

hope that they recover). All of these diverse predictions are now supported by experimental data. Kahneman and Tversky's research—for which Kahneman was awarded the Nobel Prize in Economics in 2002 (Tversky was deceased, and thus ineligible)—supplied the foundation for **behavioral economics**. This new field seeks not only to characterize how factors like value and probability affect real-world choices, but to identify particular interventions that may improve decision making in a variety of contexts.

Reward and Utility

Understanding the neural basis of decision making requires first considering how the brain processes rewards that motivate choices. Rewards that have direct benefits for fitness, such as food, water, and sex, are called **primary reinforcers** (or *unconditioned reinforcers*). When decision-making experiments are conducted in non-human animals, the rewards are typically primary reinforcers that are valued even in the absence of training, such as drops of juice following successful performance of an experimental task. Many real-world decisions, at least for humans, do not involve primary reinforcers. Money, for example, has no intrinsic value but can be used to obtain other rewards; thus it is described as a **secondary reinforcer**. The removal of an aversive outcome following an action is called **negative reinforcement**, while the delivery of an aversive outcome is called **punishment**.

Two caveats are important to raise. First, the following sections will discuss learning and motivation specifically in the context of value-guided decision making. Other types of learning about positive and negative reinforcers (e.g., aversive responses to threatening stimuli, fear conditioning) involve brain circuitry important for affect and emotion, as was discussed in Chapter 10. Second, the brain regions and neurotransmitters discussed hereafter do not only contribute to decision making, but also have broader relevance for a wide range of adaptive behavior.

Dopamine: Pleasure or motivation?

For any animal to make behaviorally useful decisions, it must be able to predict the choices that lead to particular outcomes and to evaluate the reinforcement value of those outcomes, computing what economists refer to as subjective value or utility. Studies of the neural responses to rewards and their evaluation have generally implicated neural circuits that use the neurotransmitter **dopamine**. This important neurotransmitter contributes to many cognitive functions. As discussed in the Introductory Box, damage to dopamine neurons can lead to a difficulty in controlling motor movements, as evident in Parkinson's and Huntington's diseases. Dysfunction of the dopamine system has also been linked to a number of psychiatric conditions, including schizophrenia, depression, and attention deficit hyperactivity disorder. Conversely, other neurotransmitters also contribute to reward evaluation and decision making, as will be described later in this chapter. It is important, therefore, to understand the specific roles that dopamine neurons play in reward processing—while also recognizing that those roles are not always unique to dopamine or limited only to the processing of rewards.

Two important structures in the midbrain contain dopamine neurons: the substantia nigra and the **ventral tegmental area**, or **VTA** (**Figure 14.3**). Neurons in the substantia nigra contribute to motor control through their interconnections with the basal ganglia (see Chapter 5). Neurons in the VTA contribute to reward evaluation through their projections to the **nucleus accumbens** in the basal ganglia (sometimes referred to as the *ventral striatum*), the amygdala, and the hippocampus, as well as to cortical regions like the medial frontal lobe.

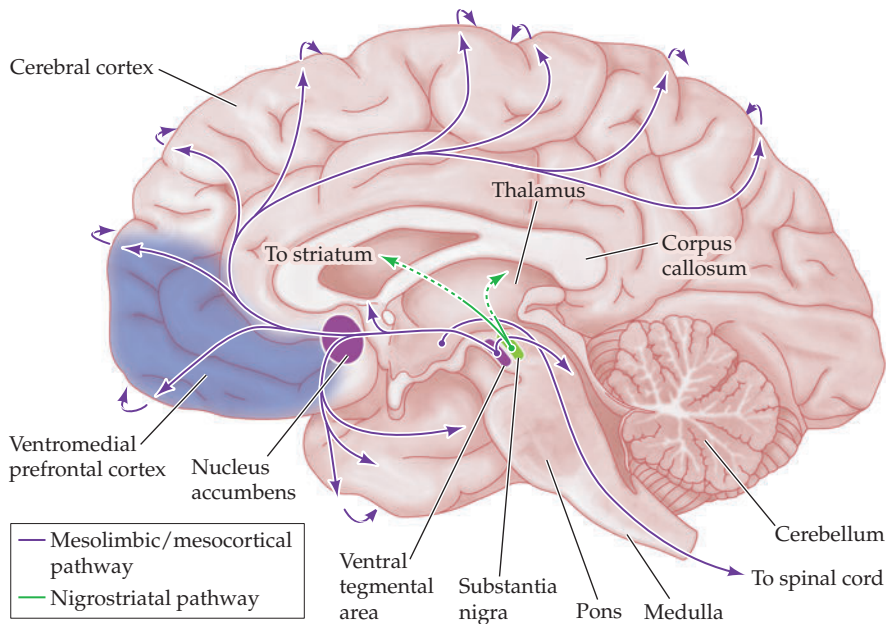


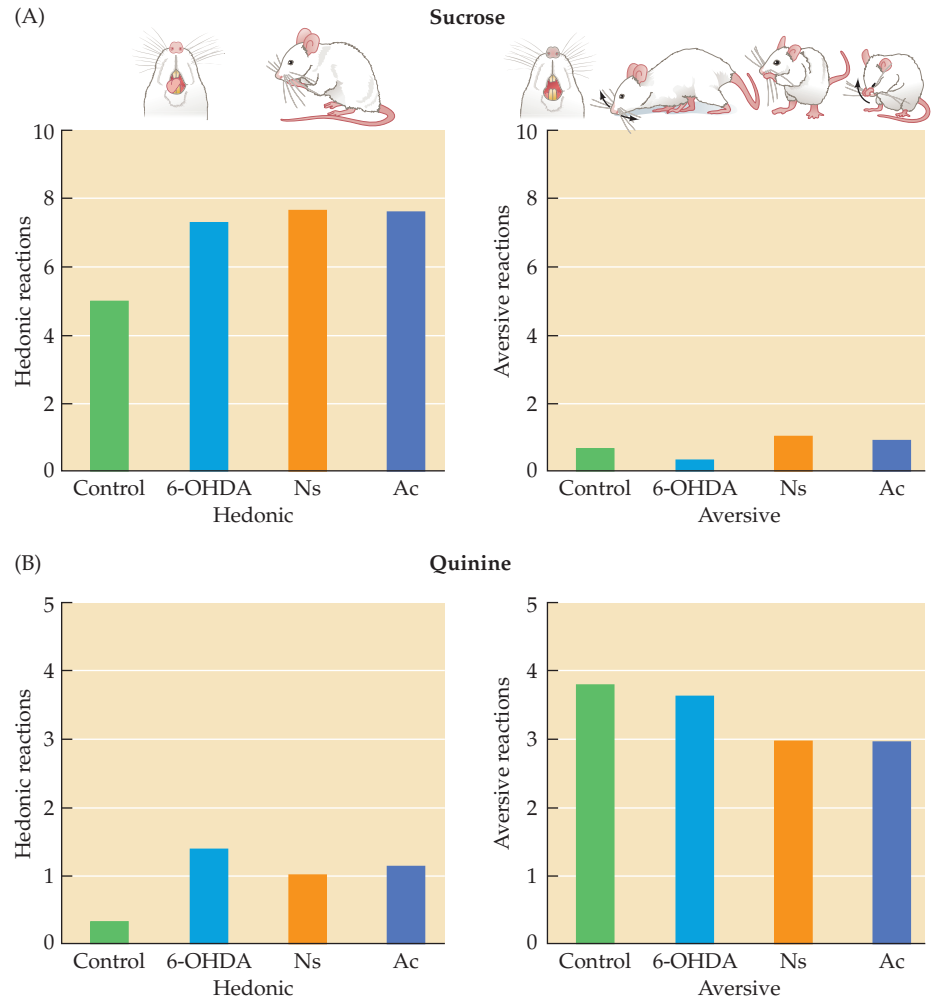
Figure 14.3 Dopaminergic pathways The activity of neurons in the substantia nigra and ventral tegmental area of the midbrain is modulated by the neurotransmitter dopamine. These neurons project broadly throughout the brain, notably to the nucleus accumbens, other regions of the basal ganglia, and the ventral and medial parts of the prefrontal cortex.

Several lines of evidence have linked dopamine neurons in these circuits to reward processing. Particularly notable have been discoveries that essentially all drugs of abuse—with the possible exception of some psychedelic compounds like LSD—exert their addictive influence through alterations in the function of these dopamine neurons. Moreover, direct stimulation of dopamine neurons serves as a particularly potent reward. In the 1950s, neurophysiologists James Olds and Peter Milner implanted electrodes in the brains of rats along a white matter tract called the medial forebrain bundle. While this tract contains axons from a diverse array of brain regions, an important component comes from dopamine neurons in the brainstem that project onto the nucleus accumbens. If the rats were given a stimulating current from these electrodes each time they performed a particular behavior (e.g., sitting in a specific spot in their cage), they would perform that behavior repeatedly to the exclusion of everything else—neither eating or drinking, nor engaging in sex—potentially until death. On the basis of this intriguing evidence, dopamine came to be popularly known as the “pleasure chemical,” a substance responsible for the feelings of liking and pleasure that accompany rewards as diverse as drugs, sex, and money.

This simple story—that dopamine signals pleasure—is incorrect. In a seminal study, psychologist Kent Berridge and his colleagues injected into the nucleus accumbens a chemical that was toxic for dopamine neurons, causing broad damage to the mesolimbic dopamine pathways. Without dopamine function, the rats became aphagic and would not eat even if starving. Yet these same rats exhibited normal reactions to pleasurable and aversive tastes (**Figure 14.4**). A sweet taste (sucrose solution delivered into the rats’ mouths) generated tongue protrusions and licking, whereas a bitter taste (a quinine solution) led to mouth gaping and blocking the throat with the tongue. From these and similar results, Berridge argued that dopamine supported not liking, but *wanting*, or the motivation to pursue a reward. (This property has also been labeled *seeking*, to emphasize that dopamine pushes an organism into action.)

Figure 14.4 Dopamine neurons are necessary for wanting, not liking

Rats treated with the neurotoxin 6-hydroxydopamine (6-OHDA) show dramatic decreases in the number of dopaminergic neurons within the striatum, compared to control animals. For example, in the experiment whose data are shown here, administration of 6-OHDA resulted in greater than 90 percent dopamine depletion in all rats (6-OHDA group), with subsamples showing greater than 98 percent depletion in the striatum (Ns) and greater than 99 percent depletion in the nucleus accumbens (Ac). These rats became aphagic, such that they would not seek out and eat food that was freely available in their cage. However, they showed normal hedonic or “liking” reactions (e.g., tongue protrusions and paw licking) to a sucrose solution delivered directly to their mouths (A). They also showed normal aversive reactions (e.g., gaping mouths, chin rubbing, and face washing) when that sugar water was made bitter by the addition of quinine (B). (After Berridge and Robinson 1998.)



This perspective helped make sense of a counterintuitive feature of the stimulation research described earlier: even though the stimulation of this fiber tract seemed to evoke sensations of pleasure in the rat, dopamine release was not observed. Only when stimulation was randomly and unexpectedly delivered by the experimenters was dopamine release observed. It is now thought that sensations of pleasure accompany the activity of neurons sensitive to opioids and endocannabinoids, on the basis of evidence that direct delivery of those neurotransmitters to the ventral striatum (and other regions) increases the intensity of reactions associated with liking.

More recent work using functional neuroimaging has shown that activation in the nucleus accumbens is elicited by a wide range of motivationally relevant stimuli, including juice, money, social signals, donations to a favorite charity, and even the costly punishment of a noncooperative individual. Common to all of these cases is the reinforcement of a desirable association, whether learning about rewards or potentiating behavior. Conversely, any breakdown of motivation is accompanied by a diminishment of activation in the dopamine system.

In the phenomenon of reward undermining, an external reward actually reduces someone's engagement in a task. Participants in an fMRI study played a simple game in which they tried to stop a timer at exactly a target number. Subjects found this game rewarding and tried to perform well, even if they were receiving no explicit rewards—as shown by robust activation in the ventral

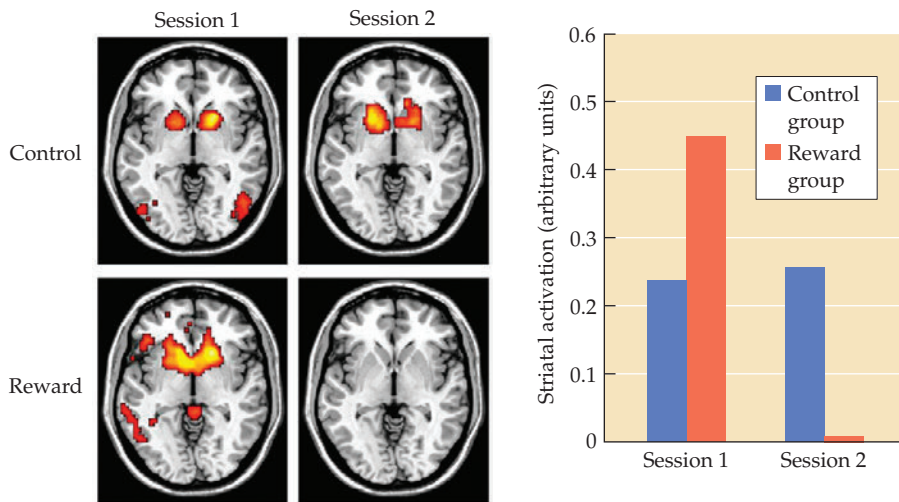


Figure 14.5 Reward undermining Economists have long recognized the possibility of reward undermining or “motivational crowding out,” in which the delivery of a reward eliminates the intrinsic motivation to perform a task. In this fMRI experiment, two groups of participants played a simple stopwatch-timing game. The control group was not paid for playing the game, but successful plays in the game still evoked activation in the ventral striatum in each of two sessions. The reward group received monetary rewards for good performance in its first session of the game, during which time robust striatal activation was observed. The monetary rewards ceased in the second session, however, and the striatal activation was abolished. (After Murayama et al. 2010.)

striatum in a control group of unpaid participants across two sessions (Figure 14.5). Members of another experimental group were paid for their performance in the first session, and successful trials that resulted in payment led to activation in the ventral striatum. But when the experimental group was then told that there would be no monetary rewards in the second session, the activation in the ventral striatum disappeared. The participants still performed the task, but they now lacked any intrinsic motivation for performing well.

Reward prediction error

The previous section set forth the now-standard perspective that some midbrain dopamine neurons and their striatal targets support motivational processes, specifically links between actions and rewards. Left unanswered, however, is what information these dopamine neurons actually encode. Do they encode the value of the reward, the value of an action, or something else? Over the past 15 years, studies have converged to support a single and perhaps counterintuitive theory: phasic changes in the activity of dopamine neurons do not signal rewards themselves, but instead signal *changes in information*.

In a series of experiments, neurophysiologist Wolfram Schultz and his colleagues investigated how VTA neurons change their firing rate in response to rewarding stimuli. Monkeys with electrodes implanted in the VTA learned to press a lever when they saw a light, whereupon they would receive a squirt of juice (Figure 14.6). At the beginning of training, when the monkey was naïve about the experimental task, the delivery of the juice reward evoked a significant increase in VTA activity. As the monkey learned the task, however, the VTA neurons fired less and less to the juice reward and began firing more and more to the light. With learning,

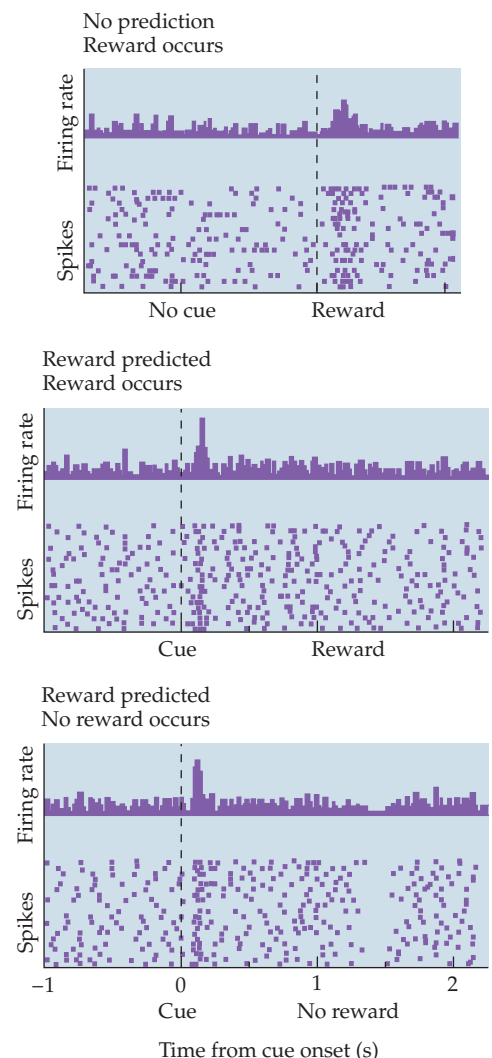


Figure 14.6 Activity of midbrain dopaminergic neurons in response to expected and omitted rewards Neurons in the ventral tegmental area (VTA) increase their firing rate in response either to unexpected rewards (top panel) or to cues that predict future rewards (middle panel), and show no change in response when a reward is delivered as expected (middle panel). However, they decrease their firing rates when a reward is expected but does not occur (bottom panel). (After Schultz et al. 1997.)

Reward prediction error (temporal difference)

Discount factor

Predicted reward at this time

$$\delta(t) = r(t) + \gamma \hat{V}(t + 1) - \hat{V}(t)$$

Actual reward at this time

Predicted reward in the future

Figure 14.7 Reward prediction error Computational models of reward learning have been used to explain the signals carried by dopamine neurons. The most influential model contends that dopamine neurons carry a signal called *reward prediction error (RPE)*, which combines three quantities: the actual reward received at a given point in time, new information gained about future rewards, and the reward that was predicted to be received. The RPE signal shares the essential properties of dopamine neurons (e.g., as shown in Figure 14.6), including increased firing to cues that provide new information about future rewards and no change in response to rewards that are fully predicted. (Equation from Schultz et al. 1997.)

therefore, the response of these particular dopamine neurons transferred from an unconditioned reinforcer to a conditioned stimulus that predicted that reinforcer. Moreover, if an expected reward did not occur, then the activity of the neurons decreased below baseline firing rates.

Subsequent studies by the same group suggested that the activation of VTA neurons is proportional both to the probability of reinforcement—with lower-probability rewards evoking greater activity—and to the magnitude of reinforcement. These results have since been replicated using functional neuroimaging studies in human participants, for whom activation in the ventral tegmental area (and its projection targets) increases in response to unpredictable rewards compared to predictable rewards.

Together, these studies indicate that dopamine neuron activity—at least in a subset of dopamine neurons—provides an index of **reward prediction error (RPE)**, which can be defined through a simple equation that combines two basic factors: (1) the difference between the reward that was received compared to what was expected, and (2) the change in information about future rewards. Note that the value of information depends on when a reward is expected (Figure 14.7). If a reward will not occur for a long time, then new information about that reward has less effect on midbrain dopamine neuron signals (for a similar phenomenon, see the discussion on temporal discounting later in the chapter). The reward prediction error signal can be used to guide behavior by means of **temporal difference learning**. The key idea in temporal difference learning is that successive states of the world (and, in turn, predictions about those states) are correlated over time.

As foreshadowed earlier in the chapter, the reward prediction error signal is substantially similar to the idea of reference dependence in prospect theory, in that both calculate subjective value in terms of transient changes from a current state, not absolute levels of reward. It is computationally efficient to track only unexpected changes in the world rather than the entire history of events over a long period of time. By taking advantage of the fast temporal resolution of EEG signals, very rapid brain responses can be identified that track unexpected reward outcomes (Figure 14.8). These rapid responses, which have been localized to the dorsomedial prefrontal cortex around the anterior cingulate gyrus, may in turn engage executive control processes to address ongoing task demands (see Chapter 13).

As described already, midbrain dopamine neurons project to many regions, most notably the basal ganglia (e.g., nucleus accumbens) and the ventromedial prefrontal cortex. An intriguing possibility is that information about rewards and punishments might modulate processing essentially everywhere in the brain, shaping all of the other cognitive functions considered in this textbook. Support for the idea of distributed reward processing comes from multivoxel pattern analysis (see Chapter 2) of fMRI data collected while human participants played a simple economic game (Figure 14.9). This analytic technique showed that classifiers that could distinguish winning and losing outcomes could be constructed from voxels in nearly all brain regions, including both primary

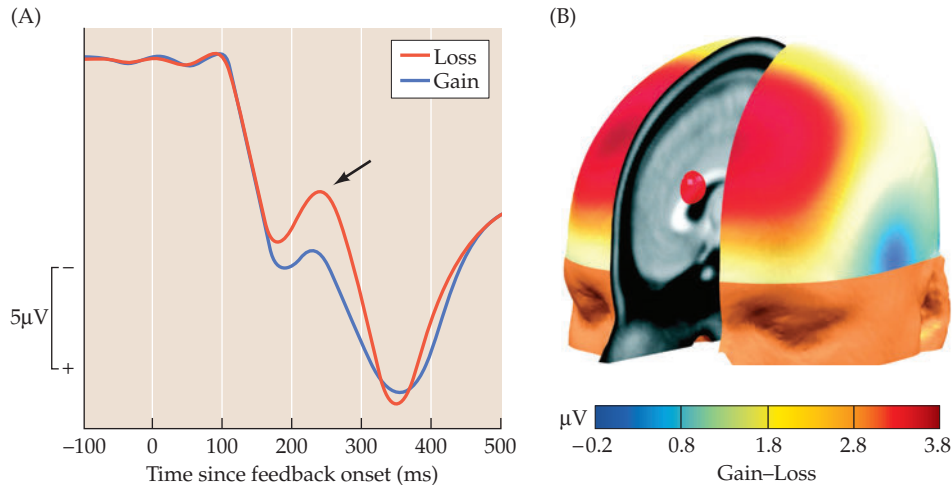


Figure 14.8 Rapid prefrontal cortex responses to monetary gains and losses

(A) By measuring brain functioning using EEG, cognitive neuroscientists can identify very rapid changes in brain activity associated with winning or losing money. Feedback about an unexpected monetary loss generates a negative-polarity ERP component within 200 milliseconds (black arrow in figure). (B) Source localization methods have linked this component to a dipole within the anterior cingulate gyrus, a region that contributes to the engagement of control systems within the brain, as introduced in Chapter 13. μV = microvolts. (After Gehring and Willoughby 2002.)

sensory regions and parts of prefrontal cortex. Similarly, studies using single-unit recording have revealed neurons sensitive to reward throughout much of the brain. While it seems clear that the midbrain dopamine system plays a critical role in the processing of reward, much remains to be understood about how other regions contribute to reward processing, especially in the context of decision making (**Box 14A**).

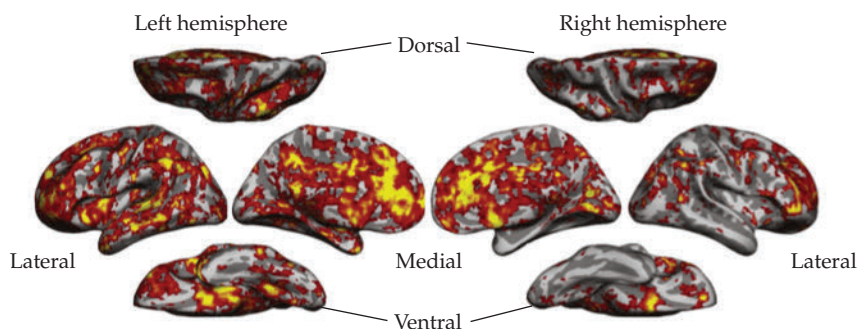


Figure 14.9 Distributed influence of reward information throughout the brain

Researchers used multivoxel pattern analysis to identify brain regions where the local pattern of fMRI activation could predict whether a participant won or lost a simple competitive game (here, “matching pennies”). The reward outcome could be decoded from sets of voxels in nearly all regions, including both cortical and subcortical structures, suggesting that reward information modulates neuronal activity throughout the brain. Shown in the red-to-yellow color map are voxels contributing significant classification power (at a corrected $p < 0.001$). Of note, when the significance threshold was relaxed to a $p < 0.05$ threshold, more than 80 percent of the voxels in the brain passed significance testing. (After Vickery et al. 2011.)

BOX 14A LEARNING VALUES AND FORMING HABITS

Decision making is often considered to be an active, controlled process. When we think about the decisions that we have made in the past, the examples that leap most readily to mind typically involve conscious reflection about consequences and goals (e.g., selecting a college). Many of the examples used in this chapter do involve such goal-directed decisions; the fMRI activation shown in Figure 14.13, for example, represents the value associated with self-controlled choices of monetary rewards that may be delayed for several months before their receipt. Yet goal-directed choices represent only one facet of decision making. Many other important behaviors involve ongoing learning processes that relate value information to actions, often in the absence of reflection on our goals and desires.

Particularly relevant to day-to-day life are **habits**, or action tendencies that develop over time because of repeated reinforcement. Habits tend to involve relatively slow learning, but when established, they can be very resistant to change. Habits can be relatively innocuous, as when a student always sits in the same seat in the classroom. Or, in the case of a chronic drug abuser, they can be manifested in destructive, life-altering consequences. Two processes thought to be critical for reward learning and the development of habits are the ability to associate stimuli with rewards and the ability to choose motor actions that maximize future outcomes. Because of evidence implicating the basal ganglia both in stimulus-reward associations and in adaptive control of motor behavior, considerable research has been targeted at understanding how the basal ganglia's component regions may interact in reward learning and habit formation.

The basic features of reward learning can be seen in what are often called **actor-critic learning models**. Such models assume two brain systems: a *critic* that



Real and fictive errors in a simple stock market game. (A) Participants in this fMRI experiment made a series of investment decisions within virtual stock markets. At each point in time (t), the participant could choose to allocate money in any proportion of stocks (i.e., betting that the market will go up) and cash (i.e., betting that the market will go down). This task generates two types of reward prediction errors: *real* RPEs based on the money invested in stocks, and *fictive* RPEs based on the money not so invested. For example, if all money was allocated to cash and then the market rose sharply, there would be a substantial positive fictive error. (After Lohrenz et al. 2007.)

evaluates whether rewards are better or worse than expected, and an *actor* that updates the values of potential courses of behavior to increase the likelihood of future rewards. These two roles have

been linked to separate components of the basal ganglia. Neurons in the ventral striatum (e.g., in the nucleus accumbens) provide critic signals associated with reward prediction error (RPE), via input

Responses to negative outcomes

To what degree, then, is the midbrain dopamine system specifically related to positive outcomes? Nearly all neuroeconomic experiments deliver or withhold some desirable outcome. In a typical task used with human subjects, winning money evokes an increase in fMRI activation in the nucleus accumbens, compared to not winning money. Similarly, tasks used with non-human primates show an increase in firing in response to the delivery of a juice reward compared to the unexpected omission of that reward. In contrast, almost no