Learning and Memory

From Brain to Behavior

THIRD EDITION

Mark A. Gluck
Rutgers University–Newark

Eduardo Mercado
University at Buffalo, The State University of New York

Catherine E. Myers
Department of Veterans Affairs, VA New Jersey Health Care System,
and Rutgers University–New Jersey Medical School

worth publishers
Macmillan Learning
New York

can become reinforcing. If you have been studying for several hours straight, the idea of "taking a break" to clean your room or do the laundry can begin to look downright attractive. If so, you've experienced the Premack principle at work.

Interim Summary

- In operant conditioning, organisms learn to make responses under particular conditions in order to obtain or avoid outcomes: Discriminative stimulus S^D → Response R → Outcome O.
- In operant conditioning, the outcome (reinforcement or punishment) occurs only if the organism makes the response. In classical conditioning, by contrast, the unconditioned stimulus (US) occurs whether or not the organism makes a conditioned response (CR).
- Discriminative stimuli signal to the organism whether a particular response will result in a particular outcome.
- An outcome that an organism will work to obtain is called a reinforcer; an outcome that an organism will work to avoid is called a punisher. While punishment can be effective in eliminating an undesired response, it leads to more varied behavior and can be undermined by discriminative stimuli that encourage cheating, by concurrent reinforcement, or by weakness of the initial punisher. Another approach to eliminating unwanted behavior is differential reinforcement of alternative behaviors (DRA).
- Complex responses can be trained via shaping, in which progressive approximations to the desired response are reinforced, and chaining, in which organisms are gradually trained to execute a sequence of responses.
- The four basic types of operant paradigm are positive reinforcement, negative reinforcement, positive punishment, and negative punishment. The words positive and negative denote whether the outcome is added or subtracted; *reinforcement* and *punishment* denote whether the response increases or decreases as a result of learning.
- Schedules of reinforcement define whether the outcome O follows every response R, is available after some (fixed or variable) number of responses, or is available only after some (fixed or variable) time interval.
- When multiple responses are reinforced under a VI schedule, the matching law predicts that organisms will allocate time among those responses based on the relative rates of reinforcement for each response.
- Behavioral economics is the study of how organisms choose to allocate their time and resources among various responses that result in different outcomes. The bliss point is the particular allocation of resources that provides maximal subjective value to an individual.
- The Premack principle states that the opportunity to perform a highly frequent behavior can reinforce performance of a less-frequent behavior. The response deprivation hypothesis states that any behavior can be reinforcing if the opportunity to perform that behavior is restricted.

5.2 Brain Substrates

The previous section defined operant conditioning as learning an association between a discriminative stimulus S^D, a response R, and an outcome O. In studying such associations, neuroscientists are discovering that the parts of the

brain that link stimuli with responses ($S^D \rightarrow R$ learning) are different from the parts of the brain that learn about the expected outcomes (O) of those responses. While many brain areas play a role in these processes, two key areas are the dorsal striatum, which appears to be particularly important for $S^D \rightarrow R$ learning, and the orbitofrontal cortex, which appears important for learning about expected outcomes. Different brain areas may help us evaluate whether those outcomes are reinforcers or punishers.

The Dorsal Striatum and Stimulus–Response ($S^D \rightarrow R$) Learning

Voluntary motor responses occur when neurons in the motor cortex send messages to motor neurons in the muscles that control movements. The motor cortex receives its primary inputs from cortical areas that process sensory information, such as the visual cortex (V1) and the somatosensory cortex (S1), which you saw back in Figure 2.7, and also from the frontal cortex. Thus, when you see a book, this visual stimulus is registered by your visual cortex. If you decide to pick up the book, this "decision" is made in your frontal cortex, and signals from both the visual cortex and the frontal cortex travel to motor cortex, which integrates these signals and produces the appropriate instructions, resulting in your picking up the book.

Information from the sensory cortex to the motor cortex can also travel via an indirect route, through the **basal ganglia** (colored purple in Figure 5.8). The basal ganglia are a collection of *ganglia* (clusters of neurons) that lie at the base of the forebrain. One part of the basal ganglia is the **dorsal striatum** (Figure 5.8), which can be further subdivided into the *caudate nucleus* and the *putamen*. The dorsal striatum receives highly processed stimulus information from sensory cortical areas and projects to the motor cortex, which produces a behavioral response.

basal ganglia. A brain region that lies at the base of the forebrain and includes the dorsal striatum.

dorsal striatum. A region of the basal ganglia that is important for stimulus–response learning.

The dorsal striatum plays a critical role in operant conditioning, particularly if discriminative stimuli are involved. Rats with lesions of the dorsal striatum can learn operant responses (e.g., when placed in a Skinner box, lever-press R to obtain food O). But if discriminative stimuli are added (e.g., lever-press R is reinforced only in the presence of a light S^D), then the lesioned rats are markedly impaired (Featherstone & McDonald, 2004). In humans, too, individuals with damage or disruption to the striatum due to Parkinson's disease or Huntington's disease show deficits in the ability to associate a discriminative stimulus with a correct response (Ashby & Waldron, 2000; Robbins, 1996). In short, the dorsal striatum appears necessary for learning $S^D \rightarrow R$ associations based on feedback about reinforcement and punishment (McDonald & White, 1994; O'Doherty et al., 2004).

$S^D \rightarrow R$ associations that depend on the dorsal striatum tend to be relatively automatic or habitual (Balleine, Daw, & O'Doherty, 2008). Remember the well-trained rats, discussed earlier in this chapter, who would run right through a pile of food on their way to a goal box in the maze? That behavior probably reflects $S^D \rightarrow R$ learning in the striatum, making the maze-running automatic even when other behaviors (such as pausing to eat) would have resulted in reward. In this case, running is based on a history of learning in which that response resulted in desirable outcomes; but after a long period of training, the response is performed even though the outcome is no longer contingent on that action.

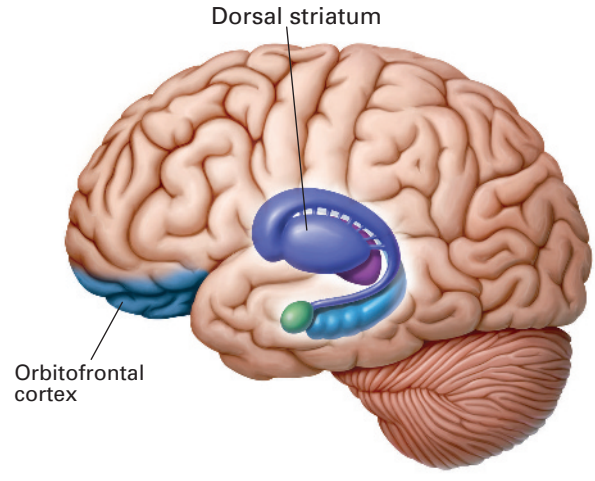

Figure 5.8 Some brain substrates of operant conditioning During operant conditioning, the dorsal striatum may help create links between the sensory cortex and the motor cortex so that stimuli can evoke appropriate motor responses ($S^D \rightarrow R$ learning). Parts of the frontal cortex, including the orbitofrontal cortex, may play a role in learning that specific responses lead to particular outcomes.

The Orbitofrontal Cortex and Learning to Predict Outcomes

$S^D \rightarrow R$ learning is, of course, only half the picture in operant conditioning. Organisms learn to predict that particular responses R (in the presence of S^D) will result in particular outcomes O. For example, you read about the negative contrast effect in the Behavioral Processes section above: monkeys may shriek in annoyance if their response earns them a less-preferred food than the one they expected, and trick-or-treaters may feel cheated if they receive pennies rather than the expected candy. Such results show that organisms don't make responses blindly but make them in anticipation of particular outcomes.

orbitofrontal cortex. An area of the prefrontal cortex that is important for learning to predict the outcomes of particular responses.

Several brain areas appear to be involved in learning to predict the outcomes of behavior. Among these are parts of the *prefrontal cortex*, including the **orbitofrontal cortex**, which lies at the underside of the front of the brain in primates (Figure 5.8), and which appears to contribute to goal-directed behavior by representing predicted outcomes (Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009; Tanaka, Balleine, & O'Doherty, 2008). The orbitofrontal cortex receives inputs conveying the full range of sensory modalities (sight, touch, sound, etc.) and also visceral sensations (including hunger and thirst), allowing this brain area to integrate many types of information; outputs from the orbitofrontal cortex travel to the striatum, where they can help determine which motor responses are executed.

Evidence that the orbitofrontal cortex plays a role in predicting the outcome of responses comes from neuronal recordings. For example, thirsty rats can be trained on a discrimination task in where the discriminative stimuli are two odors, the response R is to poke the nose into a nearby water cup, and the two possible outcomes are a tasty sucrose solution or a bitter quinine solution:

Odor 1 → R → (delay) → sucrose (reward)

Odor 2 → R → (delay) → quinine (punisher)

Here, a short delay (typically less than a second) is introduced between the response and the outcome, during which period the animal is "expecting" the outcome. During this delay, some neurons in orbitofrontal cortex fire differently, depending on whether a reward or punisher is expected (Schoenbaum, Chiba, & Gallagher, 1998). Figure 5.9a shows an example of the firing patterns of one neuron

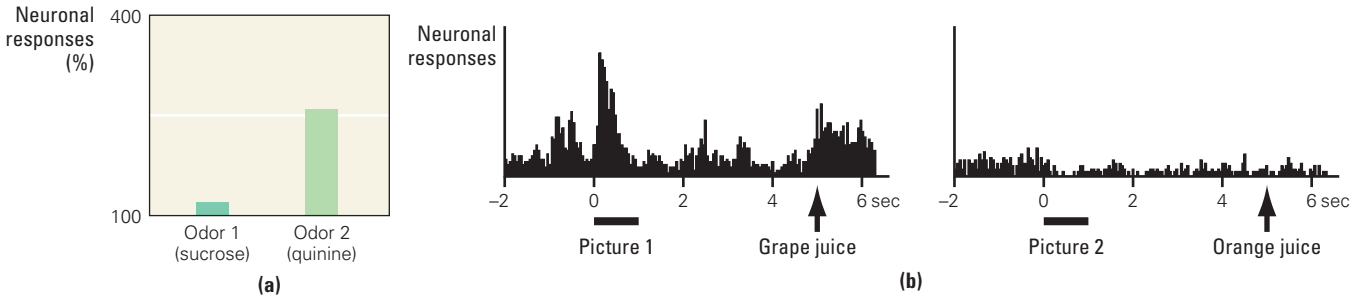

Figure 5.9 Orbitofrontal neurons code expected outcomes (a) Responses of a single neuron in orbitofrontal cortex of a rat learning that responding to odor 1 produces sucrose while responding to odor 2 produces quinine. Responses are shown during the delay between response and outcome. This neuron shows very little increase above baseline (100%) after the rat has responded to odor 1, but a strong increase after the rat has (mistakenly) responded to odor 2. Thus this neuron appears to code for expectation of quinine. (b) A single neuron in orbitofrontal cortex of a monkey trained that some pictures (e.g., Picture 1) predict that a response will be reinforced by grape juice but other pictures (e.g. Picture 2) predict that the response will be rewarded by orange juice. Each picture is presented for 1 second, with juice delivered 4 seconds later. This neuron responds strongly (height of bars indicate neuronal activity) when Picture 1 (which predicts grape juice) appears, and again when grape juice is delivered, but the same neuron does not fire to Picture 2 (which predicts orange juice) nor when orange juice is delivered. Thus, this neuron codes not only expectation of reward but also expectation of a specific outcome: grape juice.

(a) Data from Schoenbaum et al., 1998, Figure 3a. (b) Information from Tremblay & Schultz, 1999, Figure 3b.

in the orbitofrontal cortex of a rat learning such a task. This particular neuron fires strongly if the rat has just made a (mistaken) response to odor 2 and is expecting quinine, but less strongly if the rat has just made a response to odor 1 and is expecting sucrose. Thus, this neuron appears to code expectation of the punisher rather than the reward. If the contingencies are reversed, so that odor 1 now predicts quinine and odor 2 now predicts sucrose, such neurons often alter their responses to reflect the new contingencies (Stalnaker, Franz, Singh, & Schoenbaum, 2007).

Neurons in orbitofrontal cortex don't only learn whether to expect reinforcement or punishment; they even appear to code the actual identity of the expected outcome. Thus, monkeys can be trained with a set of pictures that predict whether the upcoming reward will be grape juice or orange juice. Figure 5.9b shows the responses of a single neuron that became active whenever a stimulus that predicted grape juice was presented, but not when pictures predicting orange juice were presented (Tremblay & Schultz, 1999). This same neuron also fired during the actual delivery of grape juice, but not orange juice.

Given their ability to encode specific predicted outcomes, orbitofrontal cortex neurons play an important role in helping us select between potential actions based on their expected consequences. When monkeys are trained to choose between two responses that result in different outcomes (say, water vs. Kool-Aid), most individual monkeys have a preference for one beverage over another. Given a choice between licking to obtain water or Kool-Aid, a particular monkey will alter his responses based on the amount of each he will get. If response R1 produces 1 cc of water and response R2 produces 1 cc of Kool-Aid, he may choose to lick to obtain the sugary Kool-Aid. But if R1 produces 6 cc of water, he may choose that option instead. In fact, the likelihood that the monkey will make response R1 depends on the trade-off between his individual preference of beverage and the relative amount of each he would obtain, resulting in choice behavior very much like the pigeon allocating pecks between key A and key B. Neurons in the monkey's orbitofrontal cortex respond with a strength proportional to the perceived value of each choice (Padoa-Schioppa & Assad, 2006).

Remember college student Jamie, who could spend his weekly income by distributing it among choices such as music purchases and restaurant dinners? Possibly, neurons in Jamie's orbitofrontal cortex were helping him to evaluate the potential outcomes of his actions, and to choose between them. When dinners were cheap, certain of these neurons may have responded strongly, indicating that dinners were the preferred choice. But when the cost of dining out rose, the same neurons may have responded more weakly, leading Jamie to prefer the opposite alternative for spending his money.

Mechanisms of Reinforcement Signaling in the Brain

The previous section suggested that neurons in the orbitofrontal cortex code not only the identity of an outcome (e.g., grape juice vs. orange juice) but also whether that outcome is reinforcing or not (e.g., water flavored with sucrose vs. quinine). This distinction is critical: if an outcome is reinforcing, the $S^D \rightarrow R$ association should be strengthened, increasing the likelihood that S^D evokes R in the future; if it is a punisher, the association should be weakened, decreasing the likelihood of R. How does the brain determine whether an outcome is a reinforcer or a punisher?

"Wanting" and "Liking" in the Brain

In 1954, James Olds was experimenting with delivering electrical stimulation to the rat brain. He inserted an electrode into an area that researchers now believe to have been the lateral hypothalamus. Olds waited until the rat wandered into one corner of the experimental chamber, and then he applied a brief electrical current. After a few minutes of wandering around the chamber, the rat

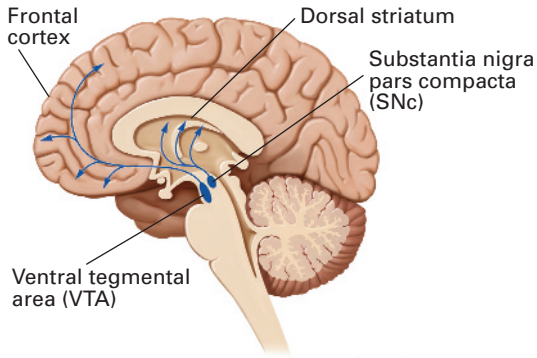

Figure 5.10 The ventral tegmental area (VTA) and substantia nigra pars compacta (SNc) The VTA (part of the midbrain) and the SNc (a part of the basal ganglia) are small regions containing neurons that project dopamine to many brain areas, including the dorsal striatum and frontal cortex.

came back to the same corner, where Olds gave it a second stimulation. The rat caught on quickly, and began to loiter in that corner of the chamber, apparently hoping for more electrical stimulation (Olds, 1955). Thus, electrical stimulation to this area of the brain seemed to be acting much like a reinforcer: increasing the probability of certain responses (in this case, hanging around the correct location).

Olds was intrigued, to say the least. He rigged a Skinner box so that the rats could press a lever to turn on the electrical stimulation. The rats were soon lever-pressing at a furious rate: as many as 700 times an hour (Olds, 1958). If allowed, rats would press the lever continuously for up to 48 hours, until they collapsed from physical exhaustion! Given a choice between electrical stimulation and food, the rats would literally starve themselves, preferring the stimulation (Routtenberg & Lindy, 1965).

Later studies identified that rats would work for electrical stimulation in several brain areas, including the **ventral tegmental area (VTA),** a small region in the midbrain of rats, humans, and other mammals (Figure 5.10). The electrodes in Olds's original studies were probably stimulating hypothalamic neurons that project to the VTA, so that the electrical current was indirectly activating this area. Because VTA stimulation was such a powerful reinforcer, some researchers inferred that the rats "liked" the stimulation, and the VTA and other areas of the brain where electrical stimulation was effective became informally known as "pleasure centers."

ventral tegmental area (VTA). A region in the midbrain that contains dopamine-producing neurons which project to the frontal cortex and other brain areas.

However, the idea of "pleasure centers" is something of an oversimplification. For one thing, rats lever pressing for electrical brain stimulation don't tend to act as if they're enjoying it; they tend to become agitated and may bite the lever instead of simply pressing it, or even scratch the walls or show other behaviors such as eating, fighting, or shredding of nesting material. This is more like the behavior of an excited animal than one who is enjoying food. Skinner, of course, would caution that we can't infer what an animal might be feeling just by watching its behaviors. Nevertheless, some researchers have suggested that electrical brain stimulation causes not pleasure but rather excitement or anticipation of reinforcement—much like the anticipation we experience when expecting a good meal or a big present (Flynn, 1972).

Currently, many researchers believe that we have separate brain systems for signaling **hedonic value—**meaning the subjective "goodness" of a reinforcer, or how much we "like" it—that are distinct from those signaling **motivational value**—meaning how much we "want" a reinforcer and how hard we are willing to work to obtain it. No matter how much we may "like" chocolate cake, most of us will not be very motivated to obtain more if we have just eaten three slices; similarly, Olds's rats doubtless still "liked" food and rest, but they were more motivated to obtain electric brain stimulation, even when starving and exhausted. In these examples, provision of a "liked" reinforcer isn't enough to evoke responding. Only when "wanting" and "liking" signals are both present will the arrival of the reinforcer evoke responding and strengthen the $S^D \rightarrow R$ association.

hedonic value. The subjective "goodness" or value of a reinforcer.

motivational value of a stimulus. The degree to which an organism is willing to work to obtain access to that stimulus.

Dopamine: How the Brain Signals "Wanting"?

The neurotransmitter dopamine is produced by neurons in several areas of the brain, including the ventral tegmental area (VTA), which projects to the frontal cortex (among other places), and also including the nearby **substantia nigra pars compacta** (SNc), which is a part of the basal ganglia that projects to the striatum (Figure 5.10). As you read above, the dorsal striatum is an important site of $S^D \rightarrow R$ association, and the orbitofrontal cortex (and other frontal areas) is important for

substantia nigra pars compacta (SNc). A part of the basal ganglia that contains dopamine-producing neurons which project to the striatum.

learning about predicted outcomes, so dopaminergic neurons in the VTA/SNc are a good place to start looking at how the brain signals motivational value.

In rats, dopamine release from the VTA/SNc is triggered by encounters with food, sex, drugs of abuse, and secondary reinforcers. In humans, PET and fMRI studies have shown that presentation of juice, cocaine, money, humor, and even video games causes heightened activity in dopamine target sites such as the striatum (Berridge & Robinson, 1998; Knutson, Fong, Adams, Varner, & Hommre, 2001; Mobbs, Greicius, Abdel-Azim, Menon, & Reiss, 2003). Even in invertebrates, such as the sea slug *Aplysia*, dopamine is released in conjunction with positive reinforcement during operant conditioning (Brembs, 2003; Nargeot, Baxter, Patterson, & Byrne, 1999).

Most researchers believe that dopamine does not simply signal hedonic value or "liking." For example, Parkinson's disease damages dopamine-producing neurons that project to the striatum. But when patients with Parkinson's disease are asked to rate the perceived pleasantness of sweet and salty tastes, their ratings are the same as those of healthy people. Apparently, the dopamine reduction in these patients causes no loss of the ability to "like" pleasurable stimuli (Travers et al., 1993).

Similar results are obtained from non-human animals. Researchers can't simply ask rats to rate the perceived pleasantness of different tastes. But researchers can infer degree of liking by watching the animals' reactions. When a sweet substance is placed in a rat's mouth, the animal shows a recognizable cluster of responses that include rhythmic movements of the mouth and protrusion of the tongue. This is sometimes called the hedonic or "yum" reaction. A bitter taste produces a different cluster of responses: gapes, shakes of the head, and wiping of the face with paws (the aversive or "ugh" reaction). Rats given injections of a drug that destroys dopaminergic neurons exhibit hedonic and aversive responses that are just as strong as or stronger than those of control rats (Berridge & Robinson, 1998). This suggests that rats with damaged dopamine systems continue to "like" and "dislike" food just as much as control rats do. What seems to change is their willingness to work for it.

The **incentive salience hypothesis** of dopamine function states that the role of dopamine in operant conditioning is to signal how much the animal "wants" a particular outcome—how motivated it is to work for it. According to this hypothesis, the incentive salience of food and other reinforcers—their ability to attract attention and motivate responding—is reduced in dopamine-depleted animals (Berridge, 1996, 2007; Berridge & Robinson, 1998). Given a choice between competing alternatives, normal animals will tend to choose their preferred reinforcer, even at the cost of a little extra work. In contrast, dopamine-depleted animals are still perfectly willing to eat a preferred food if it is placed in front of them, but they are unwilling to work hard to earn it (Salamone, Arizzi, Sandoval, Cervone, & Aberman, 2002).

incentive salience hypothesis. The hypothesis that dopamine helps provide organisms with the motivation to work for reinforcement.

A good example of this is seen in experiments where rats can choose to work for food. For example, most healthy rats prefer sugar pellets to rat chow, and they will work for the pellets by lever pressing, even if chow is freely available (Figure 5.11, green bars). Rats given a dopamine antagonist also prefer sugar to rat chow, if both are freely available. But, as shown in Figure 5.11 (red bars), if they have to work for the sugar pellets by lever pressing, they mostly settle for the free chow instead (Salamone et al., 2002). Whereas animals with normal dopamine levels prefer to work to obtain their preferred food; animals with reduced dopamine prefer not to work, even if this results in inferior food.

In an even more extreme case, mice that have been genetically engineered to be completely unable to produce dopamine will not seek food, and they generally starve to death by about 20 to 30 days of age, even if pellets are placed

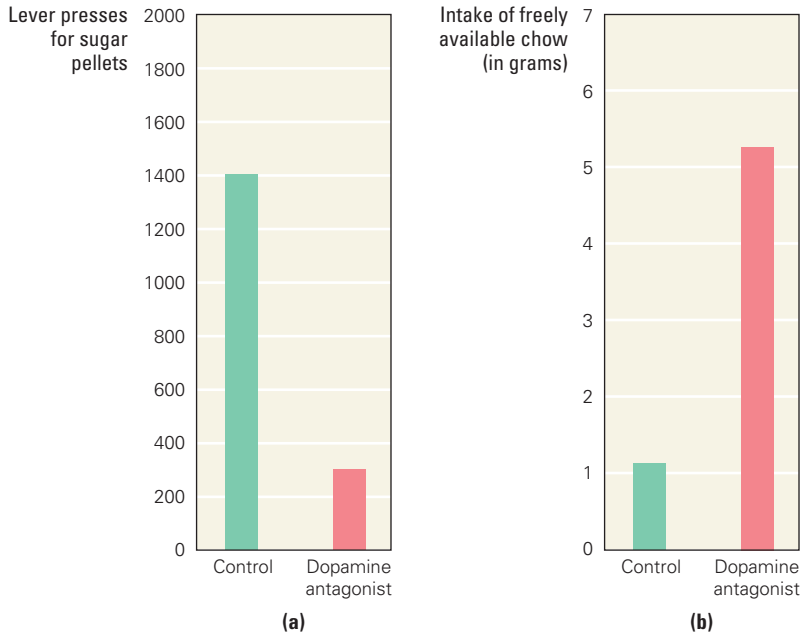

Figure 5.11 Dopamine and incentive salience If rat chow is freely available, but sugar pellets (which rats prefer) have to be "earned" by pressing a lever, control rats (green) will spend most of their time working for sugar pellets and eating relatively little free chow. Rats given a dopamine antagonist (red) are less willing to work for the sugar pellets, and instead settle for eating more of the freely available chow.
Information from Salamone et al., 2002.

directly in front of them (Palmiter, 2008). However, if the food is placed in their mouths, these animals will chew and swallow and even exhibit "yum" responses, indicating that they still "like" food and can consume it, they just lack the motivation to obtain it. These mice can be "rescued" by infecting cells in the striatum with a recombinant virus that allows the cells to produce and release dopamine; afterward, the mice eat enough normal chow to maintain body weight without further interventions.

Dopamine seems to affect incentive salience in humans, too. For example, the drug amphetamine can produce pleasurable feelings in humans, and these pleasurable feelings are not altered if the human is also given the dopamine-blocker pimozide (Brauer & de Wit, 1996, 1997). But the pimozide does suppress cravings for the amphetamine high. In other words, interfering with the dopamine system reduces "wanting" but not "liking" of amphetamine.

Conversely, increasing brain dopamine levels can increase craving. For example, in one study, humans addicted to cocaine were given the drug pergolide, which increases brain dopamine levels; the participants reported an increased craving for cocaine, but no increase in the self-reported "high" from cocaine (Haney, Foltin, & Fischman, 1998). Thus, stimulating the dopamine system increases "wanting" but not "liking" of cocaine.

The dopamine system can also be stimulated naturally by exposure to a stimulus that has previously been associated with reinforcement, increasing the "temptation power" of that stimulus (Berridge, 2012). For example, the sight of chocolate can stimulate intense desire in a chocolate-lover, even if she's not particularly hungry, while a cigarette smoker who is sincerely trying to quit may experience an overwhelming craving if he enters a room where he can see and smell others smoking. For this reason, many smokers and drug addicts who wish to quit try to stay away from environments where they are likely to encounter people using the addictive substance. Remember the concept of precommitment, discussed earlier in this chapter? Precommitment strategies can help counteract a strong $S^D \rightarrow R$ association not only by making it difficult to execute

the response but also by reducing exposure to the S^D, thus reducing the release of dopamine, which in turn helps reduce craving.

In addition to evidence that dopamine signals "wanting," there is also considerable evidence that dopamine helps strengthen learning of $S^D \rightarrow R$ associations during operant conditioning (Wickens, 2009). Although dopamine isn't required for new learning, studies show that increases in brain dopamine levels (through drug administration or by presenting a reward) do tend to enhance new $S^D \rightarrow R$ learning (Wise, 2004). Dopamine generally promotes synaptic plasticity, possibly by increasing the ability of the presynaptic neuron to activate the target neuron, and since (as you read in Chapter 2) neurons that fire together wire together, this tends in turn to strengthen the synaptic connection between those neurons (Jay, 2003). However, the effects of dopamine on neurons are notoriously variable and complicated, and a great deal remains to be clarified about this neurotransmitter and its role in learning.

Endogenous Opioids: How the Brain Signals "Liking"?

If dopamine signals "wanting," then what signals "liking" in the brain? Probably the best-studied candidate is the opioid system. Opiate receptors in the brain were discovered quite by accident in the 1970s, by researchers trying to figure out how heroin and morphine work. Heroin and morphine belong to a class of drugs called *opiates*, which bind to a class of neuronal receptors called *opiate receptors*. Rather than assume that the brain evolved special receptors to respond to heroin and morphine, researchers suspected there might be naturally occurring brain chemicals that also activate the opiate receptors. They found a class of brain chemicals, named the **endogenous opioids**, that are naturally occurring neurotransmitter-like substances (peptides) with many of the same effects as opiate drugs. (The word *endogenous* means "originating on the inside"; *opioid* means "opiate-like.") Endogenous opioids are distributed throughout the central nervous system, and when released into the body they have a wide range of effects, including lessening the normal perception of pain and producing feelings of euphoria.

endogenous opioid. Any of a group of naturally occurring neurotransmitter-like substances that have many of the same effects as opiate drugs such as heroine and morphine; may help signal hedonic value of reinforcers in the brain.

Although there is still a great deal to be learned about the endogenous opioids, many researchers believe these substances may mediate hedonic value, or "liking." If so, the reason that heroin and morphine are so intensely pleasurable could be that they happen to activate the same brain receptors as the endogenous opioids do.

For example, morphine makes sweet food taste sweeter and bitter food taste less bitter (Rideout & Parker, 1996). It can also make pain feel less painful; morphine is used medically for patients who are enduring extreme, long-term pain (in cases where the benefits of relieving suffering outweigh the risks of morphine addiction). These patients usually report that they still feel the pain but that it doesn't trouble them as much as it did before.

Endogenous opioids are released in response to primary reinforcers, such as food, water, and sex, and they may be released in response to secondary reinforcers and pleasurable behaviors, too (Le Merrer, Becker, Befort, & Kieffer, 2009). Differences in the amount of endogenous opioid released, and in the specific opiate receptors they activate, may help determine an organism's preference for one reinforcer over another (Le Merrer et al., 2009), contributing to effects such as you saw back in Figure 5.3, where infants sucked harder to obtain sweetened water, even though plain water satisfies thirst just as effectively. Just like infants, rats normally prefer sweetened to plain water, but rats given the opioid antagonist naloxone choose the sweetened water much less often than control rats (Hayward, Schaich-Borg, Pintar, & Low, 2006).

How do "Wanting" and "Liking" Interact?

Given that "wanting" seems to be signaled by dopamine, and "liking" by the endogenous opioids, and that both contribute to driving behavior, how do these two brain systems interact? The answer is not yet clear. One possibility is that some endogenous opioids may modulate dopamine release. For example, some neurons in the VTA have opiate receptors on their dendrites that, when activated, could affect those neurons' normal tendency to release dopamine. In this model, the endogenous opioids would signal "liking," which in turn would affect the VTA's ability to signal information about "wanting." But other studies have suggested that different subpopulations of dopamine neurons might exist, conveying salience ("wanting") and valence ("liking") separately (Matsumoto & Hikosaka, 2009). The picture is complicated because some drugs, such as heroin, may manipulate both pathways: activating the "liking" system to produce a pleasurable high, while also activating the "wanting" system to produce a craving for more of the drug and the high.

Punishment Signaling in the Brain

As described above, neurons in the orbitofrontal cortex code expected outcomes—including specific anticipated reinforcers and punishers—and the dopamine and opioid systems may help code "liking" (hedonic value) and "wanting" (motivational value) of reinforcers. So what codes the aversive value of punishers? So far, there doesn't appear to be just one, singular "pain center" in the brain. Rather, both physical and emotional pain can activate multiple pathways and systems in the brain.

Physical pain often begins in the skin or musculature, where specific receptors called *nociceptors* respond to intense pressure, heat, or other stimulation that can cause damage. Messages from these receptors pass through the brainstem and thalamus to reach somatosensory areas in the cortex, such as primary somatosensory cortex (S1), which you read about back in Chapter 2 (and saw in Figure 2.7). Brain imaging studies have shown that the more intense the pain, the more activity in S1. When you shower, for example, as the water gets hotter, the more activity you'll have in S1. But although S1 encodes the physical location and intensity of pain, it does not encode how bad it "feels"—the affective component of pain. If you have spent all day freezing outside in the snow, standing under that same very hot shower may actually feel good, rather than painful. Similarly, a man swallowing wasabi-flavored snacks may gasp for breath and wipe away tears—and then reach for another handful to do it again. Clearly, not all intense stimuli are aversive. And not all aversive stimuli cause physical pain: disgusting smells, loud dischordant sounds, and social rejection can all be highly aversive, even though no physical pain occurs.

insular cortex (insula). A region of cortex lying in the fold between parietal and temporal lobes that is involved in conscious awareness of bodily and emotional states and may play a role in signaling the aversive value of stimuli.

So, how does the brain decide whether a particular stimulation is aversive? Several brain areas have been implicated, including the **insular cortex**, or **insula**, shown in Figure 5.12. The insular cortex is located in the deep fold that separates the temporal lobe from the parietal and frontal lobes, and is important for our conscious awareness of our own bodies and emotional states. One subregion of the insular cortex, the *dorsal posterior insula*, plays a role in perception of physical pain, as well as other negative emotional states such as hunger, anger, and disgust (Naqvi & Bechara, 2009; Chang, 2013). For example, the dorsal posterior insula is active when participants experience painful heat or cold (for review, see Craig, 2003), and also when they experience social rejection, such as being excluded by the other players in an online video game (Eisenberger et al., 2003) or when viewing pictures of an ex-partner after an unwanted breakup (Kross et al., 2011). The degree of activation appears to be roughly proportional to the to the magnitude of the punisher. So, for example, in a study where errors could be punished by loss of 50 cents or of 5 cents, the insula showed more activity after a larger loss (Hester et al., 2010).

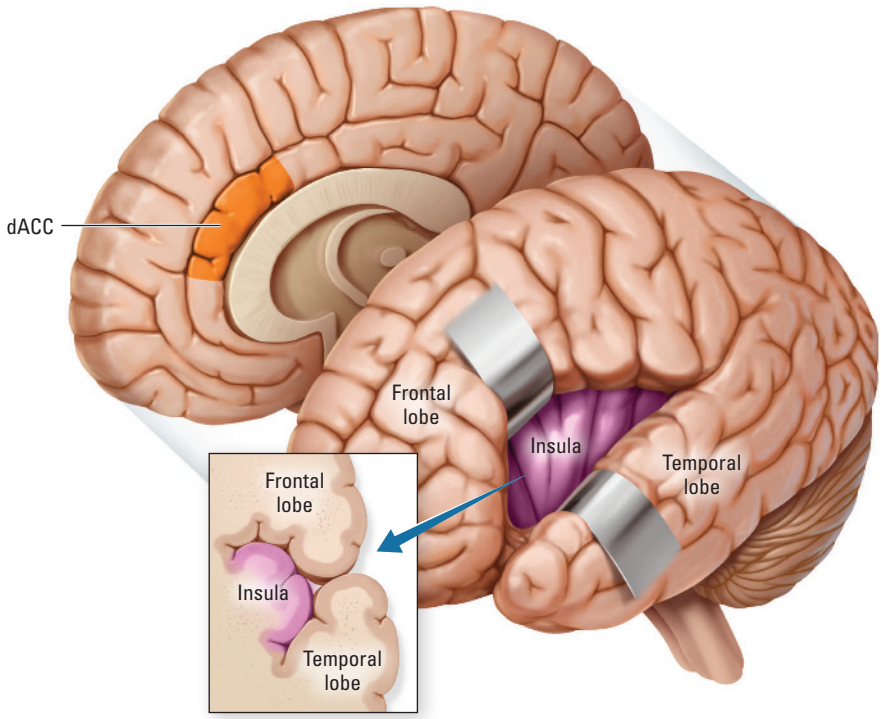

Figure 5.12 The insular cortex (insula) and dorsal anterior cingulate cortex (dACC) The insula, which lies buried in the fold separating the temporal lobe from the parietal and frontal lobes, is implicated in conscious awareness, and also plays a role in signaling the aversive value of stimuli. The dACC, which lies on the inner, or medial, surface of the prefrontal cortex, may play a role in the motivational value of punishers, helping select the actions we take in response.

Thus, just as the opioid system may signal pleasantness or "liking," the insula may be one way in which the brain determines degree of unpleasantness or "disliking." In fact, when the insula is damaged, learning to avoid unpleasant outcomes is impaired. One study found that patients with brain lesions that damaged the insula were as good as healthy controls at learning to obtain reward (point gain) but impaired at learning to avoid punishment (point loss), compared with patients whose brain damage spared the insula (Palminteri et al, 2012).

Once we've established that a stimulus is subjectively painful, the next step is to decide whether to do something about it. The **dorsal anterior cingulate cortex** (abbreviated **dACC**), which lies on the inner, or medial, surface of the prefrontal cortex, has been implicated in the motivational value of pain—the degree to which it can drive changes in behavior (Craig, 2003). Some current theories suggest that the dACC detects unexpected events (including pain) and suggests an appropriate response (Bush et al., 2002). For example, in one study, participants played a game in which they could win and lose money; during the game, neurons in the dACC responded both to errors that resulted in outright punishment and also to errors that merely resulted in no reward—but there was more activity in the former case (Simões-Franklin et al., 2010). Presumably, the worse the consequences of the error, the more motivation to change behavior.

dorsal anterior cingulate cortex (dACC). A subregion of prefrontal cortex that may play a role in the motivational value of pain.

On the other hand, dACC also shows increased activation when participants unexpectedly receive a reduced reward (Bush et al., 2002; Williams et al., 2004), and the activity level is predictive of whether participants actually change their response (Williams et al., 2004). Remember the phenomenon of negative contrast, in which monkeys and children refuse to work for a reward that is smaller than the one they've been trained to expect? In effect, the smaller-than-expected reward is functioning as a punisher, leading to decreased responding. It's possible that the dACC is recognizing this negative contrast, and signaling reduced motivation to work for the disappointing reward.

Thus, just as the brain has multiple systems for signaling the hedonic value and motivational value of reinforcers—"liking" via the opioid system and "wanting" via the dopamine system—the brain may also have multiple systems to signal the aversive value and motivational value of punishers, via brain areas such as the insula and dACC. However, much still remains to be understood about how we process and respond to punishers in the brain.

Interim Summary

- The dorsal striatum is an important brain substrate for storing stimulus–response ($S^D \rightarrow R$) associations; striatal-mediated $S^D \rightarrow R$ associations may be relatively automatic and habitual.
- The orbitofrontal cortex may be an important brain substrate for storing response–outcome ($R \rightarrow O$) associations, and in helping organisms to choose particular responses based on the expected outcomes of those actions.
- Reinforcers and punishers may activate neurons in the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc), which project dopamine to the dorsal striatum, frontal cortex, and elsewhere. Interrupting these pathways, by lesions or drugs, disrupts operant conditioning.
- The incentive salience hypothesis suggests that dopamine modulates "wanting" rather than "liking," determining how hard an organism is willing to work for a reinforcement. Dopamine also affects plasticity, possibly helping to create or strengthen $S^D \rightarrow R$ associations in the dorsal striatum and elsewhere.
- The endogenous opioids, which are mimicked by many highly addictive drugs, may signal the hedonic value ("liking") of reinforcers.
- The dorsal posterior insula is a brain region that helps us determine subjective "disliking" of painful physiological and psychological stimuli. The dorsal anterior cingulate cortex (dACC) may help determine the motivational value of punishers, which is used to guide changes in behavioral responding.

5.3 Clinical Perspectives

Through the brain's reinforcement system, animals are hardwired to seek and obtain the things they need for survival (food, water, sleep, etc.) and to avoid those things that threaten survival (pain, sickness, predators, etc.). Unfortunately, this powerful reinforcement system can go awry. As an example, consider the pleasure we feel when we eat fatty food, which ensures that we are sufficiently motivated to repeat the experience. The human brain evolved millennia ago, when our ancestors had to forage for food and could never be sure when they'd find their next meal. Fat could be stored in the body and used for energy later, when food was scarce. Under these conditions, seeking out fatty foods was a good strategy for survival. In 21st-century America, however, food is easier for most of us to obtain, but our biological drives have not changed, and many of us—still driven to obtain the taste of fatty foods—have become dangerously overweight.

Drug addiction represents another way in which the reinforcement system can malfunction (or, rather, function only too well). You read in Chapter 4 how classical conditioning can contribute to drug addiction. Another large piece of the addiction puzzle is operant conditioning: learned responding to obtain a particular kind of reinforcement. Insights from operant conditioning theory may deepen our understanding of addiction and lead to more effective treatments.