pharmacology and Experimental Therapeutics.

(E) Classical scheme of differential influence of global dopamine reinforcement signal on selectively active corticostriatal neurotransmission. The dopamine reinforcement signal (r) modifies conjointly active Hebbian synapses (α) at striatal neuron (I) but leaves inactive synapses (β) unchanged. There are about 10,000 cortical terminals and 1000 dopamine varicosities on each striatal neuron (Doucet et al., 1986; Groves et al., 1995. Drawing is based on data from Freund et al. (1984) and Smith and Bolam (1990).

inputs to dopamine neurons would increase or even generate action potentials in the absence of reward, generating a false reward signal.

Phasic Striatal Dopamine Changes

Electrical stimulation of dopamine neurons mimicking natural dopamine responses to reward induces striatal dopamine release. Systemically administered cocaine or amphetamine enhances the stimulus-induced dopamine increase by blocking the reuptake transporter (Figure 1D) (Jones et al., 1995; Gonon, 1997; Venton et al., 2003). Thus, cocaine further exaggerates the striatal dopamine changes following an excitatory dopamine signal that is already enhanced, or falsely generated, by synaptic plasticity at dopamine input synapses.

Tonic Striatal Dopamine Levels

Addictive drugs increase the tonic concentration of striatal dopamine by various mechanisms (Di Chiara and Imperato, 1988). Opiates and nicotine enhance the tonic firing of midbrain dopamine neurons (morphine: Matthews and German 1984; Johnson and North, 1992; nicotine: Lichtensteiger et al., 1982; Pidoplichko et al., 1997), although depressant effects may occur (Bonci and Malenka, 1999). Cocaine and amphetamine block reuptake transport, which enhances tonic dopamine concentrations despite suppressed dopamine neuron firing (Bunney et al., 1973). It is well known that tonic dopamine concentrations are finely regulated within postsynaptic brain areas (Chesselet, 1984). Deviations from these optimal levels, including increases of local tonic dopamine concentrations and dopamine turnover, lead to impaired striatal and cortical mechanisms underlying working memory, sensory discrimination, and planning (Murphy et al., 1996; Elliott et al., 1997; Liu et al., 2008), which may underlie the working memory deficits seen in drug abusers (Ornstein et al., 2000). Thus through their effects on tonic dopamine levels, drugs may impair striatal and cortical mechanisms beyond reward processing.

Dopamine-Dependent Transient Enhancement of Striatal Responses

Varicosities of dopamine axons impinge on dendritic spines of striatal and cortical neurons that are contacted by cortical afferents in a triad arrangement (Figure 1E) (Freund et al., 1984; Goldman-Rakic et al., 1989). This arrangement allows dopamine neurotransmission to affect postsynaptic processing. Dopamine D1 receptor activation enhances striatal postsynaptic depolarizations (Hernández-López et al., 1997). Thus increased or false dopamine signals following addictive drugs may exert a facilitatory effect on excitatory responses in the striatum, and possibly the cortex. The enhancement via D1 receptors may help the transition from drug use to addiction, whereas D2 or D3 receptors may be involved in the expression of addiction (Capriles et al., 2003; Kalivas and Volkow, 2005). Responses in subpopulations of striatal and cortical neurons reflect reward prediction and movement initiation (Hikosaka et al., 1989; Schultz and Romo, 1992; Watanabe, 1996; Hassani et al., 2001; Padoa-Schioppa and Assad, 2006). Thus, by enhancing responses to drug-predicting stimuli, and possibly reducing responses to natural rewards (Kalivas and Volkow, 2005), addictive drugs could prioritize reward prediction and movement initiation.

Dopamine-Dependent Striatal and Cortical Plasticity

The classic triad arrangement of dopamine varicosities, dendritic spines, and cortical inputs (Figure 1E) allows dopamine to enhance spike-time-dependent plasticity at active cortico-striatal and cortico-cortical synapses (Gurden et al., 2000; Reynolds et al., 2001; Wang et al., 2006; Izhikevich, 2007; Pawlak and Kerr, 2008; Shen et al., 2008). The induced long-term potentiation and long-term depression are candidate mechanisms for phasic dopamine signals to mediate behavioral learning.

Thus, by affecting striatal and cortical plasticity, addictive drugs could lead to long-lasting changes of the motor, reward, and cognitive functions of these structures.

Dopamine Reward Signals

Dopamine systems are involved in drug addiction in two important ways. First, data from neuronal stimulation, lesioning, neuropharmacology, and neurophysiology show that the dopamine system is a principal component of the brain's reward system. Its activation induces learning and approach behavior. Second, drugs of addiction constitute rewards and crucially affect neurotransmission in dopamine and associated brain systems. In bypassing sensory receptors and their activity-limiting mechanisms, they induce unphysiological dopamine stimulation, which may lead to addiction.

Prediction Error and Learning

Normal mechanisms. According to reinforcement learning theory, prediction errors act to increase or decrease the prediction of outcome value and thus mediate learning (Rescorla and Wagner, 1972; Sutton and Barto, 1981). In its most simple expression, the reward prediction error δ captures the discrepancy between received reward R and prediction V in trial t:

$$\delta(t) = R(t) - V(t),$$

and serves to update the predicted reward value V in the next trial:

$$V(t+1) = V(t) + k\delta(t),$$

with k as learning rate.

Most dopamine neurons are activated by rewards and reward-predicting stimuli. The reward response codes a prediction error; a reward that is better than predicted elicits an activation (positive prediction error, $R > V$), a fully predicted reward ($R = V$) draws no response, and a reward that is worse than predicted induces a depression (negative error, $R < V$) (Figure 2A) (Schultz et al., 1997). The response implements the teaching term of efficient reinforcement learning models (Rescorla and Wagner, 1972; Sutton and Barto, 1981) and occurs during learning (Schultz et al., 1993; Hollerman and Schultz, 1998). Stimuli not associated with prediction errors are blocked from behavioral and neuronal learning (Waelti et al., 2001).

Potential vulnerabilities. A recent modeling study hypothesized that drug addiction may result from associative learning; the drug-induced higher dopamine levels would mimic a positive dopamine prediction error signal irrespective of any true error and affect striatal plasticity (Redish, 2004). The present hypothesis is based on the mostly phasic effects of addictive drugs on phasic dopamine signals and their consequences on neuronal plasticity. While there is considerable debate about a causal role of dopamine in driving learning, the following descriptions assume that the prediction error contributes, at least to some extent, to behavioral learning. Although the electrophysiological characterization of prediction error coding cannot solve the issue, the effects of electrical and optogenetic dopamine stimulation on learning (Corbett and Wise, 1980; Tsai et al., 2009) suggest a dopamine contribution to several forms of learning.

The long-term potentiation of excitatory inputs to dopamine neurons by addictive drugs (Figure 1A) would increase the phasic activation of dopamine neurons by positive prediction errors ($R > V$) (Figure 1B) and the resulting impulse-dependent striatal dopamine release. Psychostimulants such as cocaine would further enhance this phasic striatal dopamine increase (Figure 1C). Compatible with this reasoning, even a single administration of nonaddictive dopamine receptor agonists such as L-Dopa leads to enhanced positive prediction error signals in humans (Pessiglione et al., 2006). By contrast, the reduction of striatal dopamine following negative prediction errors ($R < V$) would be blunted by the tonic dopamine increase induced by all major drugs of addiction. Finally, there would be no phasic striatal dopamine change in the absence of prediction errors ($R = V$), as the drug-induced tonic elevation of dopamine levels would not adequately mimic a phasic, positive prediction error signal. Indeed, tonic dopamine increases affect learning much less than phasic changes (e.g., Grace, 1991; Tsai et al., 2009). The enhanced influence of positive prediction errors and the blunted effect of negative prediction errors would lead to monotonically increasing striatal plasticity. Thus, the value prediction V would fail to asymptote and ultimately exceed the value of the primary reward R. Although saturation and self-regulatory mechanisms would ultimately limit the process, the resulting extreme induction of reward prediction could overstimulate approach behavior to the level of compulsion. Conditioned inhibition may

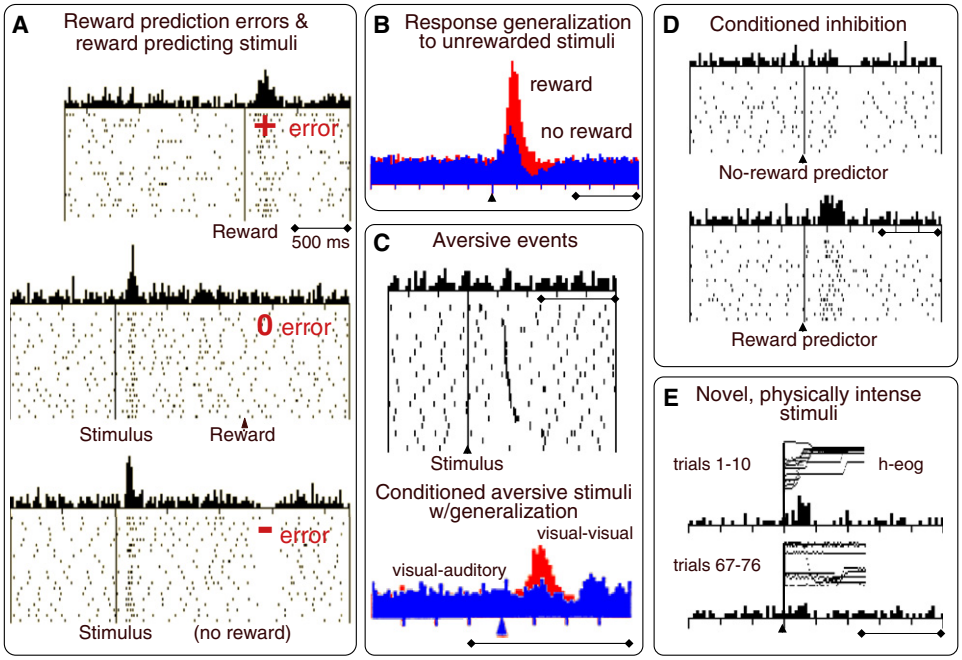

Figure 2. Phasic Increases of Neurophysiological Dopamine Activity following Real and Potential Rewards

(A) Activations following reward prediction errors occurring in 75%–90% of dopamine neurons (right) and reward-predicting stimuli (60%–75% of neurons; left). Data from Schultz (1998).

(B) Generalization of phasic activating population response from reward predicting stimulus (gray) to explicit no-reward predicting stimulus (black; conditioned inhibitor). The shorter and smaller activations to the conditioned inhibitor are partly offset by depressions. Similar generalizing activations are seen with unrewarded stimuli in up to 50% of dopamine neurons (Waelti et al., 2001; Tobler et al., 2003). Data from Tobler et al. (2003), with permission by Society for Neuroscience.

(C) Activations to aversive events. Top: Typical lack of activation to visual stimulus during active avoidance. Primary aversive events activate only about 15% of dopamine neurons. Aversive responses do not code prediction errors (Matsumoto and Hikosaka, 2009). Bottom: activating population response to conditioned aversive stimulus due to stimulus generalization in 65% of dopamine neurons (gray: aversive and appetitive stimuli are both visual), but lack of response when appetitive stimulus is auditory (black). Data from Mirenowicz and Schultz (1996), with permission by MacMillan.

(D) Depressant response to conditioned inhibitor (top) in neuron showing activating response to conditioned excitor (bottom). Data from Tobler et al. (2003), with permission by Society for Neuroscience.

(E) Phasic activations following novel, physically intense stimuli (about 80%). Activations by physically intense stimuli are substantially reduced when pseudoconditioning is controlled for (S. Kobayashi and W. Schultz, 2010, Soc. Neurosci., abstract). Overlapping h-eog traces show horizontal eye movements toward the novel stimulus; unfocused h-eogs after >60 trials indicate familiarity, accompanied by loss of dopamine responses. Data from Ljungberg et al. (1992), with permission by American Physiological Society.

not occur, as the tonic dopamine increase would blunt any neuronal depression, further contributing to compulsion.

The reduction of long-term depression by amphetamine, and possibly other psychostimulants such as cocaine, would facilitate false phasic activations of dopamine neurons by random excitations. These activations would occur without any reward and be independent of actual prediction errors. Postsynaptic striatal and cortical mechanisms would be unable to distinguish such false activations from true positive prediction error signals and react with plastic changes. Thus, any stimulus present during the action of amphetamine and cocaine would be learned as a reward predictor and facilitate approach behavior, including stimuli not associated with prediction errors that are normally blocked from learning (Waelti et al., 2001).

Reward Discrimination

The capacity to discriminate between different rewards is important for selecting the most valuable reward during decision making. Reward discrimination is limited by two processes, namely stimulus generalization, which is due to physical similarity between stimuli, and pseudoconditioning, which occurs via context conditioning by primary reinforcers (Mackintosh, 1974; Sheafor, 1975). The activating dopamine responses to stimuli consist of two components. The initial component is prone to generalization and thus discriminates poorly, whereas the second component distinguishes well between differently rewarded stimuli (Figure 2B). Generalization in the first component of dopamine responses occurs with neutral stimuli (Schultz and Romo, 1990; Waelti et al., 2001), aversive stimuli (Mirenowicz and Schultz, 1996; Joshua et al., 2008; Matsumoto and Hikosaka, 2009), explicit nonreward predicting stimuli (conditioned inhibitors; Tobler et al., 2003), and delay-predicting stimuli (Kobayashi and Schultz, 2008). Substantial fractions of dopamine neurons are activated by physically salient stimuli (Ljungberg et al., 1992), although these responses seem to be largely due to pseudoconditioning (Kobayashi and Schultz, 2010). The initial, "false," generalized or pseudoconditioned activation is often followed by a depressant response that may not entirely cancel the effects of the activation. Thus stimulus generalization may lead to net striatal dopamine release with neutral stimuli (Day et al., 2007). However, although generalization and

pseudoconditioning reduce reward discrimination, they may play a useful role in enhancing the detection of potential rewards.

The limited reward discrimination by dopamine neurons may become disastrous for drug addiction in two additive ways. First, drugs would enhance the generalized activations by neutral and aversive stimuli and the pseudoconditioned responses in rewarding contexts, along with the primary responses to rewarded stimuli and rewards. Second, the drug-induced tonic striatal dopamine increase may blunt the depression that often follows and partly compensates the generalized or peudoconditioned activation. The enhanced generalized or peudoconditioned dopamine activation would result in reduced reward discrimination by postsynaptic striatal mechanisms. Thus, under the influence of drugs, less rewarded, neutral, or aversive stimuli would lead to stronger phasic striatal dopamine changes and increased dopamine-dependent postsynaptic responsiveness and plasticity. Interestingly, addictive drugs do not usually induce withdrawal behavior before addiction develops, even though the few dopamine activations induced by aversive stimuli are likely to be also enhanced, confirming the strength of the reward over that of aversive dopamine function.

The breakdown of dopamine reward discrimination does not necessarily mean impairment of reward discrimination in general. Populations of neurons in the striatum, orbitofrontal cortex, and many other reward structures discriminate well between different rewards, including liquids, foods, and drugs (Carelli and Deadwyler, 1994; Bowman et al., 1996; Chang et al., 1998; Tremblay and Schultz, 1999; Hassani et al., 2001). Obviously, these reward discriminations rely on inputs other than dopamine inputs. In this way, general reward discrimination may be maintained during drug action, even when dopamine-dependent learning mechanisms lose discrimination.

Punishment and Conditioned Inhibition

Due to the long-term consequences, drugs should be considered as aversive; however, they rarely induce avoidance behavior as the aversive effect is overwhelmed by the rewarding component. Furthermore, drug use prevents drug users from receiving other rewards, including money, salaries, and friends. Thus drugs have a conditioned inhibitory component.

Minor fractions of dopamine neurons are activated by aversive stimuli when stimulus generalization is ruled out, and many dopamine neurons show depressions of activity (Figure 2C) (Mirenowicz and Schultz 1996; Matsumoto and Hikosaka, 2009). Conditioned inhibitors lead to occasional activating stimulus generalization responses (Figure 2B), but depressant responses prevail (Figure 2D) (Tobler et al., 2003).

Dopamine neurons should respond with depression of activity to the aversive and inhibitory functions of the aversive component of drugs, which would reduce the overall dopamine response to drugs. However, the enhanced response generalization and the tonic increase in dopamine induced by all major drugs of addiction would blunt the effects of depressant dopamine responses on striatal mechanisms and thus annihilate the potentially moderating aversive component of addictive drugs.

Novelty

Novelty induces attention, modulates the learning rate parameter of associability learning rules, and thus enhances learning (Mackintosh 1975; Pearce and Hall, 1980). Novelty enhances existing dopamine responses and induces activations to stimuli of sufficient minimal intensity (Figure 2E) (Ljungberg et al., 1992). Addictive drugs may increase the novelty response of dopamine neurons at their inputs and increase striatal impulse-dependent dopamine release. This effect may enhance behavioral learning via striatal plasticity, which may be conceptually linked to the learning rate parameter that determines the impact of the prediction error.

Temporal Discounting

Temporal discounting refers to the observation that later rewards have lower subjective value than earlier rewards. Decision makers need to control their usual preferences for earlier rewards (impulsivity) to avoid loss of potentially large rewards. Temporal discounting is associated with reduction of neuronal responses to later rewards in all major reward structures. Addictive drugs may affect impulsivity, enhance temporal discounting, and lead to disadvantageous choices via their effects on neuronal reward responses.

Behavioral preferences for sooner over later rewards indicate that delayed rewards lose subjective value even though their objective value remains unchanged (Ainslie, 1975; Mazur, 2002). The factors underlying temporal discounting include impatience, impulsivity, value deterioration, and fear of loss. Neurophysiological responses to reward-predicting stimuli decrease with increasing delays in orbitofrontal, dorsolateral prefrontal and parietal cortex, striatum, and dopamine neurons (Figure 3A) (Roesch and Olson, 2005; Roesch et al., 2007, 2009; Kim et al., 2008; Kobayashi and Schultz, 2008; Louie and Glimcher, 2010). Correspondingly, human blood oxygen-level dependent (BOLD) responses to reward delay-predicting stimuli measured in functional magnetic resonance imaging (fMRI) decrease with increasing delays in ranges between seconds and months in ventral striatum and medial prefrontal cortex, correlating with individual degrees of temporal discounting (Figure 3C) (Kable and Glimcher, 2007; Gregorios-Pippas et al., 2009). Response decreases across a few seconds may reflect an innate discounting mechanism that evolved for dealing with value deterioration of natural rewards. The dopamine prediction error response at the time of the reward increases with the delay (Figure 3B).

Temporal discounting may contribute to drug addiction as myopia on immediate rewards and neglect of larger distant rewards. Addicted individuals discount nondrug rewards such as money steeper than controls do (Figure 3D), which reflects drug exposure rather than a premorbid trait (Bickel et al., 1999). Furthermore addicts discount drugs more than nondrug rewards (Bickel and Marsch, 2001). Steeper discounting with the highly valuable drugs compared with the less valuable money is at odds with the usually steeper discounting of smaller compared with larger reward values when drugs are not involved (Kirby and Marakovic, 1996), suggesting that normal behavioral discounting is overridden and enhanced by drugs.

Behavioral discounting is correlated with neuronal discounting in nonaddicted individuals (Figure 3C) and might show similar relationships in drug addicts. Steeper neuronal discounting may be based on increased neuronal responses to reward-predicting stimuli in striatum and frontal cortex and produce

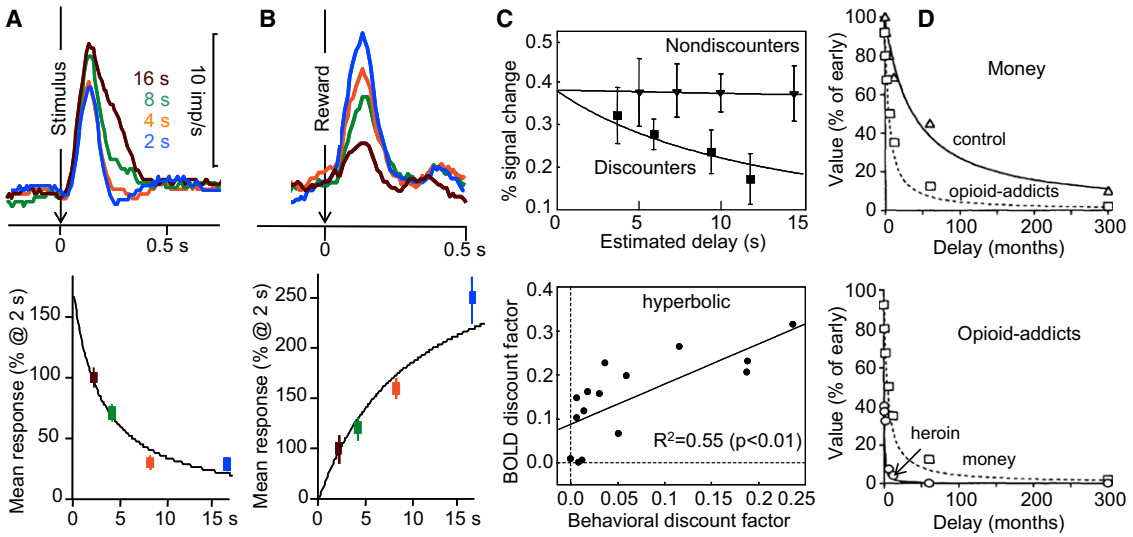

Figure 3. Temporal Discounting

(A) Temporal discounting of dopamine responses to reward delay-predicting stimuli. Stimulus-reward delays were 2 (brown), 4 (green), 8 (orange), and 16 s (blue), each being indicated by a different, small fractal stimulus. Averaged normalized population responses from 54 electrophysiologically recorded single dopamine neurons (top) suggest hyperbolic discounting (bottom). From Kobayashi and Schultz (2008), with permission by Society for Neuroscience.

(B) Increasing responses of dopamine neurons to delivery of identical liquid reward magnitude following increasing reward delays. Top: Averaged normalized population responses from 33 dopamine neurons. Bottom: Hyperbolic increase of neuronal activity with delays. From Kobayashi and Schultz (2008), with permission by Society for Neuroscience.

(C) Temporal discounting of fMRI BOLD responses to reward delay-predicting stimuli in human ventral striatum. Top: Hyperbolic BOLD discounting in seven participants showing behavioral discounting, but absent BOLD discounting in seven behavioral nondiscounters. Each objective interval (4, 6, 9, and 13.5 s) was indicated by a different, small fractal stimulus. Subjective stimulus-reward delays were individually estimated by peak-interval procedure; their averages (triangles and squares) were used for curve-fitting. Bottom: Correlations between BOLD and behavioral hyperbolic discount factors (15 participants). From Gregorios-Pippas et al. (2009), with permission by American Physiological Association.

(D) Temporal discounting in opioid addicts. Top: Steeper discounting of monetary reward in addicts compared with controls. Bottom: Steeper discounting in addicts for heroin compared with money. From Madden et al. (1997), with permission by American Psychological Association.

higher reward prediction as a result of several drug effects. Altered prediction error responses could increase striatal plasticity and enhance neuronal responses and reward prediction; positive dopamine prediction error responses at the time of the reward (Figure 3B) would be increased because more discounted reward value leads to larger prediction error; negative prediction error signals would be blunted by increased tonic dopamine levels. Further factors could be changes in dopamine-dependent presynaptic and postsynaptic plasticity (Figures 1A and 1E) and higher phasic dopamine changes with cocaine- or amphetamine-induced reuptake blockade (Figure 1C).

According to the dual systems account, an impulsive ("beta") system is driven by the value of immediate rewards, whereas an inhibitory, cognitive ("delta") system restrains immediate behavioral reactions and mediates responses to delayed rewards (Phelps and Pollack, 1968; Laibson, 1997). Human imaging shows BOLD responses in ventral striatum during choices for sooner over later rewards (beta system), suggesting impulsive valuation, and in dorsolateral prefrontal cortex during choices of later rewards that may reflect the action of the behavioral control system (delta system; McClure et al., 2004).

Addictive drugs may also affect discounting according to the dual system account. Enhancement of tonic dopamine levels in prefrontal cortex by drugs may impair the prefrontal control system (delta), possibly by altering mnemonic and planning mechanisms in this cortical area (Murphy et al., 1996; Elliott et al., 1997). In addition, addictive drugs would increase stronger approach behavior via enhanced neuronal responses to reward-predicting stimuli which would further challenge behavioral control. The two mechanisms would lead to additive impairments in discounting and result in more impulsive drug approach.

Influence of Background Reward on Prediction (Contingency)

Reward predictions induce approach behavior and provide essential information for informed decisions. Such predictions inform about the reward occurring with a stimulus relative to no stimulus (background). Thus reward predictions are influenced by background reward. The assessment of background reward requires working memory and discrimination against stimuli. Disturbance of these processes by addictive drugs may lead to incorrect predictions resulting in inadequate and enhanced approach behavior and compulsion.

A stimulus predicts a reward when more reward occurs with the stimulus than without the stimulus. This can be achieved in two ways. In conventional experiments, reward is only given during the stimulus, which amounts to stimulus-reward pairing. In more realistic situations, some rewards may occur also without stimuli. Thus reward predictions should take the background reward into account and provide differential information about reward during stimuli relative to background. For example,

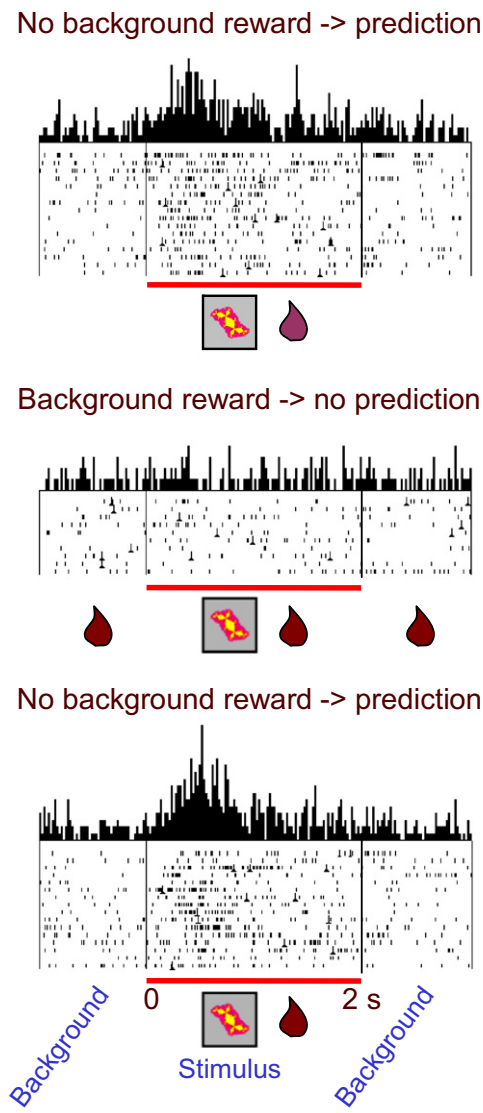

Figure 4. Influence of Background Reward on Prediction (Contingency) in a Single Amygdala Neuron

The neuron responds to the rewarded stimulus when background is unrewarded (top), loses the response when background reward increases to that of the stimulus (middle), and regains the response when background reward drops again (bottom). Thus the neuronal response to the reward-predicting stimulus depends entirely on the background reward, thus fulfilling the contingency requirement of animal learning theory. The visual stimulus (fractal during red-lined period) and reward (purple drop) were identical during the stimulus period in top, middle, and bottom graphs. From Bermudez and Schultz (2010a), with permission by American Physiological Society.

elevating background reward without changing stimulus reward would reduce the reward prediction by the stimulus as some reward occurs anyway; the reward becomes less contingent on the stimulus. By contrast, reducing background reward alone increases reward prediction by the stimulus, as the reward becomes more contingent on the stimulus. In animal learning theory, reward contingency, rather than stimulus-reward pairing, constitutes the fundamental requirement for reward prediction and learning (Rescorla, 1967).

Most laboratory experiments manipulate reward only during a stimulus. Neurons in all reward structures show activating or depressant responses to such stimuli. Manipulations of background reward use contingency to demonstrate true predictions rather than simple stimulus-reward pairing. The amygdala is necessary for true reward prediction, as its lesion makes rats insensitive to background reward increases; the animals continue to respond to stimuli that have lost their predictive properties (Ostlund and Balleine, 2008). Correspondingly, stimulus responses of amygdala neurons are sensitive to changes in background reward (Figure 4) (Bermudez and Schultz, 2010a). By taking background reward into account, these neurons are sensitive to contingency and thus code true reward predictions.

Assessing reward contingencies requires appropriate processing of events during the background and relating them to the specific stimulus. Necessary processes involve working memory about reward occurrence in the background, discrimination between background stimuli and the specific stimulus in order to attribute the reward changes to the background, and comparison between background reward and stimulus reward. Working memory and discrimination are known to be impaired by tonic dopamine increases in frontal and temporal cortex (Murphy et al., 1996; Elliott et al., 1997; Liu et al., 2008). Thus it is certainly possible that addictive drugs affect the assessment of reward contingency via their influence on tonic dopamine levels in striatum, frontal cortex and, in particular, the amygdala, with its contingency-sensitive neurons.

Drug addicts often suffer from mental disorders including anxiety, depression, and schizophrenia. Due to the cognitive deficits associated with these disorders, drug addicts may subjectively perceive their environmental background as less rewarding than normal. Even without mental illness, environmental background rewards are reduced when peer pressure toward drug use excludes nonusers, which constitutes a known factor in initial drug consumption. Thus the reduction of background reward would make stimuli associated with drug rewards more valuable and induce approach behavior.

The contingency requirement for learning may offer explanations for drug approach and addiction behavior by focusing not only on the attractive power of drug-associated stimuli, but by also considering the role of background rewards whose perception is likely changed in addicts.

Neuronal Risk Signals

In most natural situations rewards occur infrequently and vary considerably. The resulting incomplete knowledge about rewards introduces risk and uncertainty. The assessment of risk is a fundamental component of reward processing, as risk affects the subjective valuation of rewards and is important for optimal choices. Risk assessment relies on the correct estimation of reward probabilities. As there are no sensory receptors for probability, the brain needs to derive probability from the frequency of past events. The standard deviation of probability distributions is often termed ''risk,'' as it refers to the variation or spread in the distribution. Other valid measures of risk include the statistical variance (squared standard deviation) and the informational entropy. Here, risk is viewed as a form of

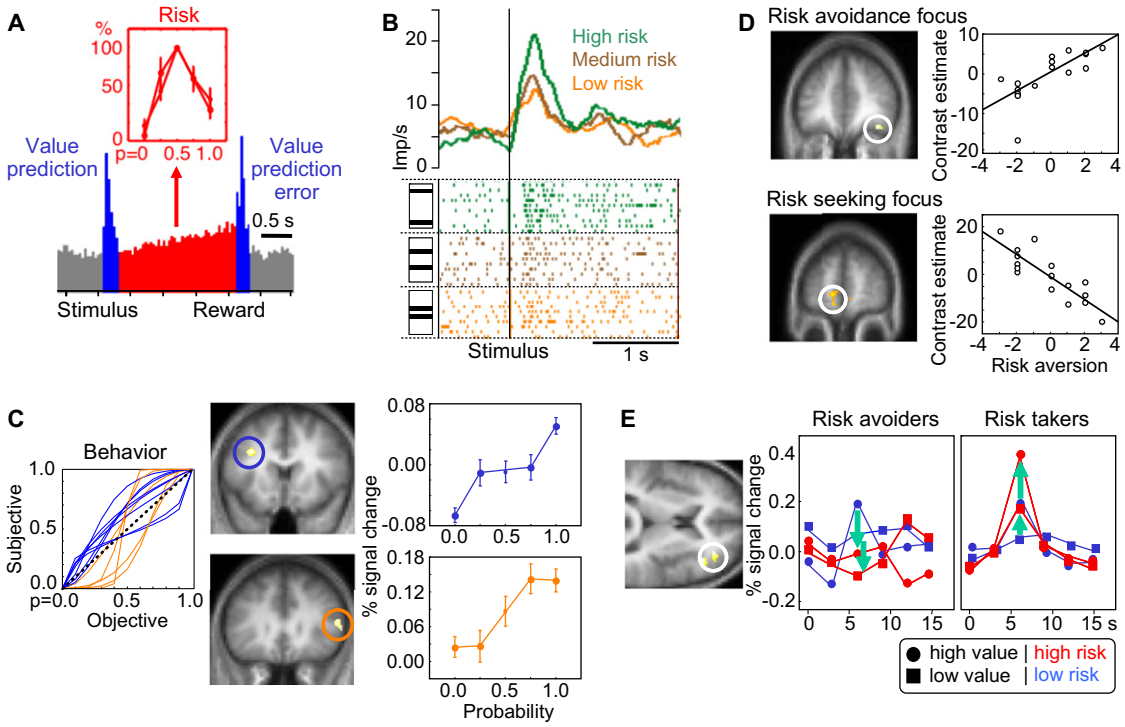

Figure 5. Neuronal Risk Signals

(A) Risk signaling in primate dopamine neurons during the stimulus-reward interval of a Pavlovian task with probabilistic rewards (averaged population activity from 44 neurons). The risk signal (center, red) is distinctively different from the value prediction and prediction error signal at the time of the stimulus and reward, respectively (blue). From Fiorillo et al. (2003), with permission by American Association for the Advancement of Science.

(B) Risk signaling in primate single orbitofrontal neuron. The response occurred to the stimulus predicting two equiprobable ($p = 0.5$) reward magnitudes indicated by the height of two bars; the three stimuli altered pseudorandomly. The response increased monotonically with the standard deviation of the three binary probability distributions (risk). From O'Neill and Schultz (2010), with permission by Cell Press.

(C) Neuronal probability distortions in human prefrontal cortex correlating with behavioral probability distortions (left). Different prefrontal areas were activated in subjects showing typical inverted S-shaped distortions (top, blue) and atypical S-shaped distortions (bottom, orange). From Tobler et al. (2008), with permission by Society for Neuroscience.

(D) Relationship of frontal risk signals to subjective risk preferences. Risk aversion was determined by behavioral choices between safe and risky outcomes. Activations in human imaging increased with risk aversion (top, potential risk warning signal) and decreased with risk aversion in medial frontal cortex (bottom, risk seeking signal). From Tobler et al. (2007), with permission by American Physiological Society.

(E) Differential influence of risk on prefrontal value signal in risk avoiders and risk takers. The neuronal responses code reward value, as shown by the dotted and squared curves, and are decreased in risk avoiders and increased in risk takers. From Tobler et al. (2009), with permission by National Academy of Sciences.

uncertainty rather than the probability of losing, which constitutes negative value. Risk-taking behavior can be enhanced by at least three mechanisms, namely genuine changes toward risk taking ("I love the risk"), overvaluation of high outcomes ("I don't care for small change"), and optimistic distortions of probabilities of above-average outcomes ("today is my lucky day").

Neuronal Risk Signals

In addition to phasically signaling reward prediction errors, about one-third of dopamine neurons show a separate, slower, and more sustained activation during the interval between a stimulus and a reward (Fiorillo et al., 2003). The activation shows an inverted U-shaped relationship to probability similar to that of standard deviation, variance, and entropy, and does not correlate with expected value, which increases monotonically with probability (Figure 5A). Additional tests with binary equiprobable ($p = 0.5$) reward distributions hold entropy constant and confirm the coding of standard deviation or variance rather than reward value. Thus the slow, sustained dopamine signal codes the risk of rewarding outcomes. Although the latency of this activation is too long to participate in decision processes, it might affect the dopamine released by a subsequent prediction error response and potentially influence learning via the associability term of attentional learning rules (Mackintosh, 1975; Pearce and Hall, 1980).

A group of neurons in orbitofrontal cortex signal reward risk distinct from reward value (Figure 5B) (O'Neill and Schultz, 2010). The risk responses increase monotonically with higher standard deviations of binary equiprobable distributions of reward amounts. Their latency is sufficiently short to allow them to participate in decision making before overt choices occur. Movement-related activity in posterior cingulate cortex increases with risk, possibly reflecting the subjective value of risky rewards (McCoy and Platt, 2005).

Addictive drugs acting on dopamine mechanisms are likely to affect risk signals in dopamine neurons and dopamine-innervated structures such as orbitofrontal cortex. The dopamine risk signal may be enhanced by the action of drugs on long-term potentiation and depression at input synapses of dopamine

neurons. The influence of this enhanced signal on postsynaptic neurons would be boosted by reduced dopamine reuptake with psychostimulants and by increased dopamine-dependent postsynaptic responsiveness and plasticity. These mechanisms could have two separate consequences. First, an enhanced dopamine risk signal would boost the effects of positive prediction error signals and blunt the effects of negative error signals, enhancing the associability term of attentional learning rules. The consequences could be runaway synaptic strengths in postsynaptic neurons. Second, enhanced dopamine-dependent postsynaptic responsiveness and plasticity in orbitofrontal neurons might reduce the distinction between risk and value signals in orbitofrontal neurons, which might come to respond to the higher values of risky outcomes while neglecting lower outcomes, thus confounding risk and value signals and producing an overly optimistic value signal. These two effects could produce risk taking and lead to compulsive approach to drugs.

Variations between Individuals

Behavioral economics suggests that humans perceive instructed outcome probabilities in a distorted manner, often overestimating low probabilities and underestimating high probabilities below p = 1.0. In close correspondence to behavioral measures, neuronal responses in human prefrontal cortex show similar probability distortions, suggesting that this brain region codes probability in a subjective rather than objective manner (Figure 5C) (Tobler et al., 2008; Hsu et al., 2009). Addictive drugs may induce new or exacerbate existing probability distortions by increasing tonic dopamine levels, which impair mnemonic processes (Murphy et al., 1996; Elliott et al., 1997) that are necessary for adequate probability estimation. This mechanism may lead to altered risk assessment, which relies on probability estimation, and may induce enhanced risk taking if probabilities of high outcomes become overestimated.

Stimuli associated with higher risk elicit increasing activations in human striatum and orbitofrontal cortex (Preuschoff et al., 2006; Tobler et al., 2007). The risk-related activations covary with individual risk aversion in lateral orbitofrontal cortex and with risk-seeking in medial frontal cortex (Figure 5D). Risk affects decision making by influencing the subjective values of competing outcomes. Indeed, the terms "risk aversion" and "risk taking" indicate that risk reduces or enhances subjective reward value. This concept constitutes a basic characteristic of economic utility theory. As a direct neuronal correlate, risk enhances neuronal value responses in lateral prefrontal cortex of human risk seekers and reduces value responses in risk avoiders (Figure 5E) (Tobler et al., 2009). Thus, both the coding of risk itself and the influence of risk on value coding correlate with behavioral risk preferences of individuals, suggesting subjective rather than objective processing. Because these mechanisms occur in frontal cortical areas that are innervated by dopamine neurons, basically all phasic and tonic alterations of dopamine functions by addictive drugs may affect risk preferences.

Initial Drug Episodes

When thrill-seeking, peer pressure, or mental disorders encourage a few initial drug taking episodes, neuronal risk signals might change under the influence of these drugs and lead to changes in behavioral risk preferences. If these changes turn into an increased willingness to take risks, individuals may engage in further drug taking and develop an addiction. Thus altered neuronal risk processing could play a particularly detrimental role during initial, preaddictive stages of drug taking.

Adaptive Reward Coding

Neuronal responses to reward are optimized for reward discrimination. The mechanism involves adjusting the range of neuronal coding to the range of currently available reward values or, more specifically, matching probability distributions of neuronal responses to probability distributions of reward values. The adaptive process may be disturbed when addictive drugs alter reward valuation or prediction, perception of reward range, or the adjustment process itself. The consequences would be suboptimal learning and reward discrimination.

Normal Mechanisms

Whereas the processing capacity of the brain is subject to biological constraints, the number of possible rewards is almost unlimited. Dedicating equal processing capacity to all possible rewards would lead to low slopes of reward-response functions and poor reward discrimination. However, the amount of available rewards can vary considerably, and often only a few rewards are available. An efficient mechanism could maintain good discrimination by matching neuronal processing to currently available rewards and neglecting rewards that are absent. Such adaptation would adjust learning to current demands, optimize reward-response slopes for reward discrimination, and improve behavioral choices. Indeed, the behavioral contrast effect in experimental psychology and the reference dependency in experimental economics demonstrate that outcomes are valued within changing frames of reference and that identical outcomes are valued differently depending on which alternatives are available (Tinklepaugh, 1928; Crespi, 1942; Black, 1968; Kahneman and Tversky, 1979). If neuronal processing focuses on current probability distributions of reward values, reward-response slopes adapt and reward discrimination becomes optimal. Adaptations occur to the main parameters of probability distributions (Figure 6A), namely expected value and standard deviation. Thus, appropriate adaptation requires correct assessment of these parameters.

Prediction errors for updating reward values need to distinguish between two principal types of reward fluctuation: changes in expected value (step changes), and changes in standard deviation (stochastic variations) (Figure 6B top). Prediction errors of the same magnitudes are indistinguishable between step changes and stochastic variations (Figure 6B middle). However, a prediction error that is not particularly meaningful in stochastic variations becomes quite meaningful when it reflects a major step change above small fluctuations; here learning should be strong (Behrens et al., 2007; Speekenbrink and Shanks, 2010). Scaling prediction errors to standard deviation would relate them to the underlying nature of fluctuation and make them more meaningful (Figure 6B bottom) (Nassar et al., 2010).

Responses in many reward neurons of orbitofrontal cortex, striatum, and amygdala adapt to the expected value and standard deviation of reward distributions (Figures 6C and 6D) (Tremblay and Schultz, 1999; Cromwell et al., 2005; Hosokawa et al., 2007; Padoa-Schioppa, 2009; Kobayashi et al., 2010; Bermudez

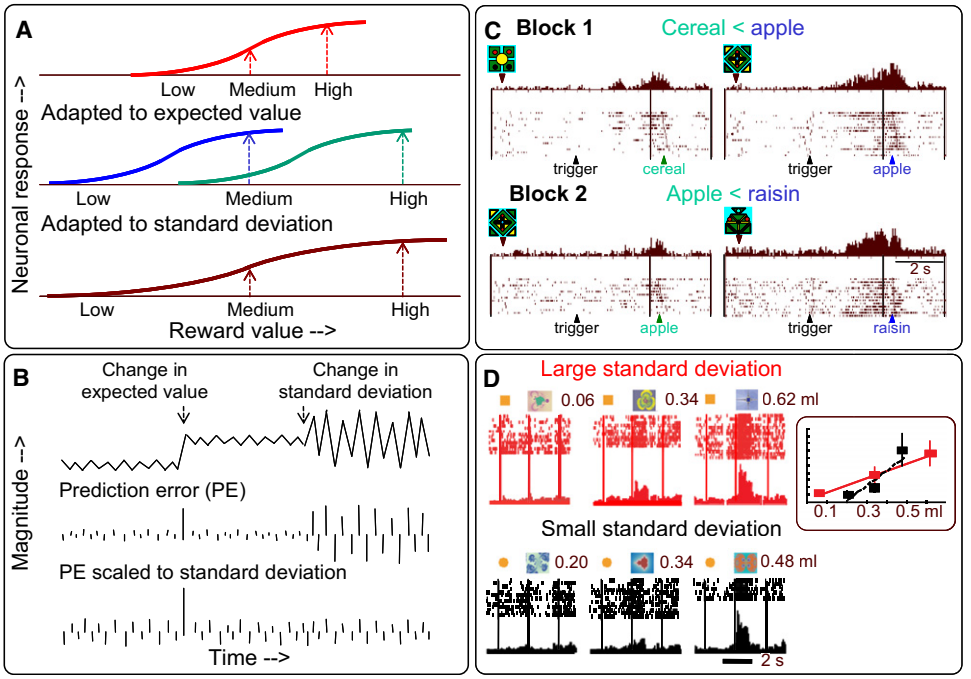

Figure 6. Adaptive Reward Coding

(A) Schematics of adaptation of neuronal responses to expected value and standard deviation of reward value distributions.

(B) Schematics of reward fluctuations and adaptation of prediction errors (received reward value minus expected value). Top: The two principal types of fluctuation, major change (left) and large stochastic variation (right). Middle: Prediction errors as such do not distinguish between fluctuation types. Bottom: Prediction errors scaled via division by standard deviation distinguish major changes from stochastic variations of same size.

(C) Adaptation of neuronal reward response in orbitofrontal neuron to approximate expected value. Reward values as derived from behavioral preferences were cereal < apple < raisin. Visual instructions predict type of reward, and trigger stimuli elicit an arm movement leading to the reward. Data from Tremblay and Schultz (1999), with permission by MacMillan.

(D) Adaptation of neuronal response in orbitofrontal neuron to standard deviation of reward volume (ml). Inset shows change of reward-response slope in the neuron shown in main part (result from two linear regressions on reward magnitude). Data from S. Kobayashi and W. Schultz (2010, Soc. Neurosci., abstract), with permission by Society for Neuroscience.

and Schultz, 2010b). The prediction error response of dopamine neurons seems to be scaled by standard deviation (Tobler et al., 2005); the underlying arithmetic division could involve shunting inhibitions along dendrites and soma. The phenomenon is also found in human orbitofrontal cortex and striatum (Breiter et al., 2001; Nieuwenhuis et al., 2005). Ventral striatal lesions in rats reduce behavioral reward contrast (Leszczuk and Flaherty, 2000). These data suggest a role of the dopamine system and other reward structures in adaptation to reward value.

Potential Vulnerabilities

Coding of expected value and standard deviation. Adaptive reward coding requires appropriate neuronal processing of experienced and predicted probability distributions of reward values and their key parameters. Basically all phasic and tonic alterations of presynaptic and postsynaptic dopamine functions by addictive drugs may alter the assessment of reward value and standard deviation. Inadequate processing of reward contingency would compromise the predictive components of adaptive coding.

Adaptation process. Addictive drugs may compromise the adaptation process itself. The scaling of dopamine responses by shunting inhibition could be compromised by altered long-term potentiation or depression at dopamine input synapses. The effects of altered dopamine prediction error responses, boosted by blockade of dopamine reuptake with amphetamine and cocaine, could lead to changes in short-term postsynaptic dopamine-dependent plasticity in striatum, orbitofrontal cortex, and amygdala, which might constitute a substrate for the adaptation. The working memory and discrimination impairments following increased tonic dopamine levels (Murphy et al., 1996; Elliott et al., 1997) could destabilize the adjustment to the current reward distribution.

Consequences. Malfunctioning adaptive coding would lead to suboptimal reward discrimination (Figure 7). The mechanisms include saturation of neuronal coding at lower or higher reward values (Figures 7A–7C and 7F), incomplete use of neuronal coding range (Figures 7A, 7B, 7D, and 7E), and unnecessarily flat reward-response slopes (Figure 7D). As a result, the neurons would be unable to accurately assess and discriminate the values of different rewards, which may contribute to an exaggerated subjective valuation of drugs. Wrong scaling of reward prediction error signals may lead also to confusion between step changes and stochastic variations and induce unwarranted learning and unstable approach behavior.

Decision Making

The acquisition of rewards, including addictive drugs, involves choices between different, predictable options. Economic

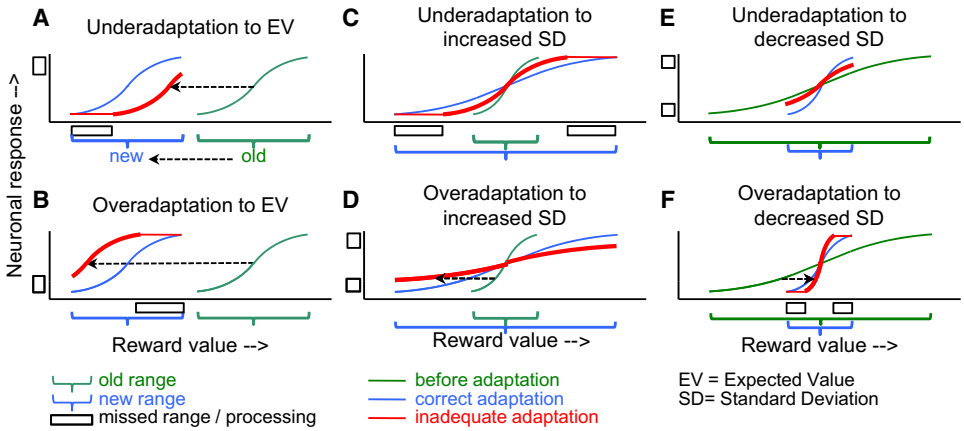

Figure 7. Malfunctioning Adaptive Coding Leading to Suboptimal Reward Discrimination

(A) Underadaptation of neuronal coding to changed expected value (EV) results in insufficient shift of new coding range, lack of coding of lower reward values (open rectangle below x axis), and incomplete use of upper neuronal coding range (open rectangle at y axis).

(B) Overadaptation to EV results in saturation at higher values and incomplete use of lower neuronal coding range (open rectangles at x and y axes, respectively).

(C) Underadaptation to increased standard deviation (SD) results in saturation at lower and higher values (open rectangles below x axis).

(D) Overadaptation to increased SD results in incomplete use of lower and upper neuronal coding range (open rectangles at y axis) and unnecessarily flat reward-response slope.

(E) Underadaptation to decreased SD results in incomplete use of lower and upper neuronal coding range (open rectangles at y axis).

(F) Overadaptation to decreased SD results in saturation at lower and higher values (open rectangles below x axis).

decision processes use values of predicted rewards as inputs and engage comparisons between these values. Because most valuation and comparison mechanisms engage the dopamine system and its postsynaptic structures, decision making is likely to be disrupted by the impact of drugs on dopamine mechanisms.

Normal Mechanisms

An economic decision variable uses information from a number of heterogeneous neuronal reward signals in a form that is appropriate for the specific comparison mechanism underlying the decision process. These variables serve as inputs to decision processes or as their outputs toward the execution of behavioral choices. Current simple decision models are based on comparisons between subjective values of predicted options, including hedonic aspects. All other parameters contributing to decisions, including reward delay and risk, would be transformed into value in order to participate in decisions (e.g., via temporal discounting and risk aversion). Thus all valuation processes described above will influence decision making.

In reinforcement theory, predicted action value represents a key input variable (the reward value obtained for a specific action irrespective of that action being chosen) (Sutton and Barto, 1998). Object value could serve an analogous role (the value of a specific reward object irrespective of being chosen). Chosen value (the value of the chosen option) is an output variable of the decision process. Neurons in striatum code action value (Figure 8A; Samejima et al., 2005; Lau and Glimcher, 2008), orbitofrontal neurons code object value (Figure 8B; Padoa-Schioppa and Assad, 2006), and striatal and orbitofrontal neurons code chosen value (Padoa-Schioppa and Assad, 2006; Lau and Glimcher, 2008). Dopamine reward prediction error signals are thought to update these decision variables with current reward values. The comparison between values would involve, in the most simple case, a winner-take-all mechanism that transforms a graded difference into an all-or-none distinction. This model is adequate for economic choices between alternatives with known values (Sutton and Barto, 1998).

In other decision models, value comparisons employ the ratio of reward probabilities of all predicted options as decision variable (log-likelihood ratio; see Gold and Shadlen, 2007). Neurons in parietal cortex code this decision variable during probabilistic reward choice (Figure 8C; Yang and Shadlen, 2007). Concepts of sensory discrimination and decision making suggest that more sophisticated decisions involve gradual accumulation of evidence and can be modeled by competing bounded diffusion processes (Ratcliff et al., 1999; Gold and Shadlen, 2007). In economic decisions, the current values of the individual options may race in a random walk fashion toward specific thresholds; the option whose process reaches its threshold first determines the behavioral choice. Formal modeling suggests feasible combinations of winner-take-all and diffusion-race models (Lo and Wang, 2006). Thus, decision mechanisms require valuation, prediction, working memory, computations, comparisons, and planning.

Potential Vulnerabilities

Subjective reward values and log likelihood ratios constitute decision variables that serve as inputs to decision mechanisms. The addiction process is associated with an increased subjective valuation of drugs, irrespective of drug value at first use. Depending on the reward system under study, the value increase could be induced by the molecular action of the drug on neuronal reward mechanisms and constitute an essential component of the addiction process; alternatively, the value could increase as a simple consequence of the addiction process occurring in other brain structures. Enhanced positive reward prediction error

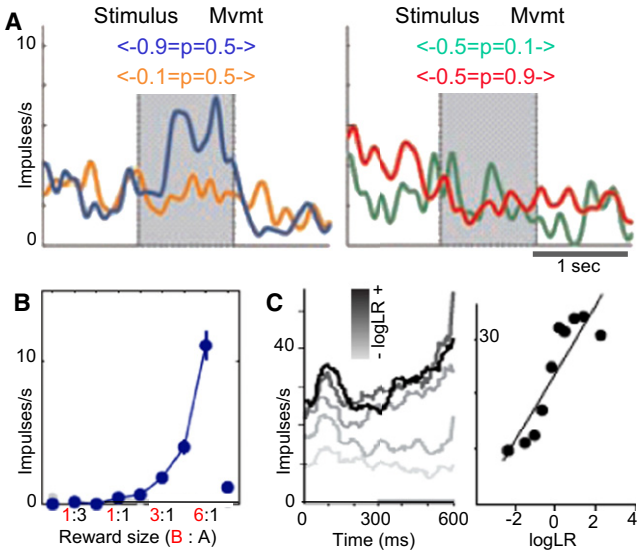

Figure 8. Neuronal Coding of Decision Variables

(A) Action value coding in a striatum neuron. This neuron coded left action value, as premovement activity correlated with reward probability for left movement trials (p = 0.9 compared to p = 0.1; left, brown versus blue), but not for right trials (right, green versus red). Data from Samejima et al. (2005), with permission by American Association for the Advancement of Science.

(B) Object value coding in orbitofrontal neurons. This neuron coded the size of reward juice B but not juice A. Data from Padoa-Schioppa and Assad (2006), with permission by MacMillan.

(C) Coding of log-likelihood ratio in a parietal cortex neuron. Neuronal activity increases with the log-likelihood ratio for a target in the response field of this neuron being rewarded. Data from Yang and Shadlen (2007), with permission by MacMillan.

signals, and blunted negative error signals, would lead to supranormal reward values of drugs. Reuptake blockade and increased dopamine-dependent postsynaptic responsiveness and plasticity would enhance the predictive neuronal coding of action value, object value, and chosen value of drugs in dopamine-innervated brain structures. Insufficient adaptive coding would produce suboptimal neuronal discrimination between reward values (Figure 7). Altered tonic striatal and frontal dopamine concentrations would impair working memory processes necessary for assessing contingencies and reward probabilities and computing log-likelihood ratios. Irrespective of the addiction mechanism acting on the individual neuron, the resulting high value of the drug would dominate the decision process and produce strong behavioral choices that appear to be beyond voluntary control.

Reward values would be integrated with subjective preferences to serve as inputs to decision mechanisms. Drugs can affect reward preferences in multiple ways via alterations of phasic and tonic dopamine processes, as described above for risk preferences.

Most of the drug-induced alterations of neuronal reward mechanisms may impact neuronal decision mechanisms. Altered tonic striatal and frontal dopamine concentrations would affect winner-take-all computations. Alterations in phasic and tonic dopamine activities might affect the electrophysiological noise in neurons, change random walks of activity in diffusion-race models, and allow drug-related activity to reach decision thresholds first. These drug-induced alterations of decision processes would affect behavioral choices between rewards and may result in exaggerated preferences for drug rewards. Some of these mechanisms may explain the general behavioral deficits in decision making observed in drug abusers (Rogers et al., 1999; Kalivas and Volkow, 2005), which, similar to the dual systems account of temporal discounting, may involve inhibitory prefrontal and impulsive subcortical reward systems (Chambers et al., 2003; Bechara, 2005).

ACKNOWLEDGMENTS

Work in the author's laboratory was supported by the Wellcome Trust, Swiss National Science Foundation, Cambridge Behavioural and Clinical Neuroscience Institute (BCNI), Human Frontiers, and other grant and fellowship agencies.

REFERENCES

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

Argilli, E., Sibley, D.R., Malenka, R.C., England, P.M., and Bonci, A. (2008). Mechanism and time course of cocaine-induced long-term potentiation in the ventral tegmental area. J. Neurosci. *28*, 9092–9100.

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. Nat. Neurosci. *8*, 1458–1463.

Behrens, T.E., Woolrich, M.W., Walton, M.E., and Rushworth, M.F. (2007). Learning the value of information in an uncertain world. Nat. Neurosci. *10*, 1214–1221.

Bermudez, M.A., and Schultz, W. (2010a). Responses of amygdala neurons to positive reward-predicting stimuli depend on background reward (contingency) rather than stimulus-reward pairing (contiguity). J. Neurophysiol. *103*, 1158–1170.

Bermudez, M.A., and Schultz, W. (2010b). Reward magnitude coding in primate amygdala neurons. J. Neurophysiol. *104*, 3424–3432.

Bickel, W.K., and Marsch, L.A. (2001). Toward a behavioral economic understanding of drug dependence: Delay discounting processes. Addiction *96*, 73–86.

Bickel, W.K., Odum, A.L., and Madden, G.J. (1999). Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. Psychopharmacology (Berl.) *146*, 447–454.

Black, R.W. (1968). Shifts in magnitude of reward and contrast effects in instrumental and selective learning: A reinterpretation. Psychol. Rev. *75*, 114–126.

Bonci, A., and Malenka, R.C. (1999). Properties and plasticity of excitatory synapses on dopaminergic and GABAergic cells in the ventral tegmental area. J. Neurosci. *19*, 3723–3730.

Bowman, E.M., Aigner, T.G., and Richmond, B.J. (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. J. Neurophysiol. *75*, 1061–1073.

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

Bunney, B.S., Walters, J.R., Roth, R.H., and Aghajanian, G.K. (1973). Dopaminergic neurons: Effect of antipsychotic drugs and amphetamine on single cell activity. J. Pharmacol. Exp. Ther. *185*, 560–571.

Capriles, N., Rodaros, D., Sorge, R.E., and Stewart, J. (2003). A role for the prefrontal cortex in stress- and cocaine-induced reinstatement of cocaine seeking in rats. Psychopharmacology (Berl.) *168*, 66–74.

Carelli, R.M., and Deadwyler, S.A. (1994). A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. J. Neurosci. *14*, 7735–7746.

Chambers, R.A., Taylor, J.R., and Potenza, M.N. (2003). Developmental neurocircuitry of motivation in adolescence: A critical period of addiction vulnerability. Am. J. Psychiatry *160*, 1041–1052.

Chang, J.Y., Janak, P.H., and Woodward, D.J. (1998). Comparison of mesocorticolimbic neuronal responses during cocaine and heroin self-administration in freely moving rats. J. Neurosci. *18*, 3098–3115.

Chesselet, M.F. (1984). Presynaptic regulation of neurotransmitter release in the brain: Facts and hypothesis. Neuroscience *12*, 347–375.

Corbett, D., and Wise, R.A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable electrode mapping study. Brain Res. *185*, 1–15.

Crespi, L.P. (1942). Quantitative variation in incentive and performance in the white rat. Am. J. Psychol. *40*, 467–517.

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.

Di Chiara, G., and Imperato, A. (1988). Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. Proc. Natl. Acad. Sci. USA *85*, 5274–5278.

Doucet, G., Descarries, L., and Garcia, S. (1986). Quantification of the dopamine innervation in adult rat neostriatum. Neuroscience *19*, 427–445.

Elliott, R., Sahakian, B.J., Matthews, K., Bannerjea, A., Rimmer, J., and Robbins, T.W. (1997). Effects of methylphenidate on spatial working memory and planning in healthy young adults. Psychopharmacology (Berl.) *131*, 196–206.

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

Freund, T.F., Powell, J.F., and Smith, A.D. (1984). Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. Neuroscience *13*, 1189–1215.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Goldman-Rakic, P.S., Leranth, C., Williams, S.M., Mons, N., and Geffard, M. (1989). Dopamine synaptic complex with pyramidal neurons in primate cerebral cortex. Proc. Natl. Acad. Sci. USA *86*, 9015–9019.

Gonon, F. (1997). Prolonged and extrasynaptic excitatory action of dopamine mediated by D1 receptors in the rat striatum in vivo. J. Neurosci. *17*, 5972–5978.

Grace, A.A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. Neuroscience *41*, 1–24.

Gregorios-Pippas, L., Tobler, P.N., and Schultz, W. (2009). Short-term temporal discounting of reward value in human ventral striatum. J. Neurophysiol. *101*, 1507–1523.

Groves, P.M., Garcia-Munoz, M., Linder, J.C., Manley, M.S., Martone, M.E., and Young, S.J. (1995). Elements of the intrinsic organization and information processing in the neostriatum. In Models of Information Processing in the Basal Ganglia, J.C. Houk, J.L. Davis, and D.G. Beiser, eds. (Cambridge, MA: MIT Press), pp. 51–96.

Gurden, H., Takita, M., and Jay, T.M. (2000). Essential role of D1 but not D2 receptors in the NMDA receptor-dependent long-term potentiation at hippocampal-prefrontal cortex synapses in vivo. J. Neurosci. *20*, RC106.

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.

Hernández-López, S., Bargas, J., Surmeier, D.J., Reyes, A., and Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca^{2+} conductance. J. Neurosci. *17*, 3334–3342.

Hikosaka, O., Sakamoto, M., and Usui, S. (1989). Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. J. Neurophysiol. *61*, 780–798.

Hollerman, J.R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. Nat. Neurosci. *1*, 304–309.

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.

Hsu, M., Krajbich, I., Zhao, C., and Camerer, C.F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. J. Neurosci. *29*, 2231–2237.

Izhikevich, E.M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. Cereb. Cortex *17*, 2443–2452.

Johnson, S.W., and North, R.A. (1992). Two types of neurone in the rat ventral tegmental area and their synaptic inputs. J. Physiol. *450*, 455–468.

Jones, S., and Bonci, A. (2005). Synaptic plasticity and drug addiction. Curr. Opin. Pharmacol. *5*, 20–25.

Jones, S.R., Garris, P.A., and Wightman, R.M. (1995). Different effects of cocaine and nomifensine on dopamine uptake in the caudate-putamen and nucleus accumbens. J. Pharmacol. Exp. Ther. *274*, 396–403.

Jones, S., Kornblum, J.L., and Kauer, J.A. (2000). Amphetamine blocks long-term synaptic depression in the ventral tegmental area. J. Neurosci. *20*, 5575–5580.

Joshua, M., Adler, A., Mitelman, R., Vaadia, E., and Bergman, H. (2008). Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. J. Neurosci. *28*, 11673–11684.

Kable, J.W., and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. *10*, 1625–1633.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica *47*, 263–291.

Kalivas, P.W., and Volkow, N.D. (2005). The neural basis of addiction: A pathology of motivation and choice. Am. J. Psychiatry *162*, 1403–1413.

Kim, S., Hwang, J., and Lee, D. (2008). Prefrontal coding of temporally discounted values during intertemporal choice. Neuron *59*, 161–172.

Kirby, K.N., and Marakovic, N.N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. Psychon. Bull. Rev. *3*, 100–104.

Kobayashi, S., and Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. J. Neurosci. *28*, 7837–7846.

Kobayashi, S., Pinto de Carvalho, O., and Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. J. Neurosci. *30*, 534–544.

Laibson, D. (1997). Golden Eggs and hyperbolic discounting. Q. J. Econ. *112*, 443–477.

Lau, B., and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. Neuron *58*, 451–463.

Leszczuk, M.H., and Flaherty, C.F. (2000). Lesions of nucleus accumbens reduce instrumental but not consummatory negative contrast in rats. Behav. Brain Res. *116*, 61–79.

Lichtensteiger, W., Hefti, F., Felix, D., Huwyler, T., Melamed, E., and Schlumpf, M. (1982). Stimulation of nigrostriatal dopamine neurones by nicotine. Neuropharmacology *21*, 963–968.

Liu, S., Heitz, R.P., Sampson, A.R., Zhang, W., and Bradberry, C.W. (2008). Evidence of temporal cortical dysfunction in rhesus monkeys following chronic cocaine self-administration. Cereb. Cortex *18*, 2109–2116.

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. J. Neurophysiol. *67*, 145–163.

Lo, C.-C., and Wang, X.-J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat. Neurosci. *9*, 956–963.

Louie, K., and Glimcher, P.W. (2010). Separating value from choice: Delay discounting activity in the lateral intraparietal area. J. Neurosci. *30*, 5498–5507.

Mackintosh, N.J. (1974). The Psychology of Animal Learning (London: Academic Press).

Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimulus with reinforcement. Psychol. Rev. *82*, 276–298.

Madden, G.J., Petry, N.M., Badger, G.J., and Bickel, W.K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: Drug and monetary rewards. Exp. Clin. Psychopharmacol. *5*, 256–262.

Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature *459*, 837–841.

Matthews, R.T., and German, D.C. (1984). Electrophysiological evidence for excitation of rat ventral tegmental area dopamine neurons by morphine. Neuroscience *11*, 617–625.

Mazur, J.E. (2002). Learning and Behavior, Fifth Edition (Upper Saddle River, NJ: Prentice Hall).

McClure, S.M., Laibson, D.I., Loewenstein, G., and Cohen, J.D. (2004). Separate neural systems value immediate and delayed monetary rewards. Science *306*, 503–507.

McCoy, A.N., and Platt, M.L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. Nat. Neurosci. *8*, 1220–1227.

Mirenowicz, J., and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature *379*, 449–451.

Murphy, B.L., Arnsten, A.F., Goldman-Rakic, P.S., and Roth, R.H. (1996). Increased dopamine turnover in the prefrontal cortex impairs spatial working memory performance in rats and monkeys. Proc. Natl. Acad. Sci. USA *93*, 1325–1329.

Nassar, M.R., Wilson, R.C., Heasly, B., and Gold, J.I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. J. Neurosci. *30*, 12366–12378.

Nieuwenhuis, S., Heslenfeld, D.J., von Geusau, N.J., Mars, R.B., Holroyd, C.B., and Yeung, N. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. Neuroimage *25*, 1302–1309.

O'Neill, M., and Schultz, W. (2010). Coding of reward risk distinct from reward value by orbitofrontal neurons. Neuron *68*, 789–800.

Ornstein, T.J., Iddon, J.L., Baldacchino, A.M., Sahakian, B.J., London, M., Everitt, B.J., and Robbins, T.W. (2000). Profiles of cognitive dysfunction in chronic amphetamine and heroin abusers. Neuropsychopharmacology *23*, 113–126.

Ostlund, S.B., and Balleine, B.W. (2008). Differential involvement of the basolateral amygdala and mediodorsal thalamus in instrumental action selection. J. Neurosci. *28*, 4398–4405.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. J. Neurosci. *29*, 14004–14014.

Padoa-Schioppa, C., and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. Nature *441*, 223–226.

Pawlak, V., and Kerr, J.N.D. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. J. Neurosci. *28*, 2435–2446.

Pearce, J.M., and Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. *87*, 532–552.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., and Frith, C.D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature *442*, 1042–1045.

Phelps, E.S., and Pollack, R.A. (1968). On second-best national saving and game-equilibrium growth. Rev. Econ. Stud. *35*, 391–402.

Pidoplichko, V.I., DeBiasi, M., Williams, J.T., and Dani, J.A. (1997). Nicotine activates and desensitizes midbrain dopamine neurons. Nature *390*, 401–404.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. Neuron *51*, 381–390.

Ratcliff, R., Van Zandt, T., and McKoon, G. (1999). Connectionist and diffusion models of reaction time. Psychol. Rev. *106*, 261–300.

Redish, A.D. (2004). Addiction as a computational process gone awry. Science *306*, 1944–1947.

Rescorla, R.A. (1967). Pavlovian conditioning and its proper control procedures. Psychol. Rev. *74*, 71–80.

Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Classical Conditioning II: Current Research and Theory, A.H. Black and W.F. Prokasy, eds. (New York: Appleton Century Crofts), pp. 64–99.

Reynolds, J.N.J., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of reward-related learning. Nature *413*, 67–70.

Roesch, M.R., and Olson, C.R. (2005). Neuronal activity in primate orbitofrontal cortex reflects the value of time. J. Neurophysiol. *94*, 2457–2471.

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.

Roesch, M.R., Singh, T., Brown, P.L., Mullins, S.E., and Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. J. Neurosci. *29*, 13365–13376.

Rogers, R.D., Everitt, B.J., Baldacchino, A., Blackshaw, A.J., Swainson, R., Wynne, K., Baker, N.B., Hunter, J., Carthy, T., Booker, E., et al. (1999). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: Evidence for monoaminergic mechanisms. Neuropsychopharmacology *20*, 322–339.

Saal, D., Dong, Y., Bonci, A., and Malenka, R.C. (2003). Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. Neuron *37*, 577–582.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. Science *310*, 1337–1340.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. J. Neurophysiol. *80*, 1–27.

Schultz, W. (2007). Multiple dopamine functions at different time courses. Annu. 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., and Romo, R. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements. I. Preparatory activity in the anterior striatum. Exp. Brain Res. *91*, 363–384.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J. Neurosci. *13*, 900–913.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.

Sheafor, P.J. (1975). "Pseudoconditioned" jaw movements of the rabbit reflect associations conditioned to contextual background cues. J. Exp. Psychol. Anim. Behav. Process. *1*, 245–260.

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D.J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. Science *321*, 848–851.

Smith, A.D., and Bolam, J.P. (1990). The neural network of the basal ganglia as revealed by the study of synaptic connections of identified neurones. Trends Neurosci. *13*, 259–265.

Speekenbrink, M., and Shanks, D.R. (2010). Learning in a changing environment. J. Exp. Psychol. Gen. *139*, 266–298.

Sutton, R.S., and Barto, A.G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. Psychol. Rev. *88*, 135–170.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning (Cambridge, MA: MIT Press).

Tinklepaugh, O.L. (1928). An experimental study of representation factors in monkeys. J. Comp. Psychol. *8*, 197–236.

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.

Tobler, P.N., O'Doherty, J.P., Dolan, R.J., and Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J. Neurophysiol. *97*, 1621–1632.

Tobler, P.N., Christopoulos, G.I., O'Doherty, J.P., Dolan, R.J., and Schultz, W. (2008). Neuronal distortions of reward probability without choice. J. Neurosci. *28*, 11703–11711.

Tobler, P.N., Christopoulos, G.I., O'Doherty, J.P., Dolan, R.J., and Schultz, W. (2009). Risk-dependent reward value signal in human prefrontal cortex. Proc. Natl. Acad. Sci. USA *106*, 7185–7190.

Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. Nature *398*, 704–708.

Tsai, H.-C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L., and Deisseroth, K. (2009). Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. Science *324*, 1080–1084.

Ungless, M.A., Whistler, J.L., Malenka, R.C., and Bonci, A. (2001). Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. Nature *411*, 583–587.

Venton, B.J., Zhang, H., Garris, P.A., Phillips, P.E.M., Sulzer, D., and Wightman, R.M. (2003). Real-time decoding of dopamine concentration changes in the caudate-putamen during tonic and phasic firing. J. Neurochem. *87*, 1284–1295.

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. Nature *412*, 43–48.

Wang, Z., Kai, L., Day, M., Ronesi, J., Yin, H.H., Ding, J., Tkatch, T., Lovinger, D.M., and Surmeier, D.J. (2006). Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. Neuron *50*, 443–452.

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. Nature *382*, 629–632.

Wise, R.A. (2002). Brain reward circuitry: Insights from unsensed incentives. Neuron *36*, 229–240.

Wise, R.A., and Bozarth, M.A. (1987). A psychomotor stimulant theory of addiction. Psychol. Rev. *94*, 469–492.

Yang, T., and Shadlen, M.N. (2007). Probabilistic reasoning by neurons. Nature *447*, 1075–1080.

Zweifel, L.S., Parker, J.G., Lobb, C.J., Rainwater, A., Wall, V.Z., Fadok, J.P., Darvas, M., Kim, M.J., Mizumori, S.J., Paladini, C.A., et al. (2009). Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. Proc. Natl. Acad. Sci. USA *106*, 7281–7288.