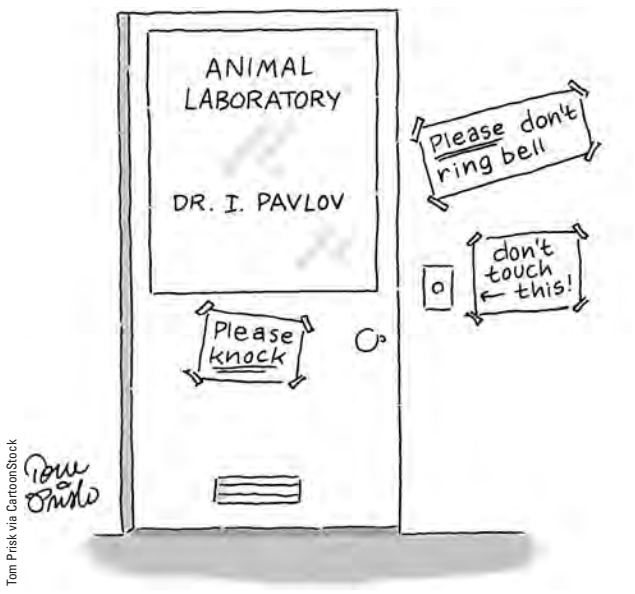
Classical Conditioning

Learning to Predict Significant Events

WHAT DO THE FOLLOWING FOUR people have in common? Four-year-old Moira, who screams with delighted anticipation for ice cream when she hears the jingle of the vendor's truck in the distance; Dan, who starts feeling anxious about the carpeting in his home when he sees rain clouds on the horizon and realizes that he left the windows open this morning; Nathalie, a former cigarette smoker, who always feels the urge to light up after sex; Sharon, who broke up with her ex-boyfriend years ago but still finds the sound of his voice arousing. It's not immediately apparent, but a little investigation will show that the link connecting Moira, Dan, Nathalie, and Sharon's reactions is Ivan Pavlov—or to be more precise, Ivan Pavlov's principle of classical conditioning. They have all had their behaviors altered by classical (or "Pavlovian") conditioning.

Most people, even if they never took a psychology course, are vaguely aware of the story of Ivan Pavlov (1849–1936) and how he trained, or "conditioned," his dogs to salivate to cues such as bells or tones that predicted the impending delivery of food; references to "Pavlov's dogs" can be found throughout popular culture, including in movies, TV shows, and many cartoons, as shown here and later in this chapter. Chapter 1 briefly introduced you to Pavlov and his training method; this chapter will explain why his work was so important and influential and why it continues to be relevant to experimental and clinical studies today.

Like many advances in science, Pavlov's discovery of classical conditioning was largely accidental. He was originally studying digestion, and he noticed that his dogs often started salivating even before they received their daily meat rations—when they saw the bowl that usually contained their food or when they heard the footsteps of the laboratory assistant who fed them (Pavlov, 1927). Initially, Pavlov viewed the premature salivation as a nuisance that interfered with his efforts to understand how the digestive system responds to food. Soon, however, Pavlov realized that he had stumbled on a way of studying how associations are formed in the brain of a dog.



Tom Prisk via CartoonStock

Why might a visitor ringing the doorbell interfere with Pavlov's experiments?

Pavlov and his assistants began a systematic study of factors that influence how an animal learns. By keeping the dogs in restraints and collecting saliva through a tube surgically inserted into their mouths (Figure 4.1), Pavlov could measure salivation in response to various cues. He began one study by first training a dog that a doorbell always preceded delivery of food; over many paired doorbell–food trials, the dog developed a stronger and stronger salivation response to the sound of the doorbell. This form of learning, in which an animal learns that one stimulus (such as a doorbell) predicts an upcoming important event (such as delivery of food), is known today as **classical conditioning** or **Pavlovian conditioning**.

There is much more to classical conditioning than dogs and saliva, however. This chapter will show why an understanding of classical "Pavlovian" conditioning (despite its seeming simplicity) is indispensable for building a behavioral and biological understanding of learning and memory. Moreover, classical conditioning is one of the few forms of learning for which the brain substrates have been worked out in precise detail, for every step, from the initial sensory input to the commands that drive the resulting motor responses. For these reasons, classical conditioning is avidly studied today by psychologists, neuroscientists, and clinical neuropsychologists, with implications far beyond what Pavlov could have anticipated when he made his accidental discovery almost one hundred years ago.

Figure 4.1 Pavlov's apparatus for studying learning
A restrained dog has a surgical tube inserted into its mouth to collect and measure salivation in response to meat placed in front of it or to a cue, such as a doorbell, that predicts delivery of the food.

4.1 Behavioral Processes

Classical Pavlovian conditioning is a way of learning about one's environment. As a child, Moira learned that the distant sound of a certain jingle predicts the imminent arrival of an ice cream truck. Since she loves ice cream, she can exploit her foreknowledge of the truck's arrival by asking her mother for money now, so she can be ready at the curb when the truck approaches. This is an example of learning to anticipate a positive event so as to take maximal advantage of it. Being able to anticipate negative events is also useful. If Dan is surprised by a sudden rainstorm in midday, he must rush home from work to close all the windows to keep rain from blowing in and soaking his carpets. Had he anticipated the rainstorm earlier, Dan could have closed the windows before he left for work.

This section begins by introducing the basic concepts and terminology of classical conditioning and then explores subsequent research into this type of learning. It describes an elegant and simple theory of conditioning, developed in the early 1970s, that helps explain a wide range of learning phenomena, and it discusses how conditioning behaviors seen in the simplest of animals help us understand the more complex cognitive behaviors observed in human learning. It ends by discussing several other facets of classical conditioning, including how attention can affect what is learned, how timing is important, and to what degree we may be biologically prepared to learn some things more easily than others.

Basic Concepts of Classical Conditioning

A dog will naturally salivate when it sees or smells food. No learning is needed for it to make this response. For this reason, psychologists call the food an **unconditioned stimulus**, or **US**, meaning a stimulus that naturally—that is, without conditioning—evokes some response. An unconditioned stimulus, such as food, evokes a natural response, such as salivation, which psychologists call the **unconditioned response**, or **UR**; their relationship does not depend on learning. Similarly, Moira's craving for ice cream and Dan's dismay at wet carpets in his house are natural—that is, unconditioned—responses to good and bad things in their lives. They both occur unconditionally without prior training. In contrast, a neutral stimulus, such as a bell that the dog has not heard before, evokes no such salivation by the dog (Figure 4.2a).

unconditioned stimulus (US). A cue that has some biological significance and in the absence of prior training naturally evokes a response.

unconditioned response (UR). The naturally occurring response to an unconditioned stimulus (US).

How Pavlov Conditioned Dogs to Salivate

After Pavlov put his dogs into the apparatus shown in Figure 4.1, he repeatedly paired the bell with food: each time the bell was rung, an assistant promptly delivered food to the dog. This resulted in the formerly neutral stimulus, the bell, becoming a **conditioned stimulus**, or **CS**, as illustrated in Figure 4.2b. After repeated presentations of the bell CS and the food US, the two became linked in the dog's mind. This training—or conditioning, as Pavlov called it—resulted in the dog learning something new: the bell predicts the food. We can evaluate the degree to which the dog learned this prediction—that is, how strongly it expects the food when it hears the bell—by measuring how much the dog salivates to the bell alone. This is shown in Figure 4.2c, in which the bell alone, the conditioned stimulus (CS), now evokes an anticipatory response, called the **conditioned response**, or **CR**, even in the absence of the food.

conditioned stimulus (CS). A cue that is paired with an unconditioned stimulus (US) and comes to elicit a conditioned response (CR).

For Moira, the truck jingle is a conditioned stimulus (CS), which, after being repeatedly paired with the arrival of the ice cream truck, evokes an anticipatory conditioned response (CR) consisting of Moira asking her mother for money and running to the street corner. When this response has been fully learned, Moira will be waiting at the street corner, money in hand, when the unconditioned stimulus (US)—the ice cream truck—arrives at her house. For Dan, the

conditioned response (CR). The trained response to a conditioned stimulus (CS) in anticipation of the unconditioned stimulus (US) that it predicts.

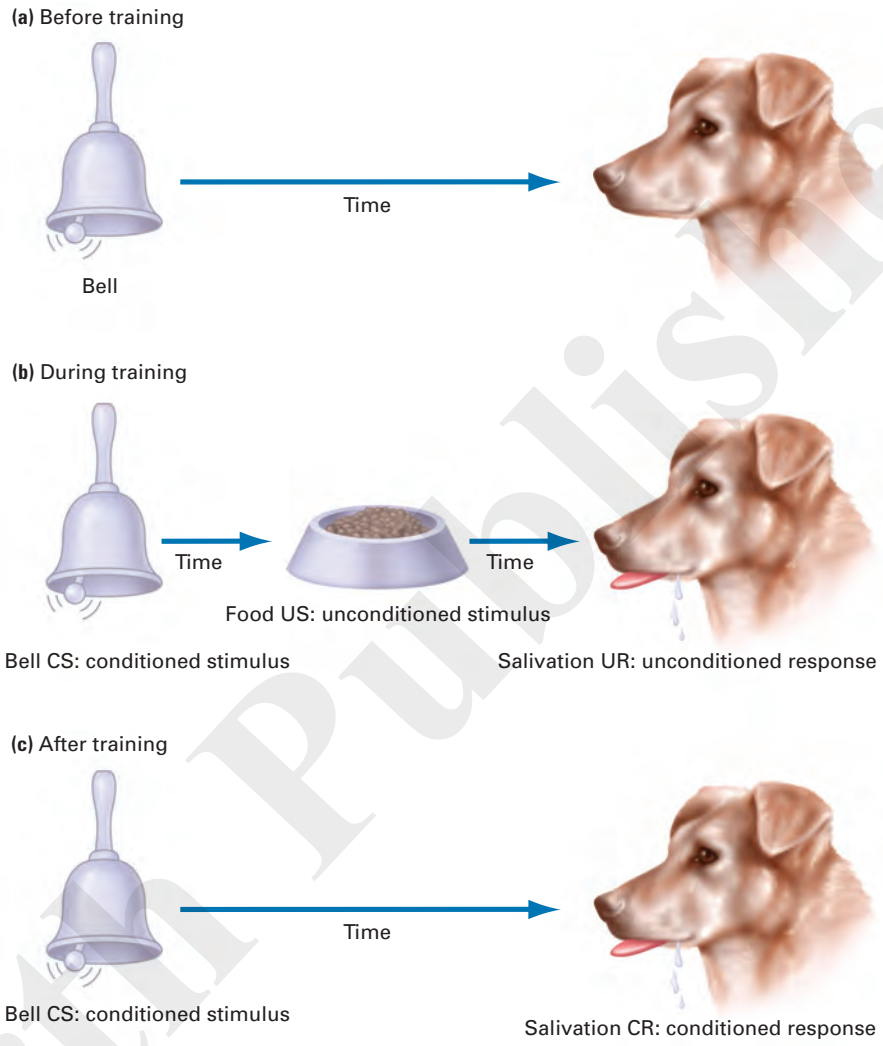


Figure 4.2 Schematic illustration of Pavlov's experiment (a) Before training, a neutral stimulus, such as a bell, evokes no response from a dog. In contrast, the presence of food, an unconditioned stimulus (US) would naturally evoke an unconditioned response (UR), salivation. (b) During training, the bell, formerly a neutral stimulus, becomes a conditioned stimulus (CS) when it is repeatedly paired with the food US to evoke a salivation UR. (c) After training, the bell, now a conditioned stimulus (CS), evokes a learned response, the conditioned response (CR) of salivation.

dark clouds in the sky are an ominous conditioned stimulus (CS), which have become associated with subsequent heavy rains through his past experiences. These experiences have taught him a *preparatory conditioned response* (CR): to close all the windows in his house before the arrival of the rain, the unconditioned stimulus (US). In all three examples—Pavlov's dog, Moira, and Dan—a learned association between a CS and subsequent US generates a CR that follows the CS. Table 4.1 reviews these terms and their relation to the three

Table 4.1 Terminology of Pavlovian conditioning, with examples

	Unconditioned stimulus, US	Unconditioned response, UR	Conditioned stimulus, CS	Conditioned response, CR
Pavlov's dog	Food	Salivation	Bell	Salivation
Moira	Ice cream truck	Appetite for ice cream	Truck jingle	Get money from Mom, run to corner before truck arrives
Dan	Rain	Closing windows while carpets get wet	Dark clouds	Closing windows before the rain starts

examples. We will continue to use all four terms—US, UR, CS, CR—throughout this chapter as well as the rest of this book, so it's a good idea to be sure you are comfortable with all four before moving on.

Test Your Knowledge

Pavlov's Experiment

Table 4.1 describes Pavlov's experiment from the perspective of the experimenter, Pavlov. From the dog's perspective, however, the scenario seems quite different. Using the cartoon here, identify Pavlov's CS and CR as viewed by the dog. (Answers appear in the back of the book.)

Appetitive Conditioning

When the US is a positive event (such as food delivery for Pavlov's dog or ice cream for Moira), the conditioning is called **appetitive conditioning**. In general, appetitive conditioning consists of learning to predict something that satisfies a desire or appetite. Food and sex are among the most powerful of appetitive USs. Recall Sharon, one of the four people described at the chapter's beginning; she was conditioned to the sound of her ex-boyfriend's voice by its past association to her having sex with him. Michael Domjan and colleagues have studied a similar form of conditioning using male domesticated Japanese quail, who will copulate readily with a sexually receptive female (Figure 4.3). When an arbitrary stimulus, such as a light CS, is paired repeatedly with access to a sexually receptive female (the US), the male quail exhibits a CR of approaching and remaining near the light (Domjan, Lyons, North, & Bruell, 1986).

appetitive conditioning. Conditioning in which the US is a positive event (such as food delivery).

Aversive Conditioning

Dan's learning that clouds predict rain damage to his home (if the windows are not shut) is an example of **aversive conditioning**, learning to avoid or minimize the consequence of an expected aversive event. Many of the procedures used for the experimental study of conditioning are examples of aversive conditioning.

In Chapter 1 you read about both B. F. Skinner, the father of behaviorist approaches to the science of learning, and W. K. Estes, a founder of mathematical psychology and learning theory. In the early 1940s, Estes was a graduate

aversive conditioning. Conditioning in which the US is a negative event (such as a shock or an airpuff to the eye).

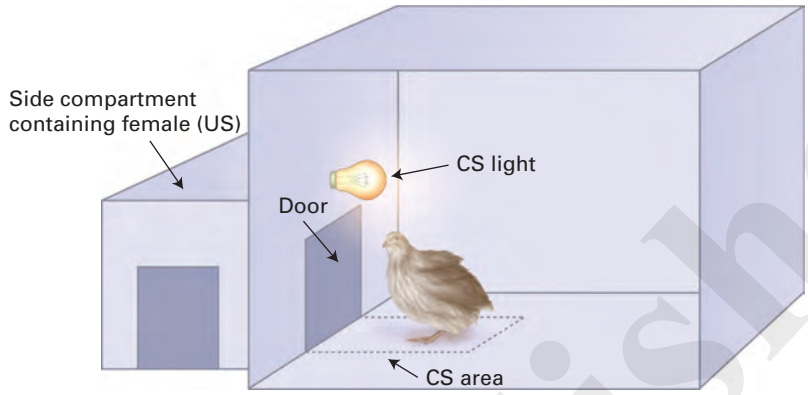


Figure 4.3 Sexual conditioning in male Japanese quail Michael Domjan and colleagues conditioned male domesticated Japanese quail to approach and remain near a light (the CS) that is associated with access through a door to a sexually receptive female (the US).

student of Skinner's. At that time, the United States had not yet entered World War II. The Germans were using a new technology—rockets—to bomb England. As Londoners heard the whine of the rocket engines approaching, they stopped whatever they were doing—eating, walking, talking—and waited for the explosions. After the rockets dropped elsewhere and people realized they were safe, they resumed their interrupted activities.

Intrigued by these stories from London, Estes and Skinner developed a new conditioning procedure for rats that was similar, in some respects, to what they imagined Londoners were experiencing. This procedure, which they called the "Conditioned Emotional Response" (CER), was a technique for studying learned fear (Estes & Skinner, 1941). Estes and Skinner placed hungry rats in a cage that delivered food pellets whenever the rats pressed a lever. The cage also had a metal grid floor wired to deliver a mild shock to the rats' feet. Normally, the hungry rats busily pressed the lever to obtain food, but if the experimenters trained the rats to learn that a tone (the conditioned stimulus, or CS) predicted an upcoming shock (the unconditioned stimulus, or US), the rats would freeze (their conditioned response, CR) when they heard the tone, interrupting their lever presses and waiting for the shock. Measuring this freezing behavior allowed Estes to quantify trial-by-trial changes in the learned response. Within a few years, this conditioned emotional response procedure became one of the most widely used techniques for studying animal conditioning, and it is still in use today. (In Chapter 10 you can read more about fear conditioning and how emotions, such as fear, influence learning and memory.)

In 1928, Cole Porter wrote "Birds do it, bees do it, even educated fleas do it," for the Broadway show *Paris*. Porter was, of course, referring to falling in love, but he could just as well have been writing about learning by classical conditioning. Even insects such as fleas and flies can be trained using classical conditioning methods. In fact, studies of classical conditioning of the fruit fly *Drosophila* have been enormously important for understanding the biology of learning (we'll see an example of such a study later in this chapter).

Figure 4.4 illustrates the behavioral procedure used in studies of fly conditioning (Dudai, Jan, Byers, Quinn, & Benzer, 1976). First the flies are placed in a container that contains one odor, designated odor 1 (Figure 4.4a), and nothing happens. Then the flies are exposed to another odor, odor 2, and in the presence of that odor (the CS), they are given a mild but aversive shock (the US). Later, the flies are placed in the middle of a container that has odor 1 at one end and odor 2 at the other end (Figure 4.4b). As the flies explore the container, they avoid the side where they smell odor 2 (which has been associated with shock) and gravitate toward the side where they smell odor 1 (which was not paired

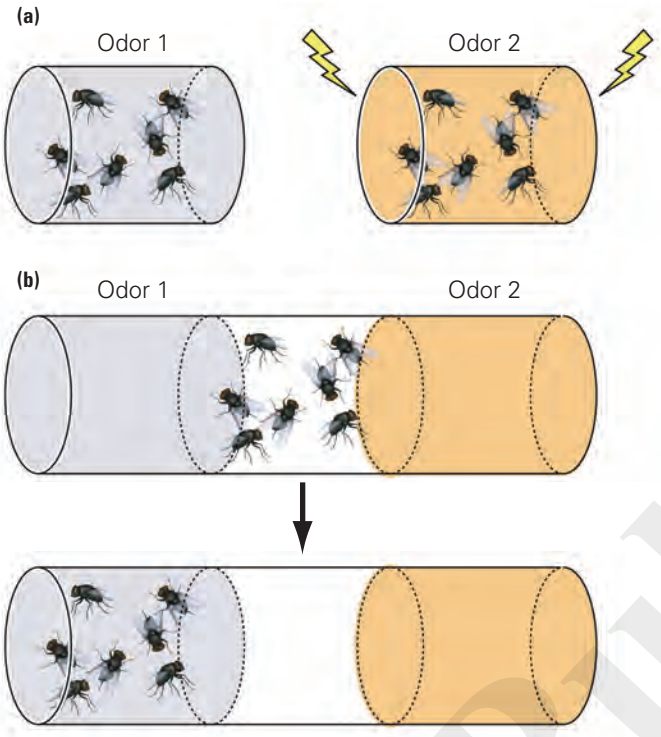


Figure 4.4 Odor conditioning in flies (a) Flies are placed sequentially in two different containers, first in one with odor 1, where they are not shocked, and then in another with odor 2, where they are shocked. (b) Later, they are placed in the middle of a container that has odor 1 at one end and odor 2 at the other end. The flies move toward odor 1, which was not associated with shock, indicating that they have learned the odor 2–shock association from their previous training.

with shock). Like Dan's keeping out the rain, the rats' and flies' avoidance of shocks are examples of aversive conditioning to predictors of a significant negative event.

Understanding the Conditioned Response

So far, we have introduced four different procedures for experimenting with classical conditioning. Two of these are appetitive conditioning procedures: Pavlov's original study with dogs and food, and Domjan's studies of quails and sex. The other two are aversive conditioning procedures: the conditioned emotional response in which rats freeze when they hear a tone that predicts a shock (Estes & Skinner, 1941), and the fly shock preparation shown in Figure 4.4.

In each of these four cases, we can ask, why does the animal exhibit the conditioned response? In all four, the conditioned response can be understood as an anticipatory response that prepares the animal for the expected US, in much the same way that Moira prepares for the arrival of an anticipated ice cream truck, or Dan prepares for a predicted rainstorm. By moving away from the odor associated with shock, the fly is more likely to avoid being shocked. By salivating in anticipation of food, the dog is better prepared to efficiently digest the food. By freezing in anticipation of a shock, the rat is better prepared to ward off danger and also avoids having ongoing motor behaviors (such as eating) disrupted by the shock. By moving toward the light, the quail is all the sooner able to mount and copulate with the female.

Mammalian Conditioning of Motor Reflexes: Eyeblink Conditioning

Another widely studied form of aversive conditioning—one that induces an anticipatory defensive response much like Dan's shutting his windows—is **eyeblink conditioning**, perhaps the most thoroughly studied form of motor reflex

eyeblink conditioning. A classical conditioning procedure in which the US is an airpuff to the eye and the conditioned and unconditioned responses are eyeblinks.

Courtesy of Ernest Hilgard

Figure 4.5 Face-slap eyeblink conditioning in the 1920s Clark Hull (standing with visor) and his young graduate student Ernest Hilgard (seated) in an early study of classical eyeblink conditioning at Yale University. Hull trained Hilgard to blink in anticipation of a slap to the face.

conditioning in mammals (Gormezano, Kehoe, & Marshall, 1983). You may recall Clark Hull of Yale from Chapter 1, one of the fathers of mathematical learning. He used (and perhaps abused) his graduate students by teaching them to blink in anticipation of a slap to the face, as shown in Figure 4.5. The subject in this photo is Ernest "Jack" Hilgard, who later went on to become a pioneer in the study of the psychology of hypnosis and the author of a leading textbook on memory.

Hull arranged for a tone (CS) to play just prior to each face slap (US). After many repeated presentations of the CS followed by the US, poor Jack began to blink (the CR) every time he heard a tone, whether or not it was followed by a slap.

For practical as well as ethical reasons, researchers no longer use the face slap as a US in human eyeblink conditioning. Instead, they often use an airpuff to the eye. This is not painful, but it does cause a reflexive—that is, unconditioned—eyeblink UR (if you don't believe this, ask a friend to blow lightly in your eye). The blink UR in human eyeblink conditioning can be measured in several different ways, one of which is through the use of electromyography (EMG) detectors of electrical activity of muscles, placed above and below the eye as shown in Figure 4.6a. The CS is a tone, typically delivered through headphones so that the volume can be regulated (and outside noises masked). With repeated pairings of the tone CS and airpuff US, subjects develop a CR: in this case, an anticipatory blink that occurs before US arrival, so that the eye is partially shut and partially protected when the airpuff occurs.

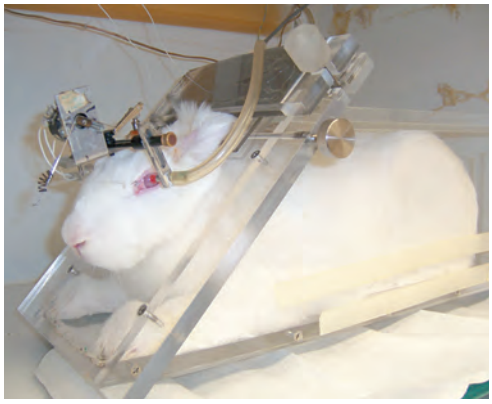
What is most important about eyeblink conditioning is the similarity of its appearance in many different species, so that the results found in one species can reasonably be expected to apply to others. Eyeblink conditioning has been shown in mice, rats, and monkeys, but one of the most common animals for the study of eyeblink conditioning has been the rabbit, because of its propensity to sit still for long periods and to blink very little except when something bothers its

Figure 4.6 Eyeblink conditioning in humans and rabbits (a) In human eyeblink conditioning, a tone CS is delivered through headphones. The US is a puff of air delivered through the rubber tube. The eyeblink CR is recorded by EMG electrodes placed above and below the eye. (b) In rabbit eyeblink conditioning, a similar rubber tube delivers the airpuff US to the rabbit in the restraining acrylic glass case; a photobeam measures the CR and UR.

(a)

Mark Gluck

(b)



Richard F. Thompson

eyes. Figure 4.6b shows a rabbit in a restraining acrylic glass case within which eyeblink conditioning is often conducted. The tube at the rabbit's left eye delivers the airpuff US while the eyeblink CR and UR are measured by photo beams.

With rabbits, just as with humans, the airpuff is the US, and the reflexive blink is the UR, as shown in Figure 4.7. Before training, the tone does not cause the rabbit to blink because it is a neutral stimulus, as shown in Figure 4.7a. However, if the airpuff US is repeatedly preceded by a tone, then the animal learns that the tone predicts the airpuff US and is a warning signal to get ready, as shown in Figure 4.7b. Eventually, the animal will blink as a response to the tone alone, as shown in Figure 4.7c. At this point, the tone has become a CS, and the anticipatory eyeblink is the CR.

To the uninformed observer, the learned conditioned response, the eyeblink CR, is identical to the automatic unconditioned response, the eyeblink UR. However, the learned CR takes place during the warning period provided by the CS (analogous to a weather report predicting rain) in advance of the US and

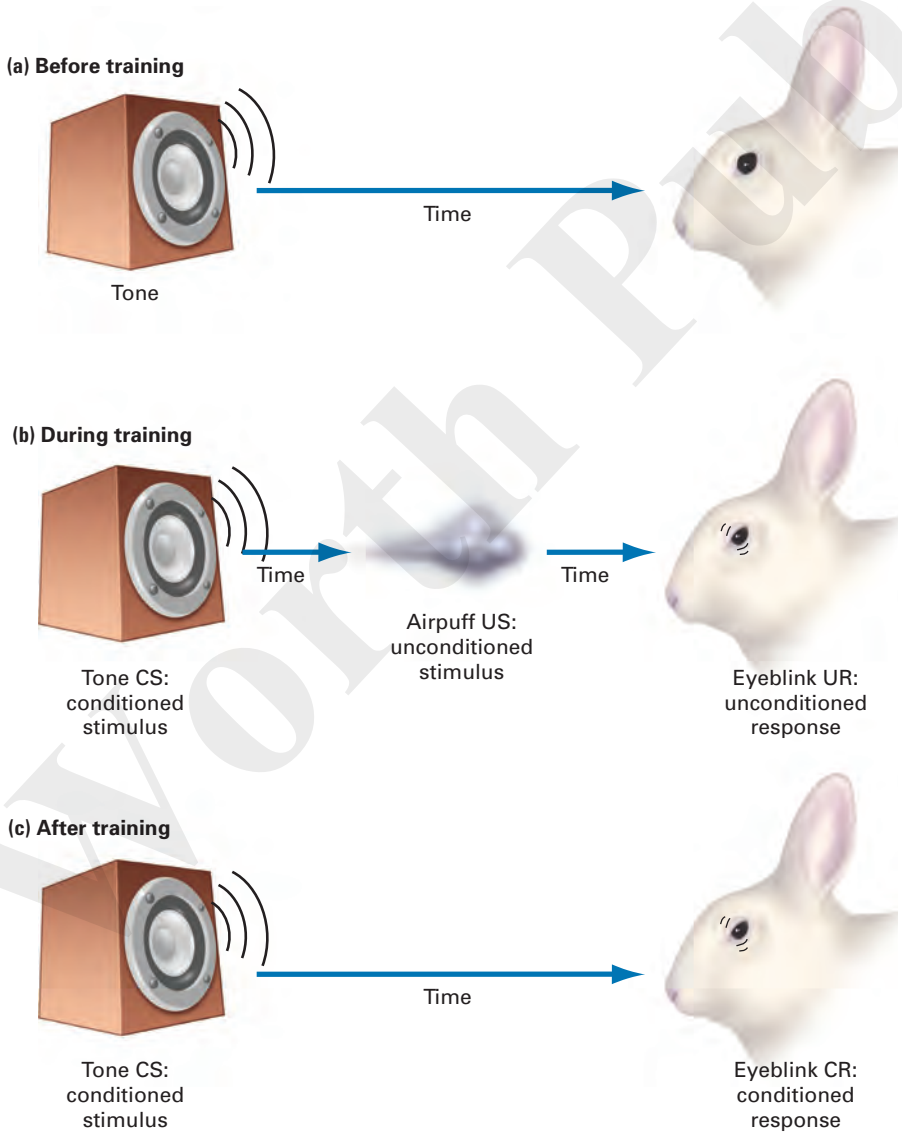


Figure 4.7 Learning progression in rabbit eyeblink conditioning (a) Before training, the tone is a neutral stimulus that has no relevance to the rabbit. (b) The tone CS followed by airpuff US causes an eyeblink UR in a naive rabbit. (c) The tone CS is followed by a blink CR in a rabbit that has undergone eyeblink conditioning. Compare these diagrams to Figure 4.2, showing analogous stages of Pavlov's conditioning procedure.

Table 4.2 Widely used classical conditioning procedures

	Unconditioned stimulus, US	Unconditioned response, UR	Conditioned stimulus, CS	Conditioned response, CR
Appetitive conditioning				
Pavlov's dog	Food	Salivation	Bell	Salivation
Quail sex	Sexually available female	Approach, mounting, and copulation	Light	Approach
Aversive conditioning				
Fly shock	Shock	Attempt to escape	Odor	Attempt to escape
Conditioned emotional response	Shock	Freezing	Tone	Freezing
Eyeblink conditioning	Airpuff	Blink	Tone	Blink

UR, adaptively protecting the eye from the onset of the airpuff. The same is true for Pavlov's original salivation study, in which the learned CR, salivation, is the same as the dog's natural unconditioned response to food, but it takes place *before* the food is presented, at the sound of the doorbell that predicts the food.

You have now been introduced to five different formats that are used for experiments in classical conditioning: two appetitive preparations and three aversive preparations. These are all reviewed in Table 4.2. To see if you really understand this material and can tell your USs, URs, CSs, and CRs apart, try to identify each in the real-world examples given in the accompanying Test Your Knowledge box.

Test Your Knowledge

Classical Conditioning in Everyday Life

Are you sure you can tell the US, UR, CS, and CR apart? Test yourself by identifying each of them in the real-world situations described below. For each, indicate if it is an example of appetitive or aversive conditioning. (Answers appear in the back of the book.)

1. Advertisements for a new sports car show a sexy female model draped over the car's hood.
2. Mark loves pizza. When he was a boy, his parents frequently had it delivered to their home. Because the pizzas often arrived only lukewarm, his parents would put the pizza, still inside the box, into the oven to heat up. This caused the box to give off a smell of burning cardboard. Now, years later, whenever Mark smells cardboard burning, he gets hungry for pizza.

Learning a New Association

How, exactly, does learning progress in eyeblink conditioning or other classical conditioning procedures? Figure 4.8 shows an eyeblink CR becoming stronger over several days of training in a rabbit eyeblink-conditioning study.

Each day, the animal received 80 training trials, each of which presented a tone followed shortly thereafter by an airpuff to the eye. The graphs (green lines) in Figure 4.8 show the extent to which the rabbit's eyelid lowers at the start of different days during the experiment; the higher the curve, the farther the eyelid has shut. Note that on the beginning of day 1, the only response is the eyeblink UR that occurs *after* the onset of the airpuff US. However, with training, an eyeblink CR emerges: By day 3, there is movement of the eyelid before the US arrives. This anticipatory blink in response to the CS is the beginning of a CR. With further training, by about day 5, a strong anticipatory eyeblink CR occurs, timed so that the eyelid is safely closed before the airpuff US occurs.

In both rabbits and humans, eyeblink conditioning is a gradual process, occurring over many trials. Figure 4.9 shows trial-by-trial changes in the percentage of human participants and rabbits giving conditioned eyeblink responses in a study of tone–airpuff conditioning in both species. Although the graphs in the figure are not identical (humans learn faster), they are quite similar. Most important to note is that in both humans and rabbits, the percentage rises over time until most trials elicit an appropriately timed predictive eyeblink CR.

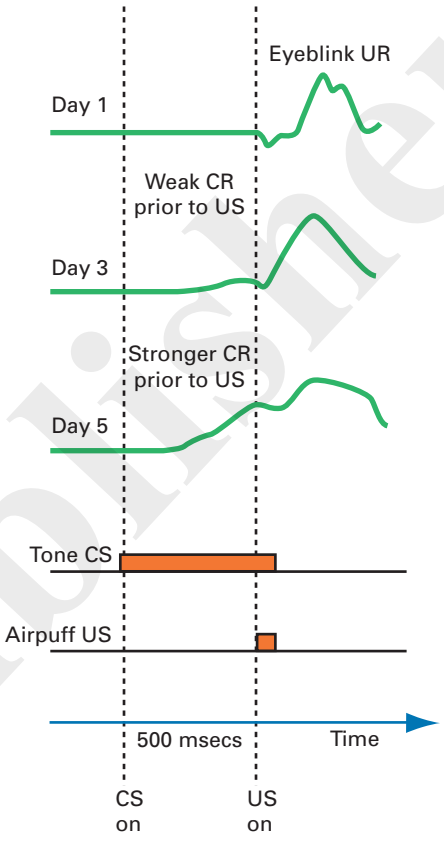


Figure 4.8 Acquisition of eyeblink-conditioning response Development of a conditioned response as measured at the beginning of day 1, day 3, and day 5 of training, using a standard tone–airpuff trial sequence. On day 1, only a UR to the eyepuff is observed, but by day 3, an anticipatory eyeblink starts to emerge. By day 5, this anticipatory CR is strong and occurs reliably before the airpuff US.

Refining the Basic Principles

Studies of classical conditioning have led to a broad range of insights into the subtle factors that influence animal and human learning. In this section we explore four of these topics, looking more closely at variations in the USs that develop, the constraints on what makes a cue a CS or a US, how old associations are extinguished when no longer applicable, and what happens when two cues, rather than one, are presented during learning.

Conditioned Compensatory Responses

Remember Dan who learned (from past bad experiences) to notice signs of impending rain and close his windows in advance of it to protect his carpets? Dan also has a swimming pool out back. If he were expecting heavy rains for several days, he might worry about the pool overflowing and damaging his lawn and house. Given the weather forecast, he might think it prudent to partially drain his pool, lowering the level a few inches before the rain arrives. When the rain does come, it will return the water level to the point it was at before he drained it, lessening the likelihood of an overflow. In this way, Dan's preparatory response (preemptively lowering the water level in the pool) compensates for the expected rise in water level and ensures that the pool never gets too full.

An analogous conditioned *compensatory response* was demonstrated by two of Pavlov's colleagues more than 60 years ago (Subkov & Zilov, 1937). These researchers injected dogs on several occasions with adrenaline (also known as epinephrine), a chemical normally produced by the adrenal glands in response to stress or anxiety. The usual effect of adrenaline is an increase in heart rate. However, the dogs' heart rate increased less and less with each subsequent injection. Such a decrease in reaction to a drug, so that larger doses are required to achieve the original effect, is known as **tolerance.** What causes tolerance to develop?

tolerance. A decrease in reaction to a drug so that larger doses are required to achieve the same effect.

To explore this question, the researchers placed their dogs on injection stands, where the dogs normally received the drug injection, but they

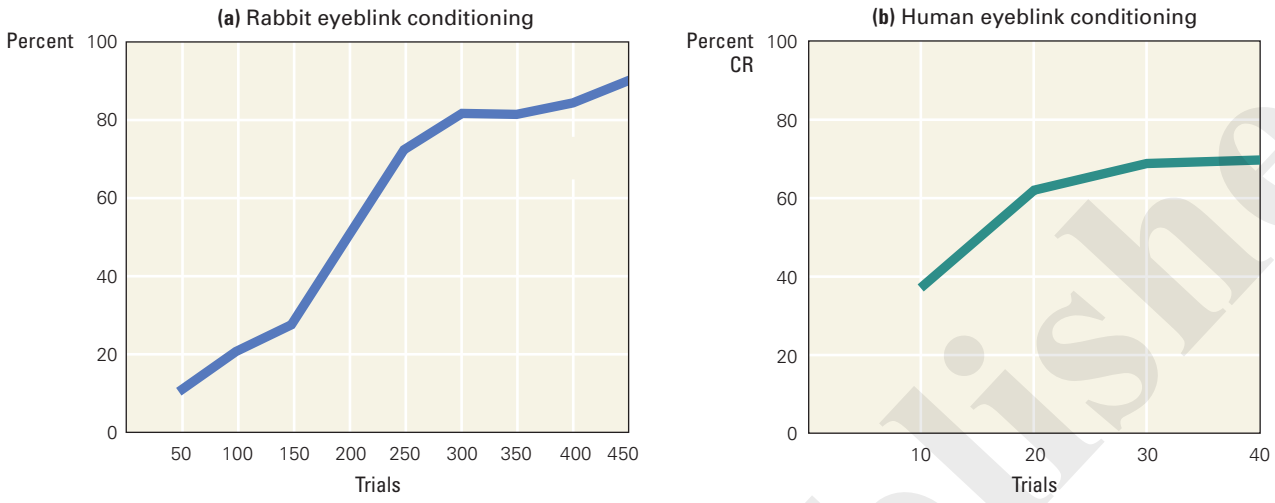


Figure 4.9 Learning curves for rabbit and human eyeblink conditioning (a) A learning curve showing the percent CRs in rabbits across blocks of training trials. (b) Analogous learning curve for human eyeblink conditioning. Although these curves are qualitatively similar, they reflect different training regimes, since the rabbits are usually trained in blocks of 1-hour trial sessions on successive days, while humans are trained in a single hour-long session.

(a) Data from Allen, Chelius, & Gluck, 2002; (b) Data from Allen, Padilla, Myers, & Gluck, 2002.

administered a neutral inert substance rather than the adrenaline. The researchers observed that this caused the dogs' heart rate to *decrease*. Apparently, the various cues (the stand, the injection) that predicted the adrenaline injection triggered a conditioned compensatory response that lowered the dogs' heart rate in anticipation of the adrenaline causing an increase in heart rate. Such automatic compensatory responses occur primarily in body systems that have a mechanism for **homeostasis**, the tendency of the body (including the brain) to gravitate toward a state of equilibrium or balance.

Much like the homeowner who acts to prevent the pool from overflowing during a storm, the dogs in these studies unconsciously used advance information about the forthcoming adrenaline injection to compensate for the drug's effect. The learned anticipatory decrease in heart rate combined with the increase produced by the drug resulted in a lower total increase in heart rate than was experienced on the first (unexpected) administration of adrenaline. Since the dogs had been conditioned to expect adrenaline after seeing cues such as the stand or the syringe, their bodies compensated by lowering their heart rates to maintain a constant level. The same compensatory mechanisms at work in certain aspects of drug addiction will be discussed in Section 4.3.

homeostasis. The tendency of the body (including the brain) to gravitate toward a state of equilibrium or balance.

What Cues Can Be CSs or USs?

The USs in a conditioning experiment are by definition events that are biologically significant, either because they are inherently positive (such as food or sex) or because they are inherently negative (such as shock or an airpuff to the eye). In contrast, a CS can be any cue in the environment, even a US. Thus, an airpuff to the eye, which is a US in the eyeblink-conditioning paradigm, can serve as the CS in another experiment, where for example an animal might learn that an airpuff predicts food delivery (the new US). Thus, stimulus cues are not inherently CSs or USs; rather, those terms define the roles the cues play in a particular learning situation.

Remember the description of Nathalie at the beginning of this chapter. She is a former smoker who gets an urge for a cigarette after sex. In Nathalie's case, sex is the CS that has become associated with cigarette smoking, the US, as shown in Table 4.3. After a person gets into the regular habit of having a cigarette after sex, the craving for and expectation of cigarettes becomes the CR. (You'll read more about addiction and conditioning later on in this chapter.) In contrast, for Sharon, who becomes aroused at the sound of her ex-boyfriend's voice, his voice

Table 4.3 Sex as a CS and a US

	Unconditioned stimulus, US	Unconditioned response, UR	Conditioned stimulus, CS	Conditioned response, CR
Nathalie	Smoking a cigarette	Reduced craving for cigarette	Sex	Desire for a cigarette
Sharon	Sex	Sexual gratification	Ex-boyfriend's voice	Sexual arousal

is now the CS and her sexual arousal is her CR. Thus, for Nathalie sex can be a CS that predicts cigarette smoking, while for Sharon, it is the US that previously followed hearing her boyfriend's voice. It all depends on the individual's unique experiences.

Extinguishing an Old Association

What do you think would happen if Moira moved to a new neighborhood where jingling trucks sold not ice cream, which she loves, but broccoli, which she hates? If each time she heard the jingle, she got broccoli (rather than ice cream), you might expect that her excited reaction to the jingle would eventually disappear. In this case, a previously acquired association would become diminished through repeated presentation of the CS (jingle) in the absence of the US (ice cream), a process known as **extinction**, which was first described in the early studies of Pavlov (1927), as noted in Chapter 1.

extinction. The process of reducing a learned response to a stimulus by ceasing to pair that stimulus with a reward or punishment.

Once it is acquired, eyeblink conditioning can also undergo extinction if the former CS (tone) is presented repeatedly without an airpuff. Eventually, the rabbit (or person) that was formerly conditioned to blink to the tone begins to learn that the world has changed and the tone no longer predicts the US. Figure 4.10 shows what happens if, after 70 trials of eyeblink-conditioning acquisition training, rabbits are given 20 trials of tone-alone extinction training (Moore & Gormezano, 1961).

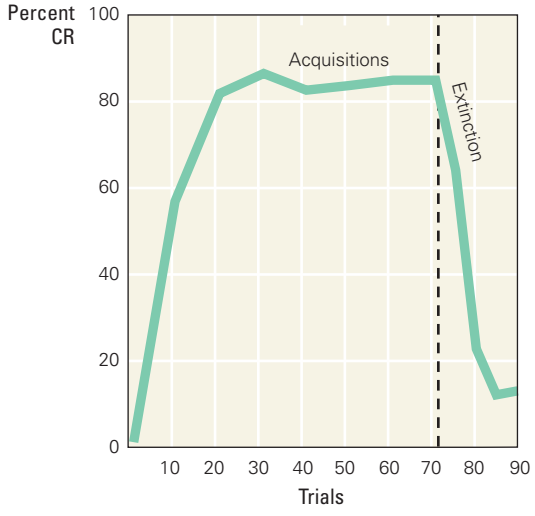


Figure 4.10 Acquisition and extinction of eyeblink conditioning Percent of rabbits exhibiting conditioned eyeblinks during 70 trials of acquisition and 20 trials of extinction.
Data from Moore & Gormezano, 1961.

It is tempting to think of the extinction in Figure 4.10 as simply the unraveling of acquisition. However, in recent years a consensus has been building in support of the idea that extinction is not just unlearning but rather a combination of unlearning and the learning of a new, opposing response to the CS. Specifically, it appears that during extinction, the CS acquires a second "don't respond" meaning that competes with the originally acquired "do respond" association. This suggests that even though the animal (or person) is no longer responding to the CS at the end of extinction training (as seen in Figure 4.10), the learned response is not gone, just unexpressed.

Some of the most compelling evidence for this view of extinction comes from studies that show the original learned response can reappear if the extinction was accompanied by a change in context (such as another room or testing chamber). In one example from the laboratory of Mark Bouton, a CS is paired with a shock in one context (context X) and then extinguished in another context (context Y); then the CS is again presented in context X, and the response is renewed (Bouton & King, 1983). As noted above, the most parsimonious explanation seems to be that there are two associations: CS–US and CS–no US; the context determines which response is retrieved (Bouton, 1991). During initial training, a CS–US association is created in context X

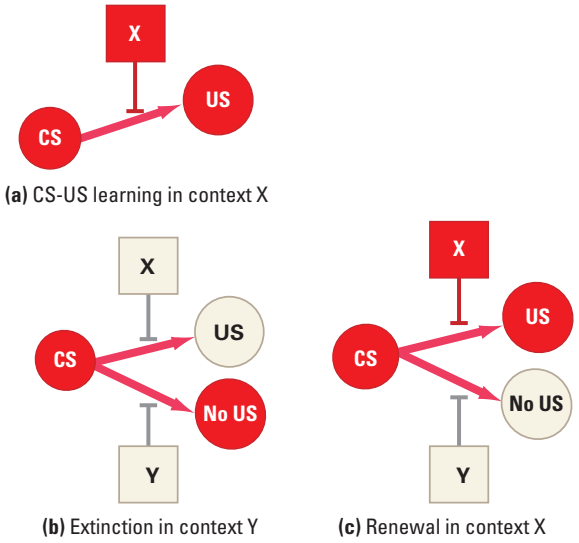


Figure 4.11 An account of extinction and renewal that involves context (a) CS–US learning in context X (b) Extinction in context Y (c) Renewal in context X
Information from Bouton, 1991.

(Figure 4.11a); then, in context Y, a CS–no US association is created. At this point, the CS–US association learned in X has not disappeared, even though there is no responding to the CS in context Y (Figure 4.11b). When the animal is returned to context X, the original CS–US association is retrieved, and the response "spontaneously" reappears (Figure 4.11c).

Another way in which extinction can be undone is if a long time passes before the animal is retested with a presentation of the CS. The return of a CR after such a delay is another example of spontaneous recovery (which you previously encountered in Chapter 3, in our discussion of habituation) and was first reported by Pavlov (1927). These findings (seen in almost all forms of classical conditioning) suggest that an association that became dormant following extinction training is not entirely lost.

Further evidence that extinction does not involve the total loss of what was previously learned comes from studies of rapid reacquisition, where a previously extinguished CS is learned more rapidly than a novel CS. Rapid reacquisition suggests that the learned association between a CS and a US is saved during extinction even if the behavioral CR is no longer evident following extinction trials.

Compound Conditioning and Overshadowing

When Robi was a young mother, her two children, Roberta and Mark, would often talk to her (or more often, shout at her) at the same time, competing for their mother's attention. As you can imagine, it was hard for Robi to pay full attention to both kids. The same competition can occur when two cues, such as a tone and light, appear together in a conditioning experiment, a paradigm known as **compound conditioning**. When an animal or person is trained in a tone + light ⇒ US compound conditioning paradigm, learning of the compound stimulus usually proceeds much like learning of a simple single-cue stimulus. However, when one of the stimuli, such as the tone, is later tested alone, it becomes apparent that something different has happened. Many studies have shown that the tone will have less association with the US if it is trained in compound than if it had been trained alone in a tone ⇒ US procedure. This suggests that two cues, like the two young children, Roberta and Mark, are competing against each other, and neither produces as much learning as it would have had it been trained alone.

Identify the conditioning paradigm illustrated here.

However, imagine if Roberta had a soft, quiet voice but Mark was prone to shouting out loud. As much as their mother might want to listen to both equally, it is likely that Mark would be the one more easily heard. Again, this can be analogous to compound conditioning studies: if one of the two stimuli, say the light, is very dim and hard to see, while the tone is loud and salient, the tone is said to overshadow the light during learning. In classical conditioning,

overshadowing occurs when a more salient cue within a compound acquires far more of the share of attention and learning than the less salient cue. For example, the learning that accrues to a loud tone in a loud tone + dim light → US compound training paradigm will be far more than to the dim light, because the loud tone is far more salient than the dim light. In contrast, if the tone was very soft and the light very bright, the opposite pattern would be seen, with the bright light overshadowing the soft tone and accruing more learning. Even so, both the cues, when tested individually, command less attention than they would have if they were trained alone and not in a compound.

compound conditioning. The simultaneous conditioning of two cues, usually presented at the same time.

overshadowing. A effect seen in compound conditioning when a more salient cue within a compound acquires more association strength, and is thus more strongly conditioned, than does the less salient cue.

Dissecting the interactions between cues that co-occur during learning led to fascinating insights into what and how animals are learning during conditioning experiments. From what you read previously in this section, it might appear that animals and people in conditioning studies are only passively responding to which cues do (or do not) go together. However, as you will see in the discussions to follow, conditioning involves a much more subtle form of learning than was previously appreciated, and the subtlety emerges most clearly when two or more cues are present during learning.

Error Correction and the Modulation of US Processing

Chapter 1 introduced Aristotle's argument that contiguity—closeness in time and space—is necessary for a new association, such as that between a CS and US, to be learned. For most of the first half of the twentieth century, psychologists believed that contiguity was both necessary and sufficient: so long as a potential CS and a US occurred with little separation in time and space, animals and people were expected to form an association between them (Hull, 1943). But would it really make sense for animals or people to learn associations between all the simultaneously occurring stimuli that they perceive? Would it even be possible?

The Informational Value of Cues

Imagine you are a struggling stock investor whose livelihood depends on correctly predicting whether the stock market will go up or down the next day. One morning Doris, an eager new stock analyst, walks into your office and says that if you hire her, she will tell you each day which way the next day's market will go. You agree, and during her first week of work, you are amazed to see that she is 100% accurate, correctly predicting each day whether the market will rise or fall. The next week, Herman comes to visit and offers you his services as a stock analyst to predict the market's movements. Would you hire him? Probably not, because he is redundant if you already have Doris; that is, Herman offers no value beyond what you are already getting from Doris. You might say that Doris's early success at predicting the stock market has blocked you from valuing Herman's similar, but redundant, ability to do the same, as summarized in Table 4.4.

A similar situation can occur in medical diagnosis. Consider a young physician who treated a patient, Janae, who came into the emergency room on several occasions in January, each time with a bad stomach ache. She reviews what Janae ate the day before each outbreak. It turns out that each time Janae got a stomach ache, she had eaten the same kind of chocolate that very morning. The physician suspects this may be the cause of the stomach ache, a bad reaction to the chocolate. The next month, Janae comes back to the emergency room, once again with a stomach ache. Again the physician asks her to list everything she ate the day of the outbreak. This time, she notices that Janae ate both chocolate

Table 4.4 Informational redundancy blocks learning

Group	Phase 1	Phase 2	Phase 3 (test)
Stock Prediction	Doris → stock market	Doris & Herman → stock market	*Hire Herman?* "No way; don't need him."
Medical Diagnosis	Janae eats chocolate → stomach ache	Janae eats chocolate & licorice → stomach ache	*Could the licorice be causing the stomach ache?* "Unlikely; Janae should enjoy licorice but avoid chocolate."
Bower & Trabasso (1965). See p. 136 and Figure 4.13	Circle → A Triangle B	Circle + top dot → A Triangle + bottom dot → B	Rectangle + top dot → ? Rectangle + bottom dot → ? "Participants have no idea which is A or B"

and red licorice just prior to her stomach ache. Although it's possible that the licorice caused the stomach ache, the physician dismisses this as unlikely, since she already believes that the chocolate is causing the stomach ache and this theory is sufficient to explain the new occurrence. As schematized in Table 4.4, the physician's recommendation to Janae, given her history of eating candy and getting stomach aches, is to avoid the chocolate but enjoy the licorice.

In both of these examples, new information—or evidence—can be viewed in terms of its informational value, how much new predictive value it gives us relative to what we already know or believe. For the stock investor, Herman offers no additional aid in predicting the market beyond what Doris already tells him. For the emergency room physician, the presence of licorice offers no new explanatory value in understanding the cause of Janae's stomach aches, given that she can already explain the problem as being caused by the chocolate. Research on learning has shown that humans and other animals are similarly sensitive to the informational value of cues in determining which associations they do or do not learn.

Kamin's Blocking Effect

Compare the story of Doris and Herman to the earlier story of siblings Mark and Roberta. In both cases there was competition going on. Mark beat out Roberta for their mother's attention because he shouted louder and was thus the more salient of the two children. In contrast, Doris beat out Herman for the stock-picking job not because she was louder, or even necessarily any better than Herman at the job, but because she got there sooner. From these stories we see that there are two ways to win a competition—by being more salient (i.e., louder) and by getting there sooner. These two situations mimic the two types of cue competition seen in conditioning paradigms. Competition based on salience is, as noted earlier, akin to what happens when overshadowing occurs in compound conditioning. But, a form of overshadowing can also occur between two cues when the associative or predictive value of one of the cues is learned earlier than the other, as suggested by the story of Doris and Herman (and by the chocolate and licorice example, as well). This kind of overshadowing due to temporal priority reflects a sensitivity to the informational value of one cue relative to another, co-occurring cue for making a prediction.

In the late 1960s, several studies of classical conditioning in animals made a similar point: for a potential CS to become associated with a US, the CS must provide valuable new information that helps an animal predict the future. Moreover, even if a given cue does predict a US, it may not become associated

with that US if its usefulness has been preempted (blocked) by a co-occurring cue that has a longer history of predicting the US. Much as Doris's predictive value blocked the hiring of Herman, and Janae's prior history of chocolate causing a stomach ache blocked the physician from believing that licorice was the cause, a prior-trained CS can block learning about another, redundant CS that is added later in training (Kamin, 1969).

In a classic study by Leon Kamin, rats were first trained that a light predicts a shock and later trained that a compound stimulus of a light and tone also predicts the shock (Kamin, 1969). Kamin found that, with this training, the rat will learn very little about the tone because the tone does not improve the rat's ability to predict the shock. This phenomenon is now formally known as **blocking;** it demonstrates that classical conditioning occurs only when a cue is both a useful and a nonredundant predictor of the future (Table 4.5).

blocking. A two-phase training paradigm in which prior training to one cue (CS1 → US) blocks later learning of a second cue when the two are paired together in the second phase of the training (CS1 + CS2 → US).

Kamin's 1969 blocking study is worth describing in detail because of its influence on subsequent theories of learning. In this study, one group of rats (the control group) was trained with a compound cue consisting of a light and a tone; this cue was reliably followed by a shock (see Table 4.5, control group, phase 2). The light and tone constituted a compound CS that the rats learned to associate with the shock US. Later, these rats would give a medium-strong CR to either the tone alone or the light alone, though not as strong a response as to both the light and tone together.

Consider, however, the behavior of Kamin's second group of rats, identified as the experimental, or pre-trained, group in Table 4.5. These rats first received pre-training in which the light by itself predicted a shock (phase 1). From this training, they learned an association between the light CS and the shock US. Next (phase 2), they were given training that paired the light-and-tone compound cue and the shock, just like the control group animals had received. However, unlike the control rats, rats in the pre-trained group were already responding strongly to the light CS when they began the phase 2 compound training. For these rats, the additional presence of the tone provided no new information for predicting the US.

Phase 3 was a testing phase. When the pre-trained rats were tested with the light alone, they continued to exhibit a strong CR to the light, much as they had at the end of phase 1. However, in phase 3, if they were tested with the tone alone, they would give almost no response at all. This suggests that they learned almost nothing about the relationship between the tone and the US, despite the compound training received in phase 2, in which the tone (combined with light) was repeatedly followed by the US (Kamin, 1969). In contrast, rats in the control group, which did not receive phase 1 pre-training, exhibited significant (albeit medium-strength) CRs to both the light by itself and the tone by itself in phase 3. Thus, the blocking phenomenon, exhibited by the pre-trained rats, can

Table 4.5 Kamin's blocking paradigm

Group	Phase 1	Phase 2	Phase 3 (test)
Control group	Rat sits in chamber; no training	Tone CS combined with light CS → shock US	Tone CS or light CS: medium CR
Experimental "pre-trained" group	Light CS → shock US	Tone CS combined with light CS → shock US	Tone CS: little or no CR (learning is "blocked")

be summarized as follows: prior training of the light → shock association during phase 1 blocks learning of the tone → shock association during compound (light + tone) training in phase 2. This is very similar to what happened with the emergency room physician. Having previously deduced that chocolate candies lead to a stomach ache, Janae's doctor failed to credit the licorice with any stomach-ache–inducing properties when she later encountered both chocolate and licorice preceding the next tummy ache. For this physician, learning the chocolate → tummy ache association blocked learning of the redundant licorice → tummy ache association (refer again to Table 4.4).

The Rescorla–Wagner Model of Conditioning

The blocking effect posed a challenge for simple theories of classical conditioning. It suggested that cues do not acquire strength solely on the basis of their individual relationships with the US; rather, *cues appear to compete with one another for associative strength*. Thus, in phase 2 of the blocking experiment in Table 4.5, the tone competes with the light, and in the case of the pre-trained group, the tone loses: Since the light already accurately predicts the US, the tone provides no additional predictive information (much as Herman provides no additional value over Doris in predicting stocks, or as knowing that Janae ate licorice presents no additional value for predicting her stomach ache if we already know that she ate chocolate; see Table 4.4).

The blocking paradigm demonstrated that contiguity between a cue and a US is not enough to elicit a CR, contrary to what Aristotle expected. For a stimulus to become associated with a US, it must impart reliable, useful, and nonredundant information (Kamin, 1969; Rescorla, 1968; Wagner, 1969). Apparently, "simple" Pavlovian conditioning is not as simple as psychologists once thought! In fact, rats (and other animals, including humans) appear to be very sophisticated statisticians. But how does one learn which are the most useful and informative cues to remember?

In the early 1970s, two psychologists at Yale, Robert Rescorla and Allan Wagner, were independently trying to understand Kamin's blocking effect and other related conditioning phenomena. Although the two researchers worked at the same university, they didn't realize that they were using the same approach to solve the same problem until they happened to take a train together to a conference and began chatting about their research. To their surprise, they realized that they had each come up with the same idea, and they decided to join forces (Rescorla & Wagner, 1972).

Rescorla and Wagner sought to understand how animals become aware of the informational value of stimuli, and they developed an elegantly simple way to formalize their approach. The key idea behind the Rescorla–Wagner model is that changes in CS–US associations on a trial are driven by the discrepancy (or error) between the animal's expectation (or prediction) of the US and whether or not the US actually occurred. This error is sometimes referred to as the **prediction error**, and how one can learn using these errors is described in the next section.

prediction error. The difference between what was predicted and what actually occurred.

Error-Correction Learning

"I have not failed," said Thomas Edison, "I've just found 10,000 ways that won't work." Like Edison, we can, and do, learn from our failures. Consider Herman, who we saw earlier didn't have much of a future as a stock analyst. His real passion is tennis, and he hopes to become a professional. He practices hours each day, focusing especially on his serve. His goal is to put the ball as far back in the serving zone as possible, but not so far that it goes over the line and is considered a fault. On his first serve of the day, he puts the ball into the middle of the serving

zone (where his opponent could easily return it). To serve the next ball better, he adjusts his stance a little, throws the ball higher, and swings harder. This time the ball goes closer to the end of the zone. On the third serve, he throws the ball a little higher still and hits it just a bit harder. This time, however, the ball goes too far; it is a fault. In response to this error, Herman again adjusts his serve. He throws the ball a little less high and hits it a little softer than the last time. This results in a perfect serve. After several tries, correcting his serve each time on the basis of how well he did previously, Herman has learned to hit a ball at just the right height and strength to make it land at the outer end of the service zone: not too close, and not too far.

Herman has learned through a process of trial and error, using a method called **error-correction learning**, in which the errors on each trial lead to small changes in performance that seek to reduce the error on the next trial. Error-correction learning can be summarized by describing three situations representing three types of errors Herman experienced: (1) when Herman hits the ball too close, he changes his next serve to hit the ball harder, so that it will land a little farther; (2) when Herman hits a perfect serve, he tries to do exactly the same thing the next time; and (3) when Herman hits the ball too far, he changes his next serve to hit the ball a little lighter so it lands a little closer next time.

Rescorla and Wagner proposed that there are three key situations to consider in interpreting a prediction error, as summarized in Table 4.6, and they are very similar to the three ways Herman learned from past errors to improve his tennis serve. One is a situation in which either no CS or a novel CS is presented followed by a US, so that the US will be unexpected; this is considered a positive prediction error because there is more US than expected. The Rescorla–Wagner theory expects that the CS → US association should increase proportional to the degree that the US is surprising; that is, the larger the error, the greater the learning. This makes sense because if you failed to predict the US, you want to increase (move in a positive direction) your likelihood of predicting it in the future (given the

Bill Losh/Getty Images

Every time Herman serves too far outside the box, he learns to correct his swing, so that next time he hits the ball not so far, exhibiting what kind of learning?

error-correction learning. A mathematical specification of the conditions for learning that holds that the degree to which an outcome is surprising modulates the amount of learning that takes place.

Table 4.6 Error correction and response in the Rescorla–Wagner model and tennis

Conditioning error	R–W model response	Tennis error	Herman's response
Positive error: CS predicts nothing or too little, but US unexpectedly occurs or is unexpectedly strong	Increase association	Ball falls short	Increase strength of serve
No error: CS predicts US, and predicted US occurs	No new learning	Ball lands perfectly	Do same thing next time
Negative error: CS predicts US, but no US occurs	Decrease association	Ball goes too far	Decrease strength of serve

same CS). It is similar to what Herman does when his serve is too short and he positively increases the strength of his serve to send it farther next time.

If, however, a well-trained CS is followed by the expected US, there is no error in prediction (the US was fully predicted by prior presentation of the CS), and thus no new learning is expected. This is similar to what happens when Herman makes a perfect serve; he doesn't want to change a thing from what he did last time. Finally, if the CS predicts a US and the US does not occur, the prediction error is considered negative, and Rescorla and Wagner expect it to be followed by a decrease in the CS → US association. This is similar to what happens when Herman hits the ball too far and he has to reduce the strength of his serve next time.

Associative Weights and Compound Conditioning

associative weight. In the Rescorla–Wagner model of conditioning, a value representing the strength of association between a conditioned stimulus (CS) and an unconditioned stimulus (US).

The Rescorla–Wagner model assumes that each CS has an **associative weight**, which is a value representing the strength of association between that cue and the US. In the blocking experiment described above and in Table 4.5, there would be two cue weights, one for light and one for tone. Think of these weights as numbers that indicate how strongly the CS predicts the US. Before any training takes place, all associative weights are 0.0, meaning that when a potential CS first appears, there is no expectation that any US will follow. These associative weights change through learning as the animal discovers which stimuli predict the US and therefore which should have strong weights.

A critical property of the Rescorla–Wagner model is that the weights associated with one cue can indirectly influence the weights accruing to other co-occurring cues. That is, if a tone and a light are both present on a trial, they will compete for associative strength (much like Doris and Herman compete for the stock analyst job). This competitive property of the Rescorla–Wagner model allows it to account for many important conditioning phenomena, especially those with complex stimuli involving the presentation of multiple stimulus elements (such as tones and lights paired together). Most importantly, this cue-competition property of the Rescorla–Wagner model allows it to account for Kamin's blocking effect as described next.

Using the Rescorla-Wagner Model to Explain Blocking

Consider the simple network model of eyeblink conditioning in Figure 4.12. For every aspect of the real world that the model represents, the model contains a node—a small element that is activated, or "turned on," when the network believes that a certain aspect of the world is or will be present. You can think of these nodes as being something like abstract neurons, or collections of neurons, which process information in a manner analogous to how activation flows through the brain. In the network model of Figure 4.12 there is an input node for each CS (tone and light), an "actual output" node for the eyeblink CR, and a "teaching node" that shows whether or not the US (airpuff) actually occurred and a CR blink would have been appropriate (that is, the "desired output").

Figure 4.12a represents the state of affairs after phase 1 training in which a tone predicts the airpuff US: the tone has acquired an association weight of 1.0 and the light has a weight of 0.0. Now, at the start of phase 2, the light and tone are presented together (Figure 4.12b). The actual output node receives activation equal to the sum of the association weights of all the active (present) cues, which in this case is 1.0 for the tone plus 0.0 for the light, for a total of 1.0. This is the output node's activation in Figure 4.12b. Since the airpuff US is also administered in this trial, the response is correct—and the error is 0.0. Since weight change is determined by the error and the error is 0.0, there is no learning on this trial—or, indeed, on any other phase 2 trial. By the end of

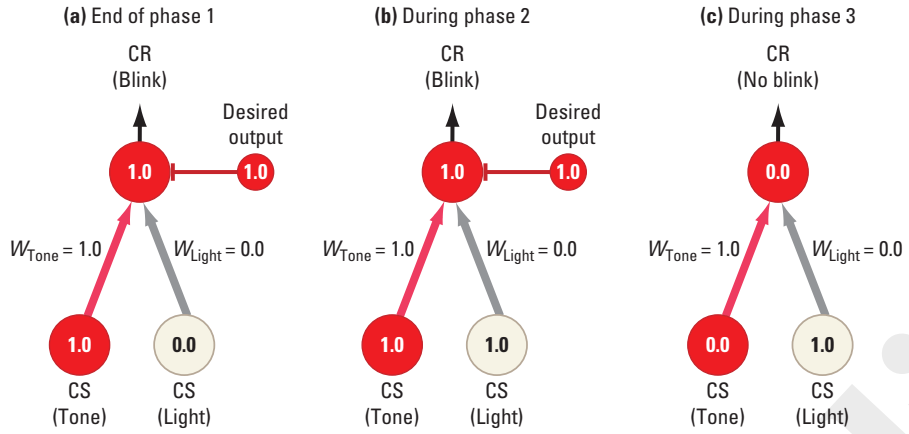


Figure 4.12 Blocking in the Rescorla–Wagner model (a) Phase 1 involves repeated pairing of the tone CS with an airpuff US. At the end of this phase, there is a strong weight for the tone input node, which in turn can activate the output node. (b) Phase 2 involves presentation of both tone and light CSs, paired with the US. Since the tone already perfectly predicts the US, causing the output node to be fully activated, there is no output error and no learning occurs. Hence, the weight from the light input node remains at 0.0. (c) Finally, in phase 3, the network is tested with presentations of light alone. Since the weight of the light node was never changed from 0.0, there is no activation of the output node, and no behavioral response occurs. Hence, the Rescorla–Wagner model correctly shows blocking.

phase 2, the network weights will still be the same as they were at the beginning of phase 2: the tone is strongly weighted, and the light is not. A subsequent test presentation of the light alone in phase 3 is shown in Figure 4.12c: There is no association weight on the light input node, so there is no activation of the output node—and therefore there is no response to the light alone. Thus this simple graphical network model representation of the Rescorla–Wagner model shows Kamin's blocking effect, just like the pre-trained animals in Table 4.5.

As compared to the experimental pre-trained animals, the control rabbits (Table 4.5) get no training at all in phase 1, so the expected US starts at 0.0 and remains at 0.0 from start to end of phase 1. In phase 2, where the tone-and-light compound cue is paired with the US, there is a big error (1.0) on the first trial when the US appears totally unexpectedly. Over the course of the training trials in phase 2, the association weights for both the tone and light CSs will rise together until both are equally (but only mildly) associated with the US—but their sum will be enough together to predict the US, leading to an expected US of 1.0 by the end of phase 2 and an error that has been brought down over the course of phase 2 from 1.0 to 0.0. In a subsequent testing phase (phase 3), a medium-strong response is given to either the tone or the light if they are presented individually, because the association weights for both tone and light alone are only at half the strength required to predict the US (since they have always been presented together as a pair, with their associations combined and summed during phase 2).

Influence of the Rescorla–Wagner Model

More than four decades after its publication, the Rescorla–Wagner model is generally acknowledged as the most influential formal model of learning. Its broad acceptance is due to its elegant simplicity and to the fact that it explains a wide range of previously puzzling empirical results. One hallmark of a successful model is that it reveals underlying connections between a series of observations that initially seemed unrelated or even contradictory.

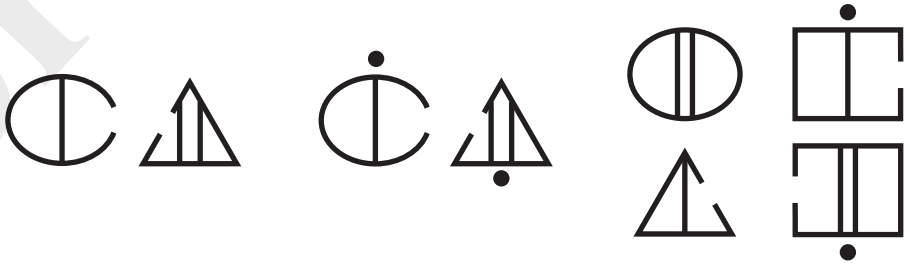
The Rescorla–Wagner model also made surprising predictions about how animals would behave in *new* experimental procedures, and experimenters rushed to test those predictions. This is another feature of a successful model: it should allow scientists to make predictions that could not otherwise be foreseen without the model. Ideally, modeling and empirical work should constitute a cycle in which the model makes predictions that, when tested, provide new empirical data. If the data match the predictions, the model is supported. If not, then the model must be revised. The revised model then generates new predictions, and the cycle continues.

Owing to its simplicity, the Rescorla–Wagner model cannot account for every kind of learning, and should not be expected to. However, many researchers have devoted their careers to showing how one or another addition to the model would allow it to explain a wider range of phenomena. With so many additions, the model may be in danger of losing some of its clarity and appeal. Nevertheless, the Rescorla–Wagner model has been the starting point from which many other promising models have been built, including the models of human learning discussed below.

Error Correction in Human Category Learning

Do concepts such as blocking and models such as the Rescorla–Wagner model apply only to classical conditioning, or might they also provide insights into higher forms of human cognition and behavior, especially those that involve prediction or categorization?

Early research evidence for blocking-like effects in humans came from work by Gordon Bower and Tom Trabasso, who used error-correction–type methods to train college students to categorize objects according to certain predefined rules (Bower & Trabasso, 1964). The students were presented with geometric figures varying in five dimensions: color, shape, number of internal lines, position of a dot, and position of a gap (non-colored examples are shown in Figure 4.13). Phase 1 of the experiment consisted of training the participants by asking them to guess whether each figure belonged to class A or class B; each time, they were told whether they had guessed correctly or not. For example, some participants were trained that all circular shapes belong in class A, while all triangular shapes belong in class B (and all other features are irrelevant), as illustrated by the two



(a) Phase 1 training (b) Phase 2 training (c) Testing

Figure 4.13 Human sensitivity to the informational value of cues (a) Examples of stimuli from phase 1 training of the Bower and Trabasso experiment, in which all circular shapes belong in class A and all triangular shapes belong in class B. (b) Examples of stimuli from phase 2. Participants are shown only circles and triangles, and the same circle → A/triangle → B rule still applies. However, now there is also a dot on the top of all class A items and a dot on the bottom of all class B items. (c) A final testing phase. Participants are given novel stimuli to see if they have learned that the dot-top → A/dot-bottom → B rule by itself can predict class membership.
Research from Bower & Trabasso, 1964.

sample stimuli shown in Figure 4.13a, and schematized in Table 4.4. Given enough trials with different stimuli, participants would deduce the rule: circle → A/triangle → B.

Once this lesson was mastered, the experimenter showed participants a slightly different set of figures: now all figures that were circular and thus belonged to class A had a dot on top, while all figures that were triangular and thus belonged to class B had a dot on the bottom (Figure 4.13b). This addition of a redundant cue in phase 2 (position of the dot) parallels the addition of the licorice stimulus in the second month of the physician trying to figure out what caused her patient's stomach ache. Participants in the Bower and Trabasso study continued to perform well by using their old rule of sorting on the basis of shape; the question was whether they would also learn that the dot position by itself predicted class membership.

To test this, the experimenters used new figures, shown in Figure 4.13c. Given a figure with no dot, all participants continued to sort the circles into class A and the triangles into class B. However, when given a figure with a new shape (rectangle), none of the participants correctly sorted on the basis of dot position. Thus, these humans performed much like the physician (who dismissed the possibility that licorice could be causing her patient's discomfort), in that they displayed little or no response to the redundant cue added in phase 2. In effect, prior learning that the shape predicted class membership appears to have *blocked* subsequent learning that the dot position also predicted class membership. More recent studies have verified that this tendency to "tune out" information that is redundant with regard to previously learned cues is pervasive in many forms of human learning (Kruschke, Kappenman, & Hetrick, 2005).

Although the fields of animal and human learning were originally closely intertwined, they became largely divorced from each other in the late 1960s and early 1970s. Animal learning at that time remained primarily concerned with elementary associative learning, while human learning studies focused more on memory abilities, characterized in terms of information processing and rule-based symbol manipulation, approaches borrowed from the emerging field of artificial intelligence. Ironically, this schism occurred just as animal learning theory was being reinvigorated by the new Rescorla–Wagner model in the early 1970s.

In the late 1980s, the expanding impact of computer simulations of network models of human learning (also called connectionist models) revived interest in relating human cognition to elementary associative learning processes. Some of this work was discussed in Chapter 1, when we reviewed David Rumelhart's contributions to the field. Because of the growing influence of these connectionist network models in cognitive psychology, the people working on simpler associative processes, such as classical conditioning, were motivated to re-explore the Rescorla–Wagner model. The connectionist network models adapted their association weights using a generalized (and more powerful) variation on the Rescorla–Wagner model to show how numerous complex human abilities (including speech recognition, motor control, and category learning) might emerge from configurations of elementary associations similar to those studied in conditioning experiments.

An example of one such linking of conditioning and cognition is a simple neural network model developed by Mark Gluck and Gordon Bower to model how people learn to form categories (Gluck & Bower, 1988). In this study, college students were asked to learn how to diagnose patients suffering from one of two nasty-sounding (but fictitious) diseases—midosis or burlosis. The students reviewed medical records of fictitious patients, who were each suffering from one or more of the following symptoms: bloody nose, stomach cramps, puffy

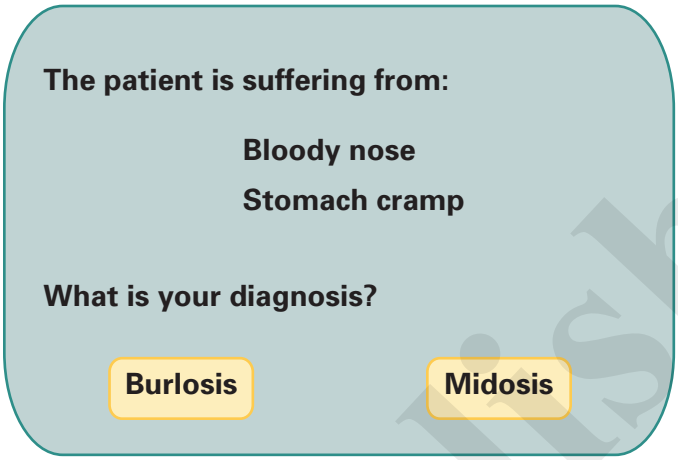


Figure 4.14 A sample training trial in Gluck and Bower's probabilistic categorization task On a particular learning trial, a research subject would see some symptoms (e.g., bloody nose and stomach cramp) and make a diagnosis, and then be given feedback as to whether or not the diagnosis was correct.

eyes, discolored gums. During the study, each student reviewed several hundred medical charts, proposed a diagnosis for each patient, and then was told the correct diagnosis (Figure 4.14). The students initially had to guess, but with practice they were able to diagnose the fictitious patients quite accurately. The fact that the different symptoms were differentially diagnostic of the two diseases helped them improve. Bloody noses were very common in burlosis patients but rare in midosis, while discolored gums were common in midosis patients but rare in burlosis.

This kind of learning can be modeled using the network in Figure 4.15. The four symptoms are represented by four input nodes at the bottom of the network, and the two disease categories correspond to the two output nodes at the top of the network. As learning progresses, the weights of the arrows leading from the symptoms to the diseases are updated according to the learning rule from the Rescorla–Wagner model, much as if the symptoms were CSs and the diseases were alternate USs.

Learning and performance in the model works as follows. As the subject learns the correct associations, the connections between nodes and diseases acquire weights that capture the true diagnosticity of each cue for each disease. For example, presentation to a research subject of the chart of a patient with the symptoms "bloody nose" and "stomach cramp" (as in Figure 4.14) is modeled in Figure 4.15 by turning "on" the corresponding input nodes. Activating

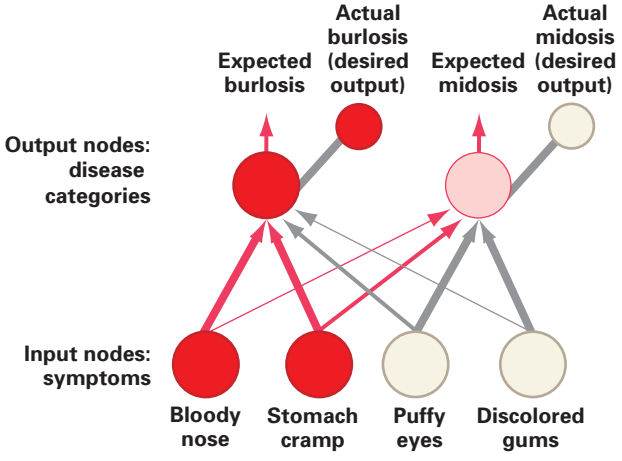


Figure 4.15 Gluck and Bower's network model of category learning The arrows from bloody nose and stomach cramp to burlosis and from puffy eyes and discolored gums to midosis are thick, indicating highly diagnostic relationships (that is, heavily weighted cues). The other cues are of only moderate diagnosticity. This figure shows a trial in which a patient presents with two symptoms, bloody nose and stomach cramp; thus, these two input nodes are active (dark red). The other two input nodes represent symptoms that are not present (puffy eyes and discolored gums), and these nodes are inactive (gray). Relative activation levels (dark red and light red) of the two "expected" category nodes are based only on the weight of the input (the current associative weights of these cues) flowing up the arrows from the active and present symptoms (bloody nose and stomach cramp).
Information from Gluck & Bower, 1988.

these two input nodes causes simulated neural activity to travel up four weighted connections, shown as arrows: two to burlosis and two to midosis (Figure 4.15). On this trial, burlosis is the correct label, indicated by the very dark "actual burlosis" node. The model is more likely to diagnose the patient as having the disease with the higher activation, namely, burlosis, which in fact is the correct diagnosis.

By analogy with the Rescorla–Wagner model, these output-node activations are equivalent to the network's *expectation* of one disease versus another. After a student guesses at a diagnosis and is told the correct answer, learning in that student is modeled by modification of the network's weights so as to reduce future error, in accordance with the Rescorla–Wagner model's error-correction learning rule. The network model shown in Figure 4.15 incorporates nothing more than the learning principle of the Rescorla–Wagner conditioning model, yet this "animal conditioning" model of human cognition accounts for variations in how the participants classified different patient charts. The model correctly predicted the percent of participants who would classify each of the 14 possible symptom charts as being midosis versus burlosis and also predicted how well the participants were later able to make judgments about the probabilities of the two diseases when they only knew one of the symptoms.

Cue–Outcome Contingency and Judgments of Causality

Another area in which classical conditioning and cognitive studies of category learning have converged is the study of cues that are only partially valid predictors of category membership. Consider, for example, what would happen if Doris, the stock analyst you recently hired, was a good but not perfect stock predictor. Suppose her predictions are correct on 3 out of every 5 days. That rate is not bad, but you yourself may already be able to make accurate predictions about the stock market 3 out of 5 days just from reading the *Wall Street Journal*. In that case, you might decide that Doris doesn't provide you with any additional useful information. If your ability to invest wisely is the same regardless of whether or not Doris is helping you, you probably wouldn't view her as a great asset to your business.

Rescorla showed a similar phenomenon in an animal conditioning experiment that provided additional support for the Rescorla–Wagner model (Rescorla, 1968). His experiment demonstrated that conditioning to a tone stimulus depends not only on the frequency of tone–US pairings but also on the frequency of the US in the absence of the tone. If the US occurs just as often without the tone as it does in the presence of the tone, then little or no conditioning will accrue to the tone. These results suggest that animals are sensitive to the *contingency* of (or degree of correlation between) the potential CS and the US. The Rescorla–Wagner model explains this effect by viewing the experimental chamber itself as a cue presented in combination with (compounded with) the experimentally manipulated tone. The experimental chamber can be thought of as the *context*, that is, the background stimuli that are relatively constant on all trials (rather than being manipulated by the experimenter), both when there is a US and when there is not; these stimuli include the sound, smell, and feel of the conditioning chamber. In the stock investor example, the context includes all the generally available information for investors, such as the stock analyses in the daily *Wall Street Journal*; the potential CSs are the extra tips occasionally provided by Doris.

In the Rescorla–Wagner model, the animal actually experiences the trials in which the tone occurs alone as trials in which a compound cue is present, a cue consisting of the tone CS in combination with the context. The Rescorla–Wagner model expects that the context will, in effect, compete with the tone for the credit of predicting the US. If the US occurs as frequently on

context-alone trials as on context-and-tone trials, the context is a more reliable cue, and thus it wins the credit and, hence, the bulk of the associative weight. Therefore, according to the Rescorla–Wagner model, the degree to which the US is contingent on the CS depends on a competition between the CS and the co-occurring background context.

Similar sensitivity to cue–outcome contingencies has also been found in studies of human causal inference. These are studies of how people deduce cause and effect in their environment. In typical experiments, people might be asked to judge which risk factors (smoking, lack of exercise, weight gain) are more or less responsible for some observable outcome, such as heart disease. These studies have shown that increasing the frequency of the outcome in the absence of the risk factor (say, the frequency of lung cancer in the absence of smoking) decreases people's estimates of the causal influence on the outcome—in much the same way that the presence of the US in the context alone decreased conditioning to the potential CS as described above. What are the implications of this finding? For one thing, it suggests that if there is a spike in the frequency of a disease (like lung cancer) but no similar increase in a risk factor (like smoking), people will come to view smoking as less harmful than they did previously. In effect, if you're going to get lung cancer anyway, why not smoke?

Stimulus Attention and the Modulation of CS Processing

Despite the many successes of the Gluck and Bower model in predicting human cognitive data, several limitations of the model became evident in further studies of human category learning. In particular, as a model of category learning, it fails to account for people's ability to actively focus their attention on one or another symptom (such as bloody nose *or* stomach cramp) or to shift or refocus this attention during learning. These limitations echo similar problems of the Rescorla–Wagner model (on which the Gluck and Bower model is based), specifically its inability to account for how attention to stimuli is modulated during learning, especially when animals or people are repeatedly exposed to stimuli that have no consequence.

To better understand where and how the Rescorla–Wagner model (and by analogy the Gluck and Bower model) falls short in accounting for attentional changes, let's consider the example of Moira, the little girl at the beginning of the chapter who learned to run to the corner with money each time she heard the ice cream truck jingle in the distance. What do you think would have happened if, soon after she arrived in the neighborhood, she continued to hear the jingle off in the distance each day, but the ice cream truck never arrived? If weeks later the truck did begin to appear again after the jingle was heard, do you think Moira would so quickly learn to associate the jingle in the distance with the forthcoming arrival of the ice cream truck? Prior exposure to the cue's having been irrelevant (like the jingle without the ice cream truck appearing) retards our ability to learn later that the cue has acquired some new predictive or associative relevance.

latent inhibition. A conditioning paradigm in which prior exposure to a CS retards later learning of the CS–US association during acquisition training.

This learning about a cue's irrelevance through exposure to the cue alone (with no associated significant event) is quantified by a measure known as **latent inhibition**, a reduction in learning about a stimulus (CS) to which there has been prior exposure without any consequence (that is, no US). The name refers to the fact that the exposure latently (that is, implicitly) appears to inhibit later learning about the cue. This phenomenon was first described in the animal conditioning literature by Robert Lubow and Ulrich Moore (Lubow &

Moore, 1959). Lubow and Moore's study was conducted using sheep and goats; however, for consistency with the rest of this chapter (and to facilitate comparison with other studies previously discussed), we will describe their latent inhibition paradigm using rabbit eyeblink conditioning, which has reliably produced the same results.

Table 4.7 The latent inhibition paradigm

Group	Phase 1	Phase 2
Control group	Animal sits in chamber	Tone CS → → → shock US
Experimental "pre-trained" group	Tone CS presented (but not US)	

Latent inhibition studies use two groups of subjects: the first group, the control group, receives no pre-training, and the second group does receive pre-exposure training, as summarized in Table 4.7. Control animals simply sit in their chambers until they are ready for the critical phase 2, in which they are trained to associate a tone CS with an airpuff-in-the-eye US. In contrast, animals in the pre-exposed group are repeatedly exposed to a tone with no US in phase 1 before they undergo the same tone training in phase 2 as the control animals do. Thus, the only difference between the two groups is that one group is pre-exposed to the tone in phase 1.

As illustrated in Figure 4.16, rabbits in the pre-exposed group learn to associate the tone with a puff of air much more slowly in phase 2 than do rabbits in the control group (Shohamy, Allen, & Gluck, 2000). The same kind of slow learning following CS pre-exposure is seen in a variety of species; for example, it is seen in human eyeblink conditioning as well (Lubow, 1973).

Latent inhibition—that is, impaired learning following cue pre-exposure—is problematic for the Rescorla–Wagner model: there is no surprise during the first phase of tone-alone exposure and thus no prediction error. Therefore, the Rescorla–Wagner model expects no learning to occur in phase 1. As a consequence, the Rescorla–Wagner model makes the incorrect prediction that the pre-exposed group should be no different from the control group at the start of phase 2, a prediction clearly disconfirmed by Lubow's studies, as well as by the data in Figure 4.16.

Latent inhibition and similar paradigms that involve learning during mere exposure to apparently neutral cues suggest that there is more going on during conditioning than the error-driven learning characterized by the Rescorla–Wagner model. Chapter 3 introduced some of these non-associative mechanisms, such as habituation and sensitization. To account for latent inhibition and other phenomena beyond the scope of the Rescorla–Wagner model, several alternative theories of conditioning have been proposed, and these are described next.

US modulation theory. Any of the theories of conditioning that say the stimulus that enters into an association is determined by a change in how the US is processed.

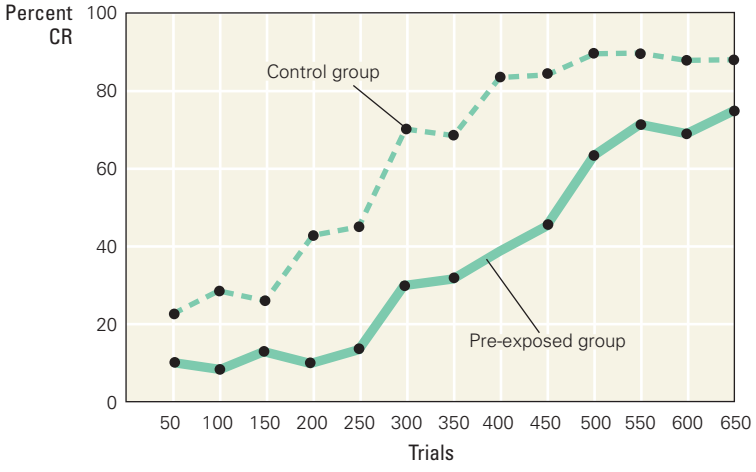


Figure 4.16 Latent inhibition in rabbit eyeblink conditioning This graph shows the percent of trials producing CRs in each block of 50 trials during the common phase 2 tone–airpuff training of the rabbits in the study. The rabbits in the control group (dotted line) learned rapidly. In contrast, the rabbits in the pre-exposed group (solid line), who had previously experienced 850 trials of tone-alone presentations in phase 1, learned much more slowly.
Data from Shohamy et al., 2000.

An Attentional Approach to Stimulus Selection

The Rescorla–Wagner model is often called a **US modulation theory** of learning because it proposes that the manner in which the US is processed determines what stimuli become associated with that US. Thus, in the Rescorla–Wagner model, the ability of the US to promote learning is modulated by how unexpected the US is, given the potential CS that precedes it. The error-correction principle of learning described above—the core idea behind the Rescorla–Wagner

model—is that error modulates the ability of a US to promote learning about the CSs that preceded it.

CS modulation theory. Any of the theories of conditioning holding that the stimulus that enters into an association is determined by a change in how the CS is processed.

An alternative class of learning theories focuses instead on the CSs, suggesting various mechanisms that modulate (either increase or reduce) the ease with which potential CSs can be incorporated into an association. For this reason, they are referred to as **CS modulation theories:** they propose that the way attention to different CSs is modulated determines which of them become associated with the US. One such theory, presented by Nicholas Mackintosh in the early 1970s, is based on the observation that people and animals have a limited capacity for processing incoming information (Mackintosh, 1975). This limited capacity means that paying attention to one stimulus diminishes (and hence modulates) our ability to attend to other stimuli.

Remember the blocking analogy in which Doris was the first to establish herself as a reliable predictor of the stock market so that later, when Herman showed up, you gave him little credit for making equally successful predictions? (See also Table 4.4.) The Rescorla–Wagner model argues that this outcome is due to the stock market (the US) already being well predicted by Doris (the first CS), so that no additional value (no learning) is attached to Herman (a potential second CS). However, Mackintosh's view of blocking is quite different. He argues that you come to devote all of your attention to Doris because she has a long history of predicting the stock market, and therefore you have no attention left to pay to Herman. The core idea of the Mackintosh theory is that a previously conditioned stimulus derives its salience from its past success as a predictor of important events (Mackintosh, 1975), and this happens at the expense of other co-occurring cues that don't get access to your limited pool of attention. In essence, Rescorla and Wagner's model lets Herman come in for an interview but doesn't consider him valuable for predicting the market, while Mackintosh's model never lets Herman in the door.

In addition to Mackintosh, several other learning theorists, most notably John Pearce and Geoffrey Hall, have proposed alternative hypotheses of how CS salience is modulated during training (Pearce & Hall, 1980). All of these models share the basic underlying idea that the changes in weighting of the CS are due to modulations of the CS, not of the US. We turn next to review one of the situations in which the CS-modulation theory of Mackintosh does a better job than the US-modulation theory of Rescorla and Wagner in explaining some aspects of learning phenomena.

An Attentional Explanation of Latent Inhibition

Recall that the Rescorla–Wagner model cannot explain cue–pre-exposure phenomena such as latent inhibition because, as a US modulation theory of learning, it only explains learning that takes place when a US is present or when previously trained cues predict the US. Thus, the Rescorla–Wagner model suggests incorrectly that no learning takes place when a neutral (previously untrained) cue is presented. In contrast, Mackintosh's model predicts that the salience of a tone as a potential CS will decrease when the tone is presented without any US because the tone develops a history of predicting nothing. According to Mackintosh, the animal treats these tone-alone trials as if they were the little boy who cried wolf. Eventually the tones (like the boy) are ignored because they don't reliably predict that anything bad or good is about to happen.

Although these CS modulation models have had many successes, especially in explaining behavioral phenomena that are not explained by the Rescorla–Wagner model, they have had less of an impact on the field of learning and memory, in part because they are more complex than the Rescorla–Wagner

model and because they don't explain as broad a range of behaviors. Moreover, as discussed earlier in this chapter, the Rescorla–Wagner model has been especially influential because it works on the same fundamental principle as the learning algorithms employed in the connectionist network models of human memory used by cognitive psychologists, including both the models of David Rumelhart and colleagues described in Chapter 1 (Rumelhart & McClelland, 1986) and the category learning model of Gluck and Bower (1988) discussed above.

Which view is correct, the CS modulation or the US modulation approach to conditioning? For many years the two camps were viewed as being in direct conflict, with each entrenched on a different side of the Atlantic Ocean: the US modulation view predominated in the United States (where Rescorla and Wagner worked), while the CS modulation view predominated in the United Kingdom (where Mackintosh, Pearce, and Hall worked). However, behavioral and biological studies of conditioning now suggest that *both* views are probably correct; that is, there are likely to be both CS modulation and US modulation mechanisms engaged in learning. As you will see in Section 4.2, part of what has helped resolve this debate is new data from neuroscience that have identified differential neural substrates for these two types of learning processes. This is one more example of the many areas where new forms of data from neuroscience have informed and helped resolve long-standing questions in psychology.

Test Your Knowledge

Contrasting the Rescorla-Wagner and Mackintosh Models

1. Fill in the blanks: The Rescorla–Wagner model explains conditioning as modulation of the effectiveness of the ____ for learning, while the Mackintosh model explains conditioning through modulation of attention to the ____.
2. From the examples below, which of these explanations of Connie's behavior would be best explained by the Rescorla–Wagner model? Which would be better explained by the Mackintosh model?
   - **a.** Connie loved the oatmeal raisin cookies so much, she devoted all of her attention to them. She didn't even bother tasting the chocolate chip cookies.
   - **b.** Connie was happy eating only the oatmeal raisin cookies, and she didn't feel any need to begin eating a new type of cookie.

(Answers appear in the back of the book.)

Other Determinants of Conditioning

Both the US-modulation model of Rescorla and Wagner and the CS-modulation model of Mackintosh have been influential in enhancing our understanding of associative learning (Rescorla & Wagner, 1972; Mackintosh, 1975). They are powerful models precisely because they reduce the behavioral process of learning to its essential elements so that we can see the underlying, fundamental principles at work. However, as a result of such simplification, these models necessarily ignore many of the more subtle facets of conditioning, such as the role of timing in conditioning and the importance of innate biases for associating different stimulus cues.

Timing

trial-level model. A theory of learning in which all of the cues that occur during a trial and all of the changes that result are considered a single event.

The Rescorla–Wagner model and the Mackintosh model both treat classical conditioning as if it were always composed of a series of discrete trials that occur one after the other. Moreover, these **trial-level models** treat each trial as a single event, resulting in a single change in learning. In reality, conditioning is more complex, and a trial consists of many events that can vary in different ways from trial to trial. For example, these models don't describe the timing of the animal's response within a given trial: does the CR occur right after the CS begins, or is it delayed until just before the US occurs? This information is lost in a trial-level model that only describes the aggregate effect of a training trial in terms of an overall association strength. Thus, one cost of having a simple and powerful model is that it can't account for every detail of the animal's behavior.

delay conditioning. A conditioning procedure in which there is no temporal gap between the end of the CS and the beginning of the US, and in which the CS co-terminates with the US.

trace conditioning. A conditioning procedure in which there is a temporal gap between the end of the CS and the beginning of the US.

One important aspect of many conditioning studies is the temporal relationship between the CS and the US. Figure 4.17a illustrates eyeblink conditioning that is conducted using an approach known as **delay conditioning**, in which the tone CS continues throughout the trial and only ends once the US has occurred (this is, in fact, how all of the animals were trained in the rabbit eyeblink-conditioning studies reported so far in this chapter). The term "delay" refers to the delay from time of onset of the CS to the onset of the US (as opposed to them occurring simultaneously). Another form of conditioning also includes a delay, but here the CS is first turned off before the US begins. This training procedure is called **trace conditioning** and is represented in Figure 4.17b; it uses a shorter CS that terminates some time before the onset of the US, requiring the animal to maintain a memory "trace" of the CS to associate with the subsequently arriving US. Although many trial-level learning models treat these types of conditioning as if they were equivalent, many studies have shown that learning behaviors, and the neural substrates associated with them, can be quite different for trace and delay-training procedures.

interstimulus interval (ISI). The temporal gap between the onset of the CS and the onset of the US.

Even within a simple delay-training procedure such as that shown in Figure 4.17a, variations in the **interstimulus interval (ISI)**, the temporal gap between the onset of the CS and the onset of the US, can have significant effects. For eyeblink conditioning in the rabbit, the optimal ISI for fastest learning is about one-quarter of a second (250 msec), as shown in Figure 4.17c.

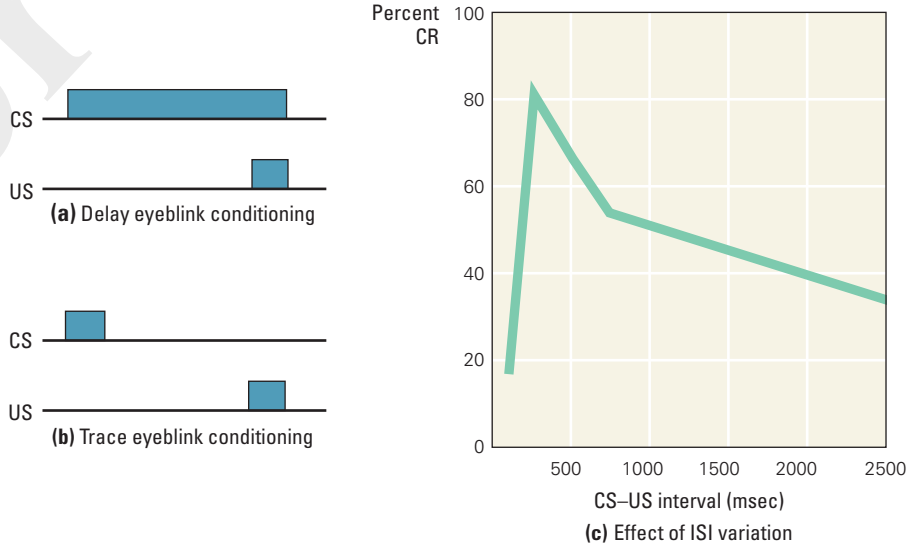


Figure 4.17 Delay and trace forms of eyeblink conditioning (a) In delay conditioning, the CS continues throughout the training trial and only terminates when the US terminates. (b) In trace conditioning, a short CS is followed by a gap before the US occurs. (c) The percentage of conditioned eyeblinks as a function of the length of the interval between the onset of CS and the US in a delay-conditioning experiment.
Data from McAllister, 1953.

Shorter or longer intervals make learning more difficult for the animal and necessitate additional training trials. One of the remarkable aspects of rabbit eyeblink conditioning is that the timing of the CR corresponds exactly to the ISI (see Figure 4.8), so that the eyelid is maximally closed at precisely the moment the onset of the US is expected.

Researchers have begun to integrate both US- and CS-modulation learning theories into unified learning theories that also accommodate some of the subtle temporal aspects of learning. One notable early example is the work of Allan Wagner, who proposed a model called SOP (for sometimes opponent process) that allows both for error-correction learning (US modulation) and for changes in the salience of CS cues (CS modulation), with these events occurring at different times through different processes (Wagner, 1981). Other researchers, too, have argued that a full understanding of classical conditioning must involve closer attention to the subtle timing interactions that occur during and between trials (Gallistel & Gibbon, 2000). The need for a better understanding of the role of timing in learning is one of the challenges at the forefront of current learning research.

Associative Bias and Ecological Constraints

The formal learning models described above imply that any arbitrary cue (such as a tone or a light) can be associated with any outcome, be it a shock or food. But is that really true? Consider Mandy, who came down with a strange illness soon after eating escargot (snails) at dinner. The same evening that she ate the snails, she also went to the movies with her date and saw a romantic comedy. Later that night, she woke up with a fever and a terrible case of hives. Both eating the escargot and watching the romantic comedy were events that preceded the illness. But whereas Mandy hasn't been able to eat escargot since that evening, she has not stopped going to see romantic comedies. What this suggests is that not all cues are equally likely to be associated with every outcome. Rather, it appears that there is an associative bias whereby some cues (such as food) are more likely to be associated with some outcomes (such as illness).

conditioned taste aversion. A conditioning preparation in which a subject learns to avoid a taste that has been paired with an aversive outcome, usually nausea.

If she gets sick next morning, will she blame the escargot or the movie she saw? What does that illustrate?

This was strikingly demonstrated in a study of **conditioned taste aversion**, conditioning in which subjects learn to avoid specific tastes (Garcia & Koelling, 1966). John Garcia and R. A. Koelling trained rats with compound stimuli consisting of an unfamiliar taste and an unfamiliar tone (a rat's version of watching a romantic comedy while eating snails). One group of rats were then injected with a poison that made them ill. A second group of rats were given an electric shock instead (Table 4.8). Which cue would the rats in each group "blame" for their illness or shock, the taste or the tone stimulus? To see which cues were most readily associated with which outcomes, the experimenters subsequently tested the rats with each of the cues independently: on some test trials the rats were given food with the same novel taste but no tone, while on other test trials, the rats were presented with the tone but no food.

Mark Gluck

What the researchers found was that the rats in the poison group were far more likely to associate the taste stimulus with their illness than to associate the tone with their illness (much as Mandy would be more likely to blame snails rather than romantic comedies for her illness). In contrast, the rats in the shock group were more fearful in the presence of the tone stimulus than when they encountered the taste stimulus. Garcia and his colleagues concluded that taste is a more effective

Table 4.8 The Garcia–Koelling taste-aversion study

Group	Phase 1	Phase 2
Poison group	Tone + taste → poisoning	Tone → ?
Shock group	Tone + taste → shock	Taste → ?

stimulus for learning to predict illness but that an audio cue is more effective for learning to predict a shock. Clearly, rats, like people, have prior biases about what should predict what. This isn't to say that you couldn't be trained to throw up at romantic comedy movies, but it would be much harder (and require more training) than Mandy learning to avoid the taste of escargot.

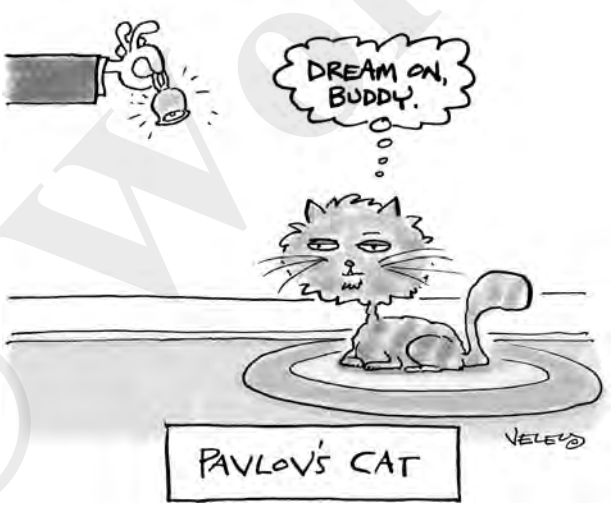
Remember the quail that were trained to associate a light with sex? Although the quail were able to learn this association following many trials of training, Domjan and colleagues found that quail could be conditioned much faster and more robustly if the CS, rather than being an arbitrary cue like a light, were something that is naturally associated in the wild with available females. These cues included the sight of a female at a distance or the sight of a female's head when the rest of her body is hidden in the underbrush (Cusato & Domjan, 1998).

Why are both Mandy and Garcia's rats more likely to associate food, rather than other cues, with getting sick? The answer may have to do with the potential *causal* relationship between eating food and getting sick that is a very real part of a person's or other animal's natural environment. In contrast, there is unlikely to be a natural causal relationship between watching a movie (good or bad) or hearing a tone and getting sick. Perhaps a sensitivity to the likely causal relationships is what guides and biases associative learning in animals. The best predictors of future events are the causes of those events, or at least their detectable indicators (Dickinson, 1980). Thus, it would make evolutionary sense for humans and other animals to be biased toward learning associations that correspond to causal relationships in the ecological niches in which the animals live and evolve.

Interim Summary

- Classical conditioning involves learning about the predictive nature of stimuli in the environment, that is, what cues predict desirable or undesirable events. If an unconditioned stimulus, US, is repeatedly and reliably preceded by a neutral stimulus, such as a bell, that neutral stimulus can become a conditioned stimulus, or CS, that evokes an anticipatory response, called the conditioned response, or CR. Conditioned responses are anticipatory responses that prepare an animal for the expected US.
  - The pairing of a potential CS with a US is not sufficient for conditioning to occur. Rather, for a CS to become associated with a US, it must provide valuable new information that helps an animal predict the future. Even if a given cue is predictive of a US, it may not become associated with that US if its usefulness has been preempted ("blocked") by a co-occurring cue that has a longer history of predicting the US.
  - Rescorla and Wagner (1972) argue that learning should occur in proportion to the degree to which the US is unexpected when it is experienced. A key assumption in the Rescorla–Wagner model is that when there are multiple CS cues present, the expectation (or prediction) of the US is calculated as the sum of the association weights of all of the cues present on that trial.

Although all animals show classical conditioning, this cartoon reminds us that not all animals will condition equally well to the same CSs and USs. If Pavlov had used cats rather than dogs, what CSs and US might he have used instead of bells and meat?



Bradford Veley/www.CartoonStock.com

- CS-modulation theories of learning (like Mackintosh's model) presume that limits in attentional capacity cause attention to one stimulus to decrease our ability to attend to other stimuli. In contrast, the Rescorla–Wagner model is a US-modulation theory of learning because it describes the learning of associations as depending on how accurately the US is predicted based on all available information. Current behavioral and biological studies of conditioning now suggest that both CS-modulation and US-modulation mechanisms are likely to be involved in learning.
- Taste is more effective than an audiovisual stimulus for learning to predict illness, while an audiovisual cue is more effective for learning to predict a shock. One interpretation of this difference is the potential causal relationship between eating food and getting sick that is part of the animal's natural ecological environment.

4.2 Brain Substrates

Pavlov was a physiologist. When he discovered associative learning in his dogs in the early 1900s, he was naturally interested in understanding the brain mechanisms responsible for it. He even conducted a few experiments examining how cortical lesions affect conditioning. However, at the beginning of the last century, the technology for observing the brain's inner workings was not highly developed. Only in recent years have scientists gained knowledge and techniques that allow detailed study of the neural circuits for conditioning. We review here two neural systems, one in mammals and the other in invertebrates, that illustrate how studies of the neural bases of conditioning have yielded insights into the circuits, cells, molecules, and genes controlling the formation of new memories.

Mammalian Conditioning of Motor Reflexes

As you saw in Figure 2.4, the **cerebellum** sits just behind and slightly below the rest of the brain and looks like a miniature brain itself. In fact, the name *cerebellum* is Latin for "little brain."

In the early 1980s, Richard Thompson and his coworkers made a startling discovery: small lesions in the cerebellum of rabbits permanently prevented the acquisition of new classically conditioned eyeblink responses and abolished retention of previously learned responses (Thompson, 1986). Thompson and his colleagues have studied the cerebellum and its role in motor-reflex conditioning for more than 25 years. Their work provides an instructive example of how support for a theory can be strengthened by converging evidence from a variety of scientific methods, such as electrophysiological recordings, brain stimulation, experimental lesions, temporary inactivation of brain structures, and genetically mutated animals (Thompson & Steinmetz, 2009).

The cerebellum has two main regions, as diagrammed in Figure 4.18. Lying along its top surface is the *cerebellar cortex*, which contains certain large, drop-shaped, densely branching neurons called **Purkinje cells**. Beneath the cerebellar cortex lies a collection of cells called the *cerebellar deep nuclei*, one of which is the **interpositus nucleus**. There are two major sensory-input pathways to the cerebellum: the CS input pathway and the US input pathway. The CS input pathway is shown in purple in Figure 4.18. (Not all the cells in the cerebellum are shown here, only the cells and pathways critical for understanding the cerebellar circuits for motor-reflex conditioning.) CS pathways from elsewhere in

Purkinje cell. A type of large, drop-shaped, and densely branching neuron in the cerebellar cortex.

interpositus nucleus. One of the cerebellar deep nuclei.

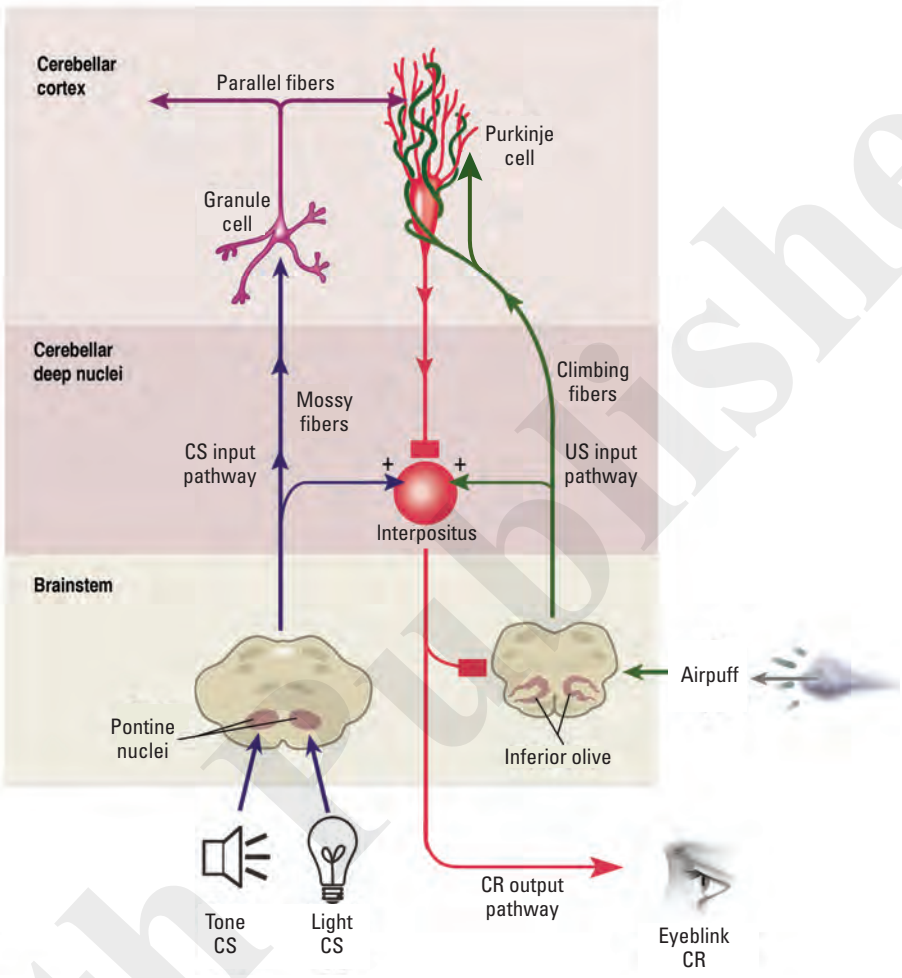


Figure 4.18 Cerebellar circuits for motor-reflex conditioning in mammals A schematic diagram of the cerebellar circuits for conditioning. The CS input pathway is purple, the CR output pathway is red, and the US input pathway is green. Excitatory synapses are shown as arrows, and inhibitory synapses terminate with a rectangle.

the brain project first to an area in the brainstem called the pontine nuclei. The pontine nuclei have different subregions for each kind of sensory stimulation. Thus, a tone CS would travel to one area of the pontine nuclei and a light CS to another. This CS information then travels up to the deep nuclei of the cerebellum along axon tracts called the mossy fibers, which branch in two directions. One branch makes contact with the interpositus nucleus. The other branch projects up toward the cerebellar cortex (by way of the granule cells and other cells not shown) and across the parallel fibers, and connects to the dendrites of the Purkinje cells.

The second sensory-input pathway, shown in green, is the US pathway. An airpuff US to the eye activates neurons in the **inferior olive**—a structure in the lower part of the brainstem—which in turn activates the interpositus nucleus. In addition, a second branch of the pathway from the inferior olive projects up to the cerebellar cortex by means of the climbing fibers (Figure 4.18). Each climbing fiber extends to and wraps around a Purkinje cell. The climbing fibers have a very strong excitatory effect on the Purkinje cells, indicated in Figure 4.18 by the large arrowhead at this synaptic junction.

inferior olive. A nucleus of cells with connections to the thalamus, cerebellum, and spinal cord.

Complementing these two converging input pathways is a single output pathway for the CR, shown in red, which starts from the Purkinje cells. The Purkinje cells project down from the cerebellar cortex into the deep nuclei, where they

form an inhibitory synapse (shown as a red rectangle) with the interpositus nucleus. To produce an eyeblink response, output from the interpositus nucleus travels (via several other intermediary cells) to the muscles in the eye to generate the eyeblink CR. You may notice that Figure 4.18 also includes an inhibitory pathway from the interpositus to the inferior olive, but we will postpone discussion of this pathway until later in the chapter. The unconditioned response (UR) pathway is not shown in Figure 4.18 because that is an innate response; it is not learned and does not originate in, or require, the cerebellum. Instead, it is a reflex circuit, similar in principle to the spinal reflexes you read about in Chapter 2.

The most important thing to note about this circuit (as diagrammed in Figure 4.18) is that there are two sites in the cerebellum where CS and US information converge and, thus, where information about the CS–US association might be stored: (1) the Purkinje cells in the cerebellar cortex and (2) the interpositus nucleus. These two sites of convergence are intimately interconnected in the output pathway: the Purkinje cells project down to the interpositus nucleus with strong inhibitory synapses.

Electrophysiological Recording in the Cerebellum

When an electrode is inserted into the interpositus nucleus (one of the two sites where CS and US information converge and the final exit point of CR information from the cerebellum), the recordings of spiking neurons during conditioned eyeblink responses display a pattern that corresponds very closely to the pattern of the eyeblinks themselves, as seen in Figure 4.19a, taken from a rabbit after one day of tone CS–US training (McCormick & Thompson, 1984). The main difference between the two patterns is that the neural activity occurs just a few milliseconds before the actual behavior. The upper blue line shows the eyeblink behavior (the extent of eyelid closure over time), while the lower graph shows

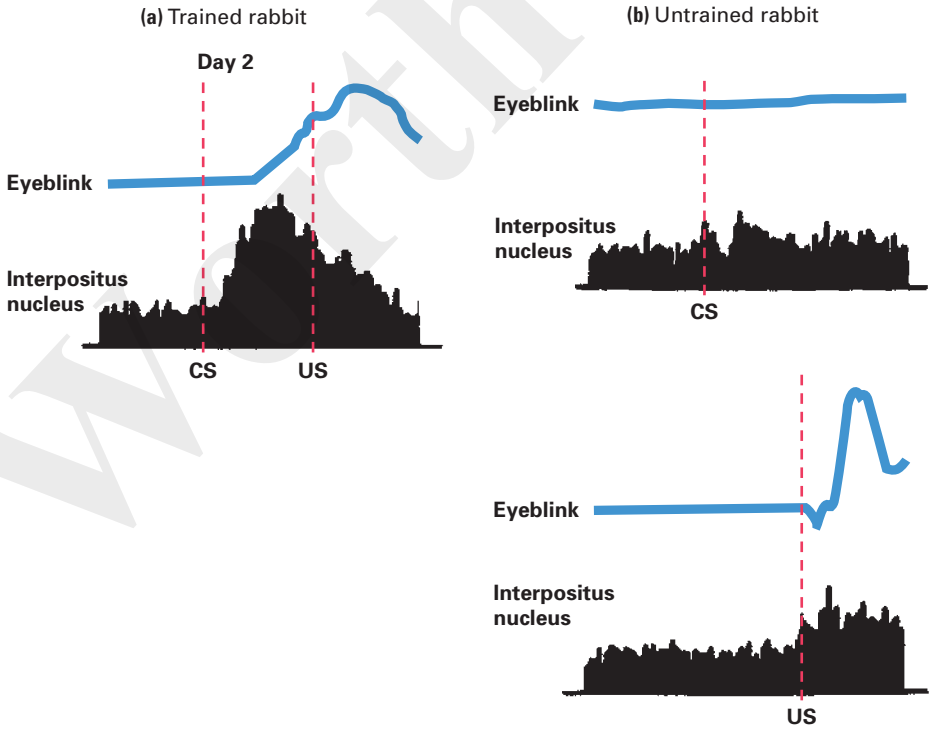


Figure 4.19 Electrophysiological recordings in the rabbit cerebellum during classical conditioning (a) Response of a trained rabbit to the CS. (b) Response of an untrained, naive rabbit to the CS alone (top) and to the US alone (bottom). The blue lines show the eyeblink behavior (the extent of eyelid closure over time), while the graphs below them show the frequency of neuronal firing in the interpositus nucleus.

Data from McCormick and Thompson, 1984.

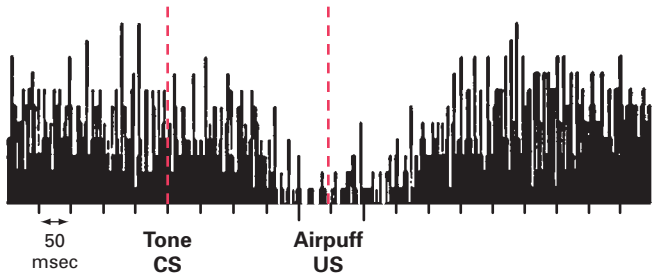


Figure 4.20 Purkinje cell activity in a well-trained rabbit The Purkinje cell's normal high rate of firing is halted in response to the CS and resumes after the US has occurred.
Data from R. F. Thompson.

the frequency of neuron firing in the interpositus nucleus, averaged over several rabbits and several trials.

Researchers have also recorded unpaired CS- or US-alone trials in naive rabbits. In both cases, where there is no CR (eyeblink), there is no activity in the interpositus nucleus, as seen in Figure 4.19b. The lack of substantial interpositus activity in a US-alone trial (despite a strong eyeblink UR) confirms that the cerebellum is responsible for conditioned eyeblink CRs only and not for the unconditioned eyeblink URs.

Figure 4.20 shows the firing rates recorded for a single Purkinje cell in a well-trained rabbit, with the time of the CS onset and the US indicated below. Purkinje cells spontaneously fire all the time, even when nothing is happening. However, in a well-trained animal, many of these cells *decrease* their firing in response to the tone CS, as shown in Figure 4.20. Why would the Purkinje cells turn off in response to a CS? Looking back at the diagram of cerebellar circuitry in Figure 4.18, note that Purkinje cells *inhibit* the interpositus nucleus, the major output pathway driving the conditioned motor response. Shutting off the Purkinje cells removes inhibition from the interpositus, freeing the interpositus to fire (as in Figure 4.19a).

Brain Stimulation as a Substitute for Behavioral Training

What if we knew exactly which pathways in your brain would change as a result of reading the words on this page? If so, we might be able to put electrodes in your brain and electrically stimulate those pathways in just the right pattern, at just the right time, to mimic the effect of reading this text. If that were possible, you wouldn't have to bother reading this book any further or studying for the final exam. Instead, you could stimulate a few neural pathways, create a little synaptic change, and then take the final exam and score an A+, even if you had never opened the textbook or sat through your professor's lectures! Science fiction, right? Unfortunately, it is still a fantasy because we don't yet know exactly where or in what way complex learning is stored in the brain. However, for simpler forms of learning, like eyeblink conditioning, this scenario is not only possible, it's been done.

Through electrical brain stimulation of the CS and US pathways shown in Figure 4.18, an experimenter can create conditioned eyeblink responses in the rabbit that are indistinguishable from those arising from behavioral training.

Recall that different parts of the pontine nuclei respond to different kinds of sensory input, such as auditory tones or visual signals, as illustrated in Figure 4.18. It is even possible to find a specific region in the pontine nuclei that responds to a *particular* tone. As a result, it is possible to condition rabbits merely by pairing electrical stimulation of the pontine nuclei (CS) with electrical stimulation of the inferior olive (US), that is, without presenting any external stimuli (airpuff or tone). After training with this type of brain stimulation, rabbits give precisely timed, reliable eyeblink responses the very first time they hear an actual tone corresponding to the pontine nuclear region that was stimulated, just as if they had been trained all along with tones and airpuffs (Steinmetz et al., 1989).

In these studies, direct stimulation of the inferior olive causes the rabbit to blink and can be substituted for an airpuff US, as shown in Figure 4.21. Similar conditioning over 4 days of training is seen whether an airpuff US (dashed line)

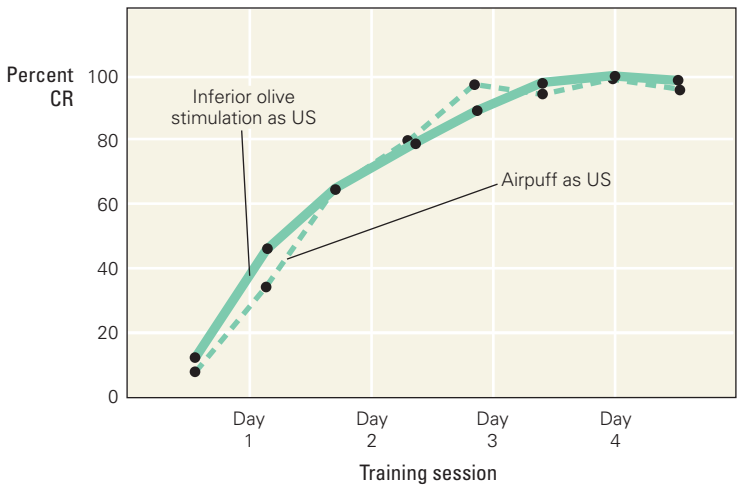


Figure 4.21 Substituting stimulation of the inferior olive for a US Four days of training using stimulation of the inferior olive as the US (solid line) produces the same amount of conditioned eyeblink response as four days of training with an airpuff US (dotted line).
Data from Steinmetz et al., 1989.

or a stimulation of the inferior olive (solid line) is used (Steinmetz, Lavond, & Thompson, 1989).

Thus, rabbits that have had their inferior olives and pontine nuclei electrically stimulated will "pass the eyeblink test" much as if they had gone through days of tone–airpuff training. Like the science fiction fantasy alluded to earlier, stimulating the correct pathways creates learning that seems indistinguishable from conditioning in a rabbit that has gone through the usual training with tones and airpuffs.

Impaired Conditioning Following Cerebellar Damage

Another experimental approach for investigating the neural bases of classical conditioning is to introduce brain lesions—that is, to selectively remove small areas of the brain—and observe the consequences. Recall that the interpositus nucleus (see Figure 4.18) projects information about the CR out of the cerebellum. Thus, without the interpositus nucleus, you would expect that there could be no CR. This is exactly what Thompson and colleagues found: removing even 1 cubic millimeter of tissue from the interpositus nucleus completely and permanently abolished all previously learned conditioned responses and prevented all future eyeblink learning.

In contrast to lesions of the interpositus, which totally abolish learned eyeblink CRs, lesions of the cerebellar cortex (including the Purkinje cells) disrupt, but do not eliminate, eyeblink conditioning. Animals with lesions of the cerebellar cortex show small, poorly timed conditioned CRs (Perret, Ruiz, & Mauk, 1993). Recently, researchers have developed mutant mice with a genetic variation that causes selective degeneration of Purkinje cells. These mutant mice are slow at learning eyeblink conditioning, much like animals that have their cerebellar cortex physically removed (Chen, Bao, Lockard, Kim, & Thompson, 1996). Together, these lesion and mutant studies provide strong converging evidence that the interpositus nucleus is involved in the formation and execution of the conditioned response, while the cerebellar cortex is involved in response timing.

Given the critical role of the cerebellum in motor-reflex conditioning, it is not surprising that patients with cerebellar damage display significant deficits in acquiring the eyeblink conditioning. Such patients are slower to learn the CR and show low overall frequency and abnormal timing of CRs (Daum et al., 1993). Interestingly, patients who have undergone surgery that spares the

deep nuclei are able to acquire a little conditioning, while patients with more extensive cerebellar damage show no conditioning at all. It is important to note that cerebellar damage does not impair all forms of associative learning. For example, cerebellar patients perform within the normal range on learning verbal associations, such as matching names with faces, which suggests that other areas of the brain play a role in these more abstract tasks (Daum et al., 1993). There is also a clear lateralization of cerebellar involvement in eyeblink conditioning: damage to the left cerebellum interferes only with conditioning to the left eye, while damage to the right cerebellum interferes only with conditioning to the right eye; this is true in both rabbits and humans (Thompson & Krupa, 1994; Woodruff-Pak & Lemieux, 2001).

Genetics offers additional insights into human eyeblink conditioning. Irene Daum and colleagues have studied several groups of patients in whom chromosomal irregularities cause abnormalities and degeneration in either the cortical Purkinje cells or the deep nuclei (Daum et al., 1993). They found that patients with genetic abnormalities of the deep nuclei are severely impaired at acquiring the eyeblink CRs, while those with abnormalities in the Purkinje cells show more mixed results. These genetic studies provide additional evidence that the deep cerebellar nuclei are essential for learning the CR, while the Purkinje cells in the cerebellar cortex exert some modulating but nonessential influence on this learning.

Error Correction through Inhibitory Feedback

As described in Chapter 2, long-term potentiation (LTP) of a synapse occurs when simultaneous activity in two adjoining neurons leads to a strengthening of the connecting synapse. LTP is a mechanism for synaptic change that occurs whenever two adjoining neurons fire at the same time and is thus much simpler than the error-correcting rule of the Rescorla–Wagner model, in which associative changes depend on many inputs (such as all the CSs present on a trial). Given its complexity, the Rescorla–Wagner model of learning probably does not describe what takes place in a learning brain at the cellular level, but the error-correction mechanisms the model predicts do appear to emerge from brain circuits.

If you look again at the cerebellar network in Figure 4.18, you will see an additional pathway within the cerebellum we have not yet discussed. This inhibitory feedback pathway projects from the interpositus nucleus to the inferior olive. In a well-trained animal, the production of a CR, through activation of the interpositus nucleus, will in turn inhibit the inferior olive from sending US information to the Purkinje cells in the cerebellar cortex (Sears & Steinmetz, 1991). This means that activity in the inferior olive will reflect the actual US minus (due to inhibition) the expected US, where the expected US is measured by the interpositus activity that drives the CR. Actual US minus expected US: sound familiar? It should. This is the same difference (actual US minus expected US) that the Rescorla–Wagner model uses to calculate the prediction error on a trial, which is then used to determine how much weight should accrue to the CS association.

If the inferior olive is where the brain codes the prediction error during conditioning, then we should be able to predict changes in the firing of the inferior olive based on the Rescorla–Wagner model (Gluck, Reifsnider, & Thompson, 1990; Gluck, Allen, Myers, & Thompson, 2001). During CS–US acquisition training, the prediction error diminishes on each successive learning trial. Thus, we should expect to see inferior olive activity in response to the US diminish the more the US is predicted by the trained CS. Eventually, when the CR is well learned, there should be very little activity in the inferior olive (that is, when

error in the Rescorla–Wagner model is close to zero). What happens matches the predictions exactly: inferior olive activity starts off high early in training and then gradually diminishes as the conditioned response is acquired (Sears & Steinmetz, 1991).

This interpretation of how the cerebellar circuits compute the changes in association weight called for in the Rescorla–Wagner model implies that Kamin's blocking effect (the clearest experimental evidence for error-correction learning) should depend on the inhibitory pathway from the interpositus to the inferior olive. This prediction was confirmed in a study by Thompson and colleagues. The researchers first trained rabbits to give reliable eyeblink responses to a tone CS and then injected a drug into the interpositus that temporarily disabled the inhibitory connection from the interpositus to the inferior olive. With this pathway disabled, they predicted, the inferior olive's activity would reflect the presence of the actual US and no longer the expected US.

The rabbits were then given phase 2 blocking training, in which a compound tone-and-light CS was paired with the US. The rabbits showed high inferior olive activity whenever the US was presented, whether or not a conditioned response was generated. As a result, in phase 3, the rabbits gave a strong response to the light CS. In other words, by disabling that one inhibitory pathway which is essential for the actual US minus expected US computation, Thompson and colleagues were able to "*block* blocking" (Kim, Krupa, & Thompson, 1998). These and related results suggest that the cerebellar–inferior olive circuit plays a role in the execution of Rescorla and Wagner's error-correction rule.

Test Your Knowledge

The Cerebellum in Motor Reflex Conditioning

1. What is the role of the Purkinje cells in the cerebellar cortex? Discuss the evidence that suggests this.
2. What are the two main cerebellar regions and the major sensory-input pathways to the cerebellum? Where do these two pathways in the cerebellum converge?
3. How do electrophysiological recordings in the rabbit cerebellum during classical conditioning demonstrate that the cerebellum is responsible for conditioned responses and not for unconditioned responses?

(Answers appear in the back of the book.)

The Hippocampus in CS Modulation

Error correction as explained by the Rescorla–Wagner model is only one mechanism at work in classical conditioning. Another, CS modulation, was suggested, as noted above, by the theories of Mackintosh and of Pearce and Hall. Here we briefly discuss some of the brain systems that appear to govern these mechanisms for modulating the processing of CS cues.

As you learned in Chapter 2, the hippocampus is a string-bean–shaped structure that lies, in humans, just inward from the ears. Figure 4.22 shows the hippocampus in various species.

The hippocampus is not necessary for learning new conditioned responses. For example, animals or humans with hippocampal damage are able to learn a basic conditioned eyeblink response quite normally. Nevertheless, electrophysiological recordings of animals show that the hippocampus is very active during conditioning, especially early in training. What role does the hippocampus

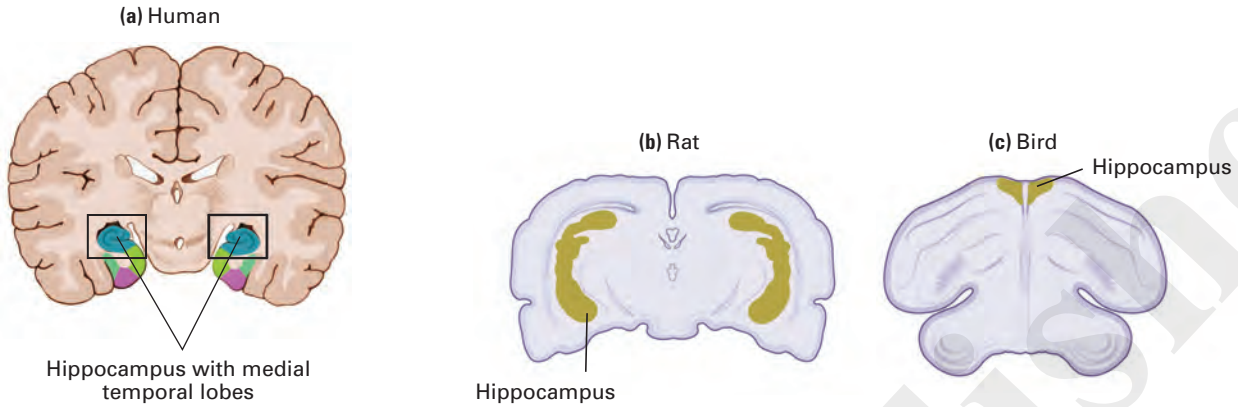


Figure 4.22 The hippocampus in different species of animals, including humans The medial (inner) part of the temporal lobes contains the hippocampus, the amygdala, and several nearby cortical areas.

play in conditioning? One way to find possible clues to its role is to look at more complex conditioning paradigms, such as latent inhibition (described in Table 4.7). As you learned in Section 4.1, latent inhibition is demonstrated when, before training, an organism is exposed to a cue unassociated with a US; later, during conditioning, the organism is then slow to learn that the cue does predict a US.

As you also learned, the Rescorla–Wagner model is *not* able to explain the phenomenon of latent inhibition. If the Rescorla–Wagner model's error-correction process cannot explain latent inhibition and if the cerebellum implements the error-correction principle, then perhaps other brain regions involved in classical conditioning besides the cerebellum are responsible for latent inhibition. Might the hippocampus be such a region? If so, then the animal learning theories that capture behavioral phenomena other than error-correction learning might provide us with some ideas of what the hippocampus may do during classical conditioning.

The CS modulation theories of Mackintosh and of Pearce and Hall, discussed earlier in this chapter, suggest that to find the system responsible for latent inhibition and related phenomena, we should look for a system involved in determining the salience of sensory cues. If the hippocampus is needed for CS modulation effects in classical conditioning, then an animal *without* a hippocampus should *not* exhibit CS modulation effects such as latent inhibition. In fact, this is exactly what researchers have found: removing the hippocampus (and associated cortical input regions) eliminates the latent inhibition effect in classical conditioning of the rabbit eyeblink reflex (Solomon & Moore, 1975; Shohamy, Allen, & Gluck, 2000).

Many other behavioral phenomena that cannot be explained by the Rescorla–Wagner model are also found to disappear in animals that have lesions to the hippocampus and surrounding brain regions. This suggests that the Rescorla–Wagner model may be better described as a model of the cerebellar contributions to motor-reflex conditioning in hippocampal-lesioned animals than as a model of conditioning in healthy, intact animals. That is to say, the model applies best to the brain regions responsible for error-correction learning, such as the cerebellum, but does not explain the additional contributions of the hippocampus.

What functional role, then, does the hippocampus play in classical conditioning of motor reflexes such as the eyeblink response? If the hippocampus is necessary for latent inhibition and other forms of CS modulation, we might infer that the hippocampus plays a role in determining how sensory cues are processed before they are used by the cerebellum to form long-term memory traces.

Further discussion of the role of the hippocampus in processing sensory relationships while remembering new facts and events will be discussed in Chapter 6. Later, in Chapter 9, we describe some specific theories about how the hippocampus modulates sensory processing in various forms of learning and memory.

Invertebrates and the Cellular Basis of Learning

Chapter 3 introduced you to the sea snail *Aplysia* and studies by Eric Kandel and colleagues on the neural substrates of two forms of non-associative learning: habituation and sensitization. To briefly recap, habituation occurs when *Aplysia's* siphon (see Figure 3.9) is repeatedly but lightly touched. Initially this results in a gill-withdrawal reflex. However, each subsequent stimulation of the siphon elicits a progressively smaller response. The circuit for this learned response includes a sensory neuron (activated by touching the siphon) that makes an excitatory synapse with a motor neuron that controls the gill withdrawal (see Figure 3.10). The neural mechanism for habituation is thought to be a progressive decrease in the number of neurotransmitter (in this case, glutamate) vesicles available in the sensory neuron's axon for each successive stimulation of the siphon. In contrast, sensitization is a global increase in responding to all or most stimuli following an unpleasant stimulus, such as an electric shock to *Aplysia's* tail. The tail shock activates modulatory interneurons that release serotonin onto the axon terminals of all the sensory neurons that project to the gill-withdrawal motor neuron. Serotonin increases the number of glutamate vesicles released when the sensory neuron is stimulated. This results in the generalized (non–stimulus-specific) increase in gill withdrawal elicited by all future stimuli, including touches on either the siphon or the mantle. The top two entries in Table 4.9 summarize the key differences between these two forms of non-associative learning.

What do you think would happen if both kinds of stimuli—touching the siphon and shocking the tail—were repeatedly paired? Tom Carew, in collaboration with Kandel and other colleagues, showed that *Aplysia's* siphon-withdrawal reflex can be classically conditioned, as illustrated in Figure 4.23a. When touching the siphon (a potential CS) is repeatedly paired with shocking the tail (the US), an enhanced siphon withdrawal (CR) results in response to subsequent touches of the siphon (Carew, Hawkins, & Kandel, 1983). The enhanced siphon-withdrawal response to the siphon-touch CS following paired training is considerably greater than the generalized sensitization that occurs from presentation of the tail shock alone. Moreover, this classically conditioned

Table 4.9 Varieties of learning in *Aplysia*

Type of learning	Associative	Stimulus specific	Mechanism(s)	Locus of effect
Habituation	No	Yes	Decrease in glutamate	Cellular process
Sensitization	No	No	Serotonin-induced increase in glutamate	Cellular process
Classical conditioning	Yes	Yes	1. Presynaptic activity–dependent enhancement of glutamate release from sensory neuron	Cellular process
			2. Postsynaptic change in receptors of motor neuron	Structural change
			3. A cascade of intracellular molecular events that activate genes in the neuron's nucleus, causing an increase in the number of sensory-motor synapses	Structural change

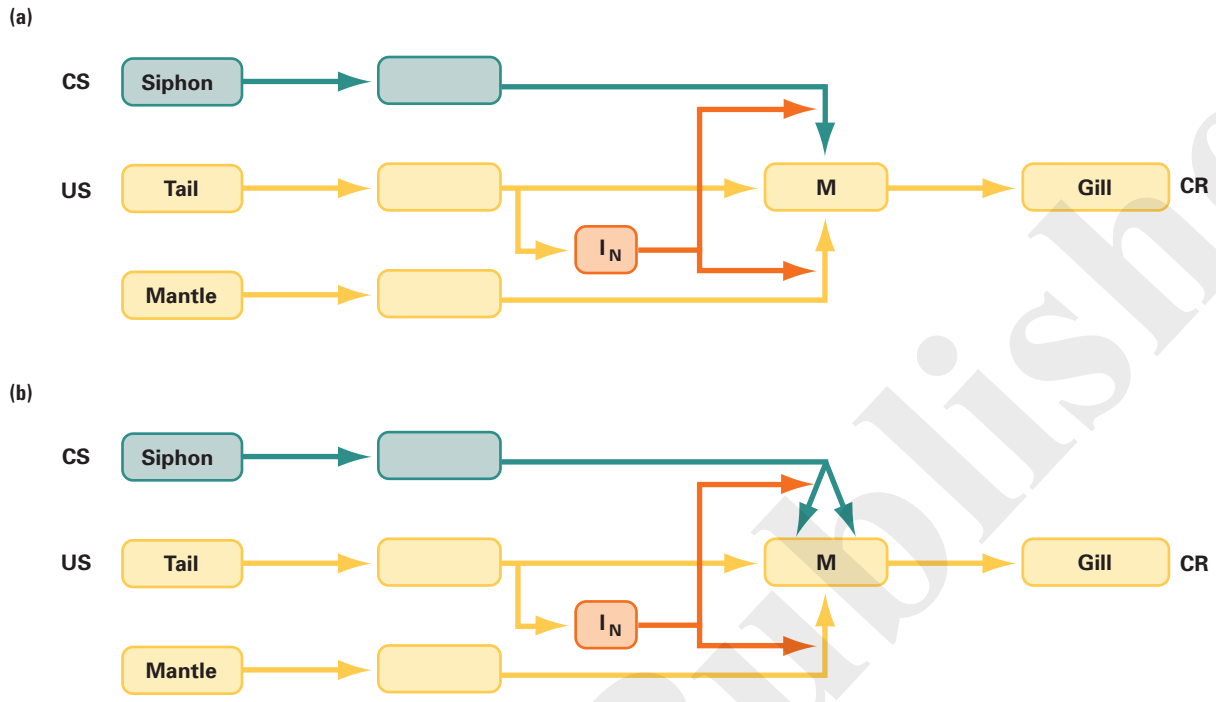


Figure 4.23 Classical conditioning in *Aplysia* (a) As with habituation and sensitization, classical conditioning (a CR) in *Aplysia* results when three sensory pathways—the siphon (CS), the tail (US), and the mantle—converge on the gill-withdrawal motor neuron (M). The tail pathway includes a secondary pathway through an interneuron (I_N), and this releases serotonin onto the other sensory synapses when the tail is shocked. (b) Long-lasting forms of classical conditioning require the formation of new synapses (the dark green arrow heads) between the sensory neurons of the siphon and the motor neuron. The new synapses are created through a molecular cascade set in motion by the serotonin released by the interneuron.

siphon-withdrawal CR is also specific to the siphon and does not generalize to other stimuli, such as a touch on the mantle.

What happens inside the nervous system of *Aplysia* when these two stimuli are paired? Kandel and colleagues demonstrated that paired training produces an increase in the glutamate vesicles that are released in the siphon's synapse on the motor neuron, much like an exaggerated form of the mechanism for sensitization described in Chapter 3 (Hawkins, Abrams, Carew, & Kandel, 1983). This implies that a cellular mechanism for classical conditioning can be understood as an elaboration of the same cellular mechanism used for sensitization.

activity-dependent enhancement. Paired training of CS and US that produces an increase in the glutamate vesicles released from sensory to motor neurons.

The pairing-specific enhancement of glutamate release in the sensory neuron synapse is called an **activity-dependent enhancement** because it depends on activation of the sensory neuron prior to the administration of the US. Earlier in this chapter, we discussed how classical conditioning of the rabbit eyeblink response is sensitive to the order and timing of the tone CS and the airpuff US. The same holds true for conditioning of *Aplysia's* siphon withdrawal: conditioning occurs only if the siphon-touch CS is presented about half a second before the tail-shock US. If the US occurs much later (more than 2 seconds after the CS) or before the CS, nothing other than nonspecific sensitization will occur. Thus, after sensory stimulation, whatever process occurs within the neuron to prime the neuron for an increase in glutamate release has a time course of about a half a second.

To summarize, Kandel and colleagues demonstrated that activation of *Aplysia's* sensory neuron has at least three consequences. First, it causes the motor neuron to fire, by the release of the neurotransmitter glutamate into

the synapse. Second, it causes a decrease in glutamate vesicles available for any subsequent stimulation of the sensory neuron, resulting in habituation. Third, it primes the synapse, through a series of intracellular events lasting about half a second, so that a subsequent presentation of the neurotransmitter serotonin (released following activation of an aversive tail shock) creates an increase in future glutamate release—resulting in a classically conditioned increase in gill withdrawal following pairing of the sensory stimulus (the CS) and the tail shock (the US).

Presynaptic versus Postsynaptic Changes during Learning

This activity-dependent enhancement of the sensory neuron's release of glutamate onto the motor neuron is a presynaptic form of synaptic plasticity, because like the mechanism for sensitization discussed in Chapter 3, it involves a change in the sensory neuron. However, the story is actually more complicated. Later studies demonstrated that there is also a postsynaptic mechanism for conditioning that involves changes in neurotransmitter receptors on the motor neuron (Bao, Kandel, & Hawkins, 1998). Thus, the mechanisms for classical conditioning in *Aplysia* involve both presynaptic and postsynaptic changes in the circuits connecting the CS and the CR, as summarized in Table 4.9.

One advantage of *Aplysia* as a model system for studying the intracellular molecular pathways of learning is that it is possible to identify key neurons (such as entire memory-trace circuits), remove them from the animals, and keep those neurons functioning in a culture dish. By isolating the key circuits for learning and studying them outside the animal, Kandel and colleagues were able to explore the question, What long-term changes in *Aplysia* circuitry could account for long-lasting forms of classical conditioning? The search for the answer to this question took scientists back to the very origins of who we are—our genes—and has given rise to an important new field, the molecular genetics of memory (and also won Kandel the Nobel Prize for Physiology or Medicine in 2001). As we review in more detail in Chapter 12, genes are stretches of DNA molecules (deoxyribonucleic acid), found in the nucleus of every cell, that encode information needed to produce protein molecules. Most people are aware of the role that genes play in determining how our bodies and brains develop during gestation in the uterus. However, our genes don't stop working after birth; rather, they play a critical role throughout our lives, continuing to maintain and guide further growth and development of our bodies and brains, including the changes that result in long-lasting forms of memory.

Long-Term Structural Changes and the Creation of New Synapses

Using recent advances in molecular biology techniques, Kandel and colleagues were able to show that the serotonin released by *Aplysia's* interneurons following a tail-shock US does more than cause a short-term increase in the sensory neuron's release of glutamate; it also launches a cascade of intracellular molecular events that set the stage for long-term structural changes in the neuron. Following multiple pairings of the CS and US, protein molecules in the sensory neuron's synapse travel back up the axon of the sensory neuron all the way to the cell body. There they switch on genes inside the nucleus of the neuron that in turn set in motion the growth of new synapses (Figure 4.23b).

More recent work by Kandel and others has identified two proteins that are found inside neurons and that play critical regulatory roles in this synapse-creation process. The first protein, CREB-1, activates genes in the neuron's nucleus that initiate the growth of new synapses. The second protein, CREB-2,

plays an opponent role, inhibiting the actions of CREB-1. The creation of new synapses during learning requires a cascade of processes inside the cell that activate CREB-1 and suppress CREB-2.

What do you think would happen if functioning of the CREB-1 protein was impaired? Kandel and colleagues demonstrated that if CREB-1 is rendered inactive by injection of molecules into the neuron that compete with CREB-1's ability to activate genes for new synapses, the circuits subsequently fail to show long-lasting forms of associative learning (Dash, Hochner, & Kandel, 1990). Most important, the inactivation of CREB-1 does not affect the short-lasting forms of learning that depend only on increased glutamate release. This study provided critical evidence for a dissociation between short-lasting forms of learning, which do not require the CREB-1 protein, and long-lasting forms, which do.

In a related study, Kandel and colleagues showed that removing the influence of the opponent protein, CREB-2, had the opposite effect: with the CREB-2 inactivated, long-lasting learning occurs rapidly at the sensory neurons, after even a single exposure to serotonin (Bartsch et al., 1995). The role of CREB molecules in modulating long-lasting forms of memory is not limited to *Aplysia*; increasing CREB-1 in fruit flies (*Drosophila*) allows them to learn much more rapidly than usual, while increasing their CREB-2 blocks the formation of long-term memories, such as those produced in the odor-conditioning task described earlier in this chapter (Yin et al., 1994). The CREB molecules also play a critical role in mammals' learning; studies in mice have shown that activity of CREB-1 in the hippocampus is critical to long-lasting but not short-term increases in neuron-to-neuron associations based on LTP (Bourtchuladze et al., 1994).

Studies of classical conditioning in *Aplysia* have demonstrated that anatomical changes in neural circuits, including the growth or deletion of synapses, are characteristic of long-lasting forms of memory. In contrast, short-term, labile forms of memory are associated with temporary intracellular changes within existing anatomical pathways, including shifts in the location, size, or number of neurotransmitter vesicles, which alter synaptic transmission efficacy. Thus, as was also discussed in Chapter 3, the transition from short-term to long-term learning may be characterized as a shift from transmission-process–based changes within the neuron to structural changes within the neural circuits (see Table 4.9).

Interim Summary

- There are two sites in the cerebellum where CS and US information converges and that might potentially be locations for the storage of the CS–US association: (1) the Purkinje cells in the cerebellar cortex and (2) the interpositus nucleus. The interpositus nucleus is the only output pathway from the cerebellum; it is the route through which the learned response travels to the motor systems that control behavior, such as an eyeblink CR.
- The inferior olive is believed to compute the degree to which a US is unexpected, providing the information necessary to implement Rescorla and Wagner's principle of error-correction learning in the cerebellum.
- The hippocampus is a structure underlying some of the CS-modulation effects in conditioning. This is consistent with data showing that an animal without a hippocampus does not exhibit CS-modulation effects such as latent inhibition.
- Kandel and colleagues demonstrated that activation of *Aplysia's* sensory neuron by an external stimulation (such as presentation of a stimulus cue)

primes the synapse, through a series of intracellular events lasting about half a second, so that a subsequent presentation of serotonin (released following activation of an aversive tail shock) creates an increase in future glutamate release, resulting in a classically conditioned increase in gill withdrawal.

- After multiple pairings of the CS and US in *Aplysia*, protein molecules in the sensory neuron's synapse travel back up the axon of the sensory neuron all the way to the cell body. There they activate genes inside the nucleus of the neuron that in turn set in motion the growth of new synapses.

4.3 Clinical Perspectives

In this final section of the chapter, we focus on two clinical applications of classical conditioning. The first involves recognition of the ways drug addiction and drug abuse are intimately linked to classical conditioning, the other harnesses classical conditioning to reduce the amount of medication needed for treating a chronic disease.

Classical Conditioning in Tolerance to Addictive Drugs

The role of learning and memory in drug addiction is a fascinating topic that we consider from several viewpoints in this textbook. In Chapter 5, we explore the neural mechanisms of reward that are impaired by most drugs of abuse. Chapter 8 discusses the role of the frontal lobes as the brain's executive controller, their importance in inhibiting inappropriate behaviors, and how this role is compromised in drug addicts. In the following discussion of drug tolerance, we see how the behavioral and biological mechanisms of classical conditioning influence another aspect of drug addiction and abuse.

Early in this chapter we discussed how automatic compensatory responses occur in body systems that have a mechanism for **homeostasis**, the tendency of the body (including the brain) to gravitate toward a state of equilibrium or balance. An addict's tolerance to drugs of abuse such as alcohol, cocaine, or ecstasy develops in the same way. As the addict's body adjusts to the drug effects (through expectation of the forthcoming "high"), larger and larger doses are required to produce the same high the addict experienced on first taking the drug. One way this happens is through conditioning: environmental cues that accompany drug use can classically condition the user to expect to receive the drug. In other words, the environmental cues (people, places, and so on) act like CSs associated with the drug (the US). The intense craving an addict feels in response to these cues is the CR and results from the body's conditioned compensatory response of lowering the levels of the brain chemicals enhanced by the drug in anticipation of the drug's arrival (more on the role of conditioning in drug addiction in Chapter 5).

A potential consequence of such conditioned tolerance is that victims of heroin who overdose are rarely novice users (Siegel, 2001). Rather, they tend to be long-time heroin addicts who have developed a high degree of tolerance to the drug but make the mistake of taking their usual dose in an unusual setting. For example, the situational cues that result in conditioned drug tolerance can include the room in which the drug is usually taken. You may recall reports of rock stars and others dying of heroin overdoses in hotel bathrooms, which were most likely far different from the settings in which they were used to taking their drug. What might have happened is that they overdosed on what was

Banana Stock/Getty Images

Why would an addict be more likely to die of an overdose in a hotel room rather than at home?

otherwise their normal dosage of drug because, in this novel setting, their body was not prepared for the large influx of drug that occurred.

Different methods of drug injection are another form of environmental cue that can become associated with drug expectation. One longtime heroin addict is reported to have died of an overdose when, looking for an accessible blood vein, he injected himself in his penis for the first time (Winek, Wahaba, & Rozin, 1999). Without the compensatory response to customary cues, the net effect of his usual dose was far greater than he was used to and resulted in his death (for reports of similar real-life cases, see Siegel, 2001; Siegel & Ellsworth, 1986).

An unusual taste to a beverage can serve as a novel situational cue influencing the effects of alcohol on the brain. This was demonstrated in a study in which college students showed greater cognitive and motor impairments when they consumed a given amount of alcohol in an unusual drink (in this case, a blue, peppermint-flavored beverage) than when they had the same amount of alcohol in a familiar drink, such as beer (Remington, Roberts, & Glauthier, 1977). Perhaps this is yet another reason why people get wilder at holiday parties, when they are drinking alcohol in sweet and bubbly holiday punches.

Research has demonstrated conditioned tolerance in a wide variety of animal species. For example, Shepard Siegel and colleagues examined the effect of cues when heroin is administered to rats (Siegel, Hinson, Krank, & McCully, 1982). Siegel gave three groups of rats a fairly large (for their body weight) dose of heroin. The first group of rats had previously received a lower dose of heroin, administered in the same cage and room where they were later tested with the larger dose (the "same-tested" group in Figure 4.24). The second group of rats had also previously received the lower dose of heroin, but in a different cage and room (the "different-tested" group). Finally, the "first-time tested" group of rats were receiving heroin for the first time.

As shown in Figure 4.24, Siegel and colleagues found that the large dose of heroin almost always (96% of the time) resulted in a fatal overdose in the "first-time tested" rats. In contrast, the "different-tested" group showed some evidence of tolerance; only 64% of these rats suffered a fatal overdose. But the "same-tested" rats, who were tested in the same environment in which

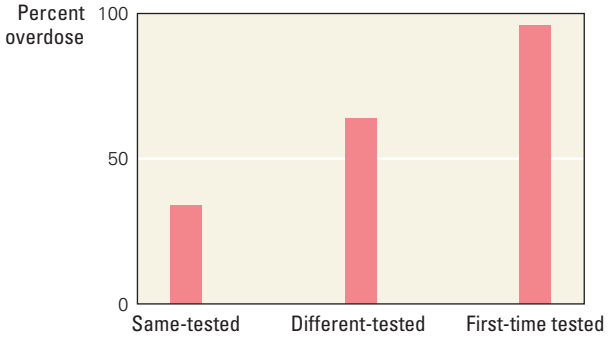


Figure 4.24 A study of drug tolerance Rats in the "first-time tested" group, which had received no prior heroin, showed the highest percentage of fatal overdose. Of those rats that had received a smaller prior dose, the ones tested in the same cage as previously ("same-tested") showed the lowest level of overdoses, while the ones tested in a different cage ("different-tested") showed an intermediate level of overdosing.
Data from Siegel et al., 1982.

they had previously been administered low doses of heroin, had a mortality rate of only 32%. Thus, these rats were protected from overdose by the conditioned tolerance that they learned during the administration of lower doses of heroin in the same setting.

If drug tolerance is a form of classical conditioning, you might expect that the same rules and principles would apply to it as to eyeblink and other forms of conditioning. This appears to be the case. Recall the paradigm of latent inhibition discussed earlier and shown in Table 4.7: pre-exposure to a CS delays subsequent learning of a CS–US association. If this latent inhibition effect applies to drug tolerance, then pre-exposing an animal to the cues that predict drug delivery should slow down development of learned tolerance to the drug. To test latent inhibition in the context of drug use, researchers have given animals an injection of an inert placebo to pre-expose them to the contextual cues (the sound and feel of getting an injection) of drug use. This pre-exposure does prove to slow down, by latent inhibition, the development of a cue–drug association where the cue is the CS and the drug is the US. In fact, this is exactly what Siegel and colleagues found in studies of morphine tolerance in rats (Siegel, 1983).

Such results provide compelling evidence for the applicability of Pavlovian analyses to learned drug tolerance. They suggest that drug tolerance and the loss of drug tolerance in novel drug-taking environments are mediated by basic processes of classical Pavlovian conditioning.

As you recall from Section 4.1, an extinction paradigm should eliminate (or at least reduce) the association between conditioned cues and conditioned response when the cues are repeatedly presented alone and not paired with the prior US. If drug craving is viewed as a conditioned response, then the same principle of extinction should also apply, and indeed it does. In carefully controlled laboratory studies, rats that became addicted to alcohol showed significant extinction of this addiction through repeated nonreinforced exposure to experimentally manipulated cues that had previously been paired with administration of alcohol (Krank & Wall, 1990). In real life, however, it is very hard to both identify and extinguish all the cues that have become associated with drug use.

Eventually, these and other studies, by deepening scientists' understanding of the role of classical conditioning in drug addiction, may provide new tools to help drug addicts overcome their addiction. For example, perhaps further research on the extinction of conditioned responses will shed light on why addicts so often relapse when they are trying to kick their drug habit. You can read more about this mechanism in "Learning and Memory in Everyday Life: Extinguishing a Drug Habit."

Reducing Medication through Classical Conditioning

Classical conditioning—specifically, the pairing of a neutral stimulus with a powerful medication—offers a way of training the body to cope with a disease without continued use of a medication that may have unfortunate side effects. In one example, an 11-year old girl suffering from the autoimmune disease lupus was able to benefit from this approach (Olness & Adler, 1992; Giang et al., 1996). The standard treatment for lupus requires suppressing a person's immune system in order to protect their tissues and organs from being targeted by their own white blood cells. The chemotherapeutic drug the girl would have had to take orally, however, had many severe and disabling side effects.

LEARNING AND MEMORY IN EVERYDAY LIFE

Extinguishing a Drug Habit

Even in a well-controlled laboratory, extinguishing a cue and keeping it extinguished turns out to be extremely difficult. Mark Bouton and colleagues, studying the extinction of conditioned responses in rats, demonstrated that manipulations such as changing the context, or just waiting a specific period, often result in the extinguished cues reacquiring their conditioned associative properties (Bouton, 2000). Is it any surprise, then, that efforts to adapt extinction methods to therapy for drug addiction have yielded only mixed results at best (Carroll, 1999; Siegel & Ramos, 2002)? The fact that the addict remains susceptible to the first drink or smoke suggests that to extinguish the drug associations, we would have to include small doses of the drugs during cue-exposure therapy to better reproduce and extinguish the cravings. Clearly, such a procedure would result in many practical and legal difficulties. Nevertheless, Bouton's work suggests three principles that can help guide anyone trying to extinguish a habit or association:

1. Since extinction effects are highly context-sensitive, cue-exposure therapy should be conducted in as many different contexts as possible, including those that are part of the patient's everyday life. This will prevent the extinction of drug craving from becoming dependent on any one context (such as a drug rehab center).
2. The extinction training should be spread out over time rather than conducted all at once, because time serves as a powerful context. A 2-week stint in a rehab clinic may not be enough to make a long-term difference. Multiple therapy sessions at different times and in different contexts are more effective.
3. Whenever possible, the cue-exposure therapy should take place in the same contexts in which the original drug habits were acquired. Thus, it is better to go through the cue-exposure therapy at home rather than in a very unfamiliar setting, such as a drug rehabilitation center.

To minimize her need to take the drug for an extended time, her doctors used classical conditioning methods in which she was initially given the drug in a liquid that tasted of cod liver oil and smelled of roses. Following that initial treatment, the doctors continued to give her the same rose-smelling cod liver oil every month but in only half of the months did they include the chemotherapeutic drug. By cutting back her total medication by half, they significantly reduced her suffering from side effects of the drug. At the same time, her body appeared to learn to suppress its own immune system in response to the compound cue of cod liver taste and rose smell, and after several years following this treatment she was still in remission from the disease.

Interim Summary

- The situational cues that result in conditioned drug tolerance can be any sensory cues associated with drug use, including the feel of the needle and the method of injection.
- Rats are protected from heroin overdose by the conditioned tolerance that they learned during the administration of lower doses in the same setting.
- Addiction can be partially reduced through Pavlovian extinction: rats who became addicted to alcohol showed significant extinction of this addiction through repeated non-reinforced exposure to experimentally manipulated cues that had previously been paired with administration of alcohol.

- Classical conditioning methods can be used to train people to suppress their own immune responses by pairing an immune-suppressing drug with a previously neutral odor and taste.

Synthesis

Classical conditioning is far more than just another behavioral process or tool for investigating brain systems: it is the mother of all memory systems. Evidence of classical "Pavlovian" conditioning surrounds us every day. Like Moira, who runs to the curb when she hears the musical jingle of an ice cream truck, and like Mandy, who has sworn off snails after a bout of illness, everyone has been conditioned by cues in the environment to predict what might follow next.

In recent years, classical conditioning experiments have moved to the forefront of research into the physiological bases of learning because of the exquisite control they afford over what stimuli are presented as cues and because of the highly refined behavioral analyses and models that have been developed as a result. Building on these analyses and models, biological research has shown how different forms of classical conditioning are mediated by different brain systems, leading to fundamental insights into the neurobiology of learning and often providing tools that help us understand various clinical brain disorders.

Thus, influenced and guided by the error-correcting model of Rescorla and Wagner, along with other elegant mathematical theories, researchers are uncovering the neural bases of conditioning in a broad range of brain systems, including the cerebellum, the amygdala (to be discussed in the context of fear conditioning in Chapter 10), and the role of dopamine in reward prediction (see Chapter 5). Links between conditioning and complex forms of cognition, such as category learning, help us see how mechanisms for learning studied in simple animal circuits can provide insights into the behavioral and neural bases of human cognition. The studies in animals also illustrate the general biological principle that evolution does not work like an engineer, creating new specialized systems for each new function. Rather, evolution works more like a tinkerer, using preexisting components, in slightly modified form, to perform new functions. The behavioral and biological processes for classical conditioning are the basic building blocks, the biological alphabet, from which more complex forms of learning emerge in all species, including humans.

KNOW YOUR KEY TERMS

activity-dependent enhancement, *p. 156*
appetitive conditioning, *p. 119*
associative weight, *p. 134*
aversive conditioning, *p. 119*
blocking, *p. 131*
classical conditioning, *p. 116*
compound conditioning, *p. 128*
conditioned response (CR), *p. 117*
conditioned stimulus (CS), *p. 117*
conditioned taste aversion, *p. 145*
CS modulation theory, *p. 142*
delay conditioning, *p. 144*
error-correction learning, *p. 133*
extinction, *p. 127*
eyeblink conditioning, *p. 121*
homeostasis, *p. 126*
inferior olive, *p. 148*
interpositus nucleus, *p. 147*
interstimulus interval (ISI), *p. 144*
latent inhibition, *p. 140*
overshadowing, *p. 129*
Pavlovian conditioning, *p. 116*
prediction error, *p. 132*
Purkinje cells, *p. 147*
tolerance, *p. 125*
trace conditioning, *p. 144*
trial-level model, *p. 144*
unconditioned response (UR), *p. 117*
unconditioned stimulus (US), *p. 117*
US modulation theory, *p. 141*

QUIZ YOURSELF

1. The relationship between a US and a UR does/does not involve learning. (p. 117)
2. A CR that precedes the US is often a _____ response. (p. 117–118)
3. In eyeblink conditioning, the blink is both a _____ and a _____, although they differ in their _____. (p. 121–122)
4. In most conditioning paradigms, extinction is faster/slower than the original acquisition of the conditioned response. (p. 127)
5. _____ conditioned responses in the body are most often the result of a biological mechanism called _____. (p. 126)
6. Evidence that extinction is more than just unlearning, comes primarily from studies that look at shifts in _____ between learning and testing. (p. 127)
7. When two cues compete to predict a US or other outcome, the one that is most strongly learned is usually the cue that is learned _____, as revealed in studies of blocking. (p. 131)
8. The principle of cue competition in learning arises in the Rescorla–Wagner model, where the association weights of two cues are _____ to generate a prediction of the US. (p. 134)
9. The Rescorla–Wagner model's account of contingency learning depends on viewing the _____ as a conditionable CS. (p. 139)
10. Latent inhibition cannot be explained by the Rescorla–Wagner model because during preexposure there is no _____. (p. 141)
11. Beneath the _____ cells of the cerebellar cortex lie the cerebellar deep nuclei, including the _____ nucleus. (p. 147)
12. CS information travels up to the deep nuclei of the cerebellum along axon tracts called the _____. (p. 148)
13. An airpuff US to the eye activates neurons in the _____, a structure in the lower part of the brainstem. (p. 148)
14. Purkinje cells inhibit/excite the interpositus nucleus, the major output pathway driving the conditioned motor response. (p. 147)
15. Animals with lesions to the cerebellum show CRs, but they are _____. (p. 151)
16. Latent inhibition and other expressions of CS modulation are impaired or eliminated by lesions to the _____. (p. 154)
17. The neural mechanism for habituation is thought to be a progressive decrease in the number of _____ neurotransmitter vesicles available in the sensory neuron's axon. (p. 156–157)
18. The _____ of the sensory neuron's release of glutamate onto the motor neuron is a presynaptic form of _____. (p. 156)
19. Two proteins found inside neurons play critical regulatory roles in the synapse-creation process. The first protein, CREB-1, activates genes in the neuron's nucleus that _____ the growth of new synapses. The second protein, CREB-2, _____ the actions of CREB-1. (p. 157–158)
20. Rats can be protected from overdose by the _____ that they learned during the administration of lower doses of heroin in the same setting. (p. 160)
21. Appealing due to its simplicity, the _____ model has proven itself to be a starting point for many promising models of learning. (p. 132)
22. The learning that takes place in order to avoid or minimize the consequences of expected aversive events is known as _____. (p. 119)
23. Rescorla demonstrated that conditioning to a tone stimulus depends not only on the frequency of tone–US pairings but also on the frequency of the US in the _____ of the tone. The results of his experiment imply that animals are sensitive to _____: the degree of correlation between a potential CS and US. (p. 139)

Answers appear in the back of the book.

CONCEPT CHECK

1. Returning to our stock analysts Doris and Herman, consider what would happen if Doris showed up every day to work but Herman only came in every now and then. On the days that Doris works alone, she does a great job of predicting the stock market. But on the days that Herman shows up, the pair of them do a lousy job. What would you think about Herman? You'd probably conclude that he is no great asset. In fact, worse than being a do-nothing, he seems to interfere with Doris's ability to perform. You might even say that Herman *inhibits* Doris's predictive value. This is similar to what occurs in conditioning procedures where a tone cue is always followed by the US except when the tone appears as part of a compound tone-and-light stimulus. On these compound trials, no US occurs. What does the Rescorla–Wagner model predict will happen to the associations during this training?
2. A recovering drug addict attends therapy sessions in which cue-exposure therapy is used. The addict is exposed to drug-related stimuli (e.g., photos of common drug-taking environments, drug paraphernalia, etc.) in the therapy center several times a week for an extended period of time. Why might this treatment fail?

Answers appear in the back of the book.