Learning

CHAPTER 6

KEY TERMS

Learning	Operant conditioning	Secondary reinforcers
Acquisition	B. F. Skinner	Generalized reinforcers
Extinction	Law of effect	Token economy
Spontaneous recovery	Edward Thorndike	Reinforcement schedules—FI, FR, VI, VR
Generalization	Instrumental learning	Continuous reinforcement
Discrimination	Skinner box	Partial-reinforcement effect
Classical conditioning	Reinforcer, reinforcement	Instinctive drift
Ivan Pavlov	Positive reinforcement	Observational learning or modeling
Unconditioned stimulus	Negative reinforcement	Albert Bandura
Unconditioned response	Punishment	Latent learning
Conditioned response	Positive punishment	Insight learning
Conditioned stimulus	Omission training	
Aversive conditioning	Shaping	
Second-order or higher-order conditioning	Chaining	
Learned taste aversion (Garcia Effect)	Primary reinforcers	

OVERVIEW

Psychologists differentiate between many different types of learning, a number of which we will discuss in this chapter. Learning is commonly defined as a long-lasting change in behavior resulting from experience. Although learning is not the same as behavior, most psychologists accept that learning can best be measured through changes in behavior. Brief changes are not thought to be indicative of learning. Consider, for example, the effects of running a marathon. For a short time afterward, one's behavior might differ radically, but we would not want to attribute this change to the effects of learning. In addition, learning must result from experience rather than from any kind of innate or biological change. Thus, changes in one's behavior as a result of puberty or menopause are not considered due to learning.

CLASSICAL CONDITIONING

Around the turn of the twentieth century, a Russian physiologist named Ivan Pavlov inadvertently discovered a kind of learning while studying digestion in dogs. Pavlov found that the dogs learned to pair the sounds in the environment where they were fed with the food that was given to them and began to salivate simply upon hearing the sounds. As a result, Pavlov deduced the basic principle of *classical conditioning*. People and animals can learn to associate neutral stimuli (for example, sounds) with stimuli that produce reflexive, involuntary responses (for example, food) and will learn to respond similarly to the new stimulus as they did to the old one (for example, salivate).

The original stimulus that elicits a response is known as the *unconditioned stimulus* (US or UCS). The US is defined as something that elicits a natural, reflexive response. In the classic Pavlovian paradigm, the US is food. Food elicits the natural, involuntary response of salivation. This response is called the *unconditioned response* (UR or UCR). Through repeated pairings with a neutral stimulus such as a bell, animals will come to associate the two stimuli together. Ultimately, animals will salivate when hearing the bell alone. Once the bell elicits salivation, a *conditioned response* (CR), it is no longer a neutral stimulus but rather a *conditioned stimulus* (CS).

Learning has taken place once the animals respond to the CS without a presentation of the US. This learning is also called *acquisition* since the animals have acquired a new behavior. Many factors affect acquisition. For instance, up to a point, repeated pairings of CSs and USs yield stronger CRs. The order and timing of the CS and US pairings also have an impact on the strength of conditioning. Generally, the most effective method of conditioning is to present the CS first and then to introduce the US while the CS is still evident. Now return to Pavlov's dogs. Acquisition will occur fastest if the bell is rung and, while it is still ringing, the dogs are presented with food. This procedure is known as *delayed conditioning*. Less effective methods of learning include:

- *Trace conditioning*—The presentation of the CS, followed by a short break, followed by the presentation of the US.
- *Simultaneous conditioning*—CS and US are presented at the same time.
- *Backward conditioning*—US is presented first and is followed by the CS. This method is particularly ineffective.

Of course, what can be learned can be unlearned. In psychological terminology, the process of unlearning a behavior is known as *extinction*. In terms of classical

TABLE 6.1

Classical Conditioning

Before Conditioning	UCS (Food)		elicits	UCR (Salivation)
During Conditioning	UCS (Food) +	CS (Bell)	elicits	UCR (Salivation)
After Conditioning	CS (Bell)		elicits	CR (Salivation)

conditioning, extinction has taken place when the CS no longer elicits the CR. Extinction is achieved by repeatedly presenting the CS without the US, thus breaking the association between the two. If one rings the bell over and over again and never feeds the dogs, the dogs will ultimately learn not to salivate to the bell.

One fascinating and yet-to-be-adequately-explained part of this process is known as *spontaneous recovery*. Sometimes, after a conditioned response has been extinguished and no further training of the animals has taken place, the response briefly reappears upon presentation of the conditioned stimulus. This phenomenon is known as spontaneous recovery.

Often animals conditioned to respond to a certain stimulus will also respond to similar stimuli, although the response is usually smaller in magnitude. The dogs may salivate to a number of bells, not just the one with which they were trained. This tendency to respond to similar CSs is known as *generalization*. Subjects can be trained, however, to tell the difference, or *discriminate*, between various stimuli. To train the dogs to discriminate between bells, we would repeatedly pair the original bell with presentation of food, but we would intermix trials where we presented other bells that we did not pair with food.

Classical conditioning can also be used with humans. In one famous, albeit ethically questionable, study, John Watson and Rosalie Rayner conditioned a little boy named Albert to fear a white rat. Little Albert initially liked the white, fluffy rat. However, by repeatedly pairing it with a loud noise, Watson and Rayner taught Albert to cry when he saw the rat. In this example, the loud noise is the US because it elicits the involuntary, natural response of fear (UR) and, in Little Albert's case, crying. The rat is the US that becomes the CS, and the CR is crying in response to presentation of the rat alone. Albert also generalized, crying in response to a white rabbit, a man's white beard, and a variety of other white, fluffy things.

This example is an illustration of what is known as *aversive conditioning*. Whereas Pavlov's dogs were conditioned with something pleasant (food), baby Albert was conditioned to have a negative response to the white rat. Aversive conditioning has been used in a number of more socially constructive ways. For instance, to stop biting their nails, some people paint them with truly horrible-tasting materials. Nail biting therefore becomes associated with a terrible taste, and the biting should cease.

Once a CS elicits a CR, it is possible, briefly, to use that CS as a US in order to condition a response to a new stimulus. This process is known as *second-order* or *higher-order conditioning*. By using a dog and a bell as our example, after the dog salivates to the bell (first-order conditioning), the bell can be paired repeatedly with a flash of light, and the dog will salivate to the light alone (second-order conditioning), even though the light has never been paired with the food (see Table 6.2).

BIOLOGY AND CLASSICAL CONDITIONING

As is evident from its description, classical conditioning can be used only when one wants to pair an involuntary, natural response with something else. Once one has identified such a US, can a subject be taught to pair it equally easily with any CS? Not surprisingly, the answer is no. Research suggests that animals and humans are biologically prepared to make certain connections more easily than others. *Learned* *taste aversions* are a classic example of this phenomenon. If you ingest an unusual food or drink and then become nauseous, you will probably develop an aversion to the food or drink. Learned taste aversions are interesting because they can result in powerful avoidance responses on the basis of a single pairing. In addition, the two events (eating and sickness) are probably separated by at least several hours. Animals, including people, seem biologically prepared to associate strange tastes with feelings of sickness. Clearly, this response is adaptive (helpful for the survival of the species), because it helps us learn to avoid dangerous things in the future. Also interesting is how we seem to learn what, exactly, to avoid. Taste aversions most commonly occur with strong and unusual tastes. The food, the CS, must be *salient* in order for us to learn to avoid it. Salient stimuli are easily noticeable and therefore create a more powerful conditioned response. Sometimes taste aversions are acquired without good reason. If you were to eat some mozzarella sticks a few hours before falling ill with the stomach flu, you might develop an aversion to that popular American appetizer even though it had nothing to do with your sickness.

Garcia and Koelling performed a famous experiment illustrating how rats more readily learned to make certain associations than others. They used four groups of subjects in their experiment and exposed each to a particular combination of CS and US as illustrated in Table 6.3.

TABLE 6.2

First-Order and Second-Order Conditioning

First-Order Conditioning	
Training:	Presentation of bell + food = salivation
Acquisition:	Presentation of bell = salivation
Second-Order Conditioning (After First-Order Conditioning Has Occurred)	
Training:	Presentation of light + bell = salivation
Acquisition:	Presentation of light = salivation

TABLE 6.3

Garcia and Koelling's Experiment Illustrating Biological Preparedness in Classical Conditioning

CS	US	Learned Response
Loud noise	Shock	Fear
Loud noise	Radiation (nausea)	Nothing
Sweet water	Shock	Nothing
Sweet water	Radiation (nausea)	Avoid water

The rats learned to associated noise with shock and unusual-tasting water with nausea. However, they were unable to make the connection between noise and nausea and between unusual-tasting water and shock. Again, learning to link loud noise with shock (for example, thunder and lightning) and unusual-tasting water with nausea seems to be adaptive.

OPERANT CONDITIONING

Whereas classical conditioning is a type of learning based on association of stimuli, *operant conditioning* is a kind of learning based on the association of consequences with one's behaviors. Edward Thorndike was one of the first people to research this kind of learning.

Thorndike conducted a series of famous experiments using a cat in a puzzle box. The hungry cat was locked in a cage next to a dish of food. The cat had to get out of the cage in order to get the food. Thorndike found that the amount of time required for the cat to get out of the box decreased over a series of trials. This amount of time decreased gradually; the cat did not seem to understand, suddenly, how to get out of the cage. This finding led Thorndike to assert that the cat learned the new behavior without mental activity but rather simply connected a stimulus and a response.

Thorndike put forth the *law of effect* that states that if the consequences of a behavior are pleasant, the stimulus-response (S-R) connection will be strengthened and the likelihood of the behavior will increase. However, if the consequences of a behavior are unpleasant, the S-R connection will weaken and the likelihood of the behavior will decrease. He used the term *instrumental learning* to describe his work because he believed the consequence was instrumental in shaping future behaviors.

B. F. Skinner, who coined the term operant conditioning, is the best-known psychologist to research this form of learning. Skinner invented a special contraption, aptly named a *Skinner box*, to use in his research of animal learning. A Skinner box usually has a way to deliver food to an animal and a lever to press or disk to peck in order to get the food. The food is called a *reinforcer*, and the process of giving the food is called *reinforcement*. Reinforcement is defined by its consequences; anything that makes a behavior more likely to occur is a reinforcer. Two kinds of reinforcement exist. *Positive reinforcement* refers to the addition of something pleasant. *Negative reinforcement* refers to the removal of something unpleasant. For instance, if we give a rat in a Skinner box food when it presses a lever, we are using positive reinforcement. However, if we terminate a loud noise or shock in response to a press of the lever, we are using negative reinforcement. The latter example results in *escape learning*. Escape learning allows one to terminate an aversive stimulus; *avoidance learning*, on the other hand, enables one to avoid the unpleasant stimulus altogether. If Sammy creates a ruckus in the English class he hates and is asked to leave, he is evidencing escape learning. An example of avoidance learning would be if Sammy cut English class.

Affecting behavior by using unpleasant consequences is also possible. Such an approach is known as *punishment*. By definition, punishment is anything that makes a behavior less likely. The two types of punishment are known as *positive punish-*

ment (usually referred to simply as "punishment"), which is the addition of something unpleasant, and *omission training* or *negative punishment*, the removal of something pleasant. If we give a rat an electric shock every time it touches the lever, we are using punishment. If we remove the rat's food when it touches the lever, we are using omission training. Both procedures will result in the rat ceasing to touch the bar. Pretend your parents decided to use operant conditioning principles to modify your behavior. If you did something your parents liked and they wanted to increase the likelihood of your repeating the behavior, your parents could use either of the types of reinforcement described in Table 6.4. On the other hand, if you did something your parents wanted to discourage, they could use either of the types of punishment described in Table 6.5.

HINT

Students often confuse negative reinforcement and punishment. However, any type of *reinforcement* results in the behavior being more likely to be repeated. The *negative* in negative reinforcement refers to the fact that something is taken away. The *positive* in positive punishment indicates that something is added. In negative reinforcement, the removal of an aversive stimulus is what is reinforcing.

Punishment Versus Reinforcement

Obviously, the same ends can be achieved through punishment and reinforcement. If I want students to be on time to my class, I can punish them for lateness or reward them for arriving on time. Punishment is operant conditioning's version of aversive conditioning. Punishment is most effective if it is delivered immediately after the unwanted behavior and if it is harsh. However, harsh punishment may also result in unwanted consequences such as fear and anger. As a result, most psy-

TABLE 6.4

Reinforcement = A Consequence That Increases the Likelihood of a Behavior

Types	Mechanism	Examples
Positive reinforcement	Adds something pleasant	Give a present
Negative reinforcement	Removes something unpleasant	Excuse from household chores

TABLE 6.5

Punishment = A Consequence That Decreases the Likelihood of a Behavior

Types	Mechanism	Examples
Positive punishment	Adds something negative	Give extra chores
Omission training (also known as negative punishment)	Removes something pleasant	Take away cell phone

chologists recommend that certain kinds of punishment (for example, physical punishment) be used sparingly if at all.

You might wonder how the rat in the Skinner box learns to push the lever in the first place. Rather than wait for an animal to perform the desired behavior by chance, we usually try to speed up the process by using *shaping*. Shaping reinforces the steps used to reach the desired behavior. First the rat might be reinforced for going to the side of the box with the lever. Then we might reinforce the rat for touching the lever with any part of its body. By rewarding approximations of the desired behavior, we increase the likelihood that the rat will stumble upon the behavior we want.

Subjects can also be taught to perform a number of responses successively in order to get a reward. This process is known as *chaining*. One famous example of chained behavior involved a rat named Barnabus who learned to run through a veritable obstacle course in order to obtain a food reward. Whereas the goal of shaping is to mold a single behavior (for example, a bar press by a rat), the goal in chaining is to link together a number of separate behaviors into a more complex activity (for example, running an obstacle course).

The terms acquisition, extinction, spontaneous recovery, discrimination, and generalization can be used in our discussion of operant conditioning, too. By using a rat in a Skinner box as our example, *acquisition* occurs when the rat learns to press the lever to get the reward. *Extinction* occurs when the rat ceases to press the lever because the reward no longer results from this action. Note that punishing the rat for pushing the lever is not necessary to extinguish the response. Behaviors that are not reinforced will ultimately stop and are said to be on an extinction schedule. *Spontaneous recovery* would occur if, after having extinguished the bar press response and without providing any further training, the rat began to press the bar again. *Generalization* would be if the rat began to press other things in the Skinner box or the bar in other boxes. *Discrimination* would involve teaching the rat to press only a particular bar or to press the bar only under certain conditions (for example, when a tone is sounded). In the latter example, the tone is called a *discriminative stimulus*.

Not all reinforcers are food, of course. Psychologists speak of two main types of reinforcers: primary and secondary. *Primary reinforcers* are, in and of themselves, rewarding. They include things like food, water, and rest, whose natural properties are reinforcing. *Secondary reinforcers* are things we have learned to value such as praise or the chance to play a video game. Money is a special kind of secondary reinforcer, called a *generalized reinforcer*, because it can be traded for virtually anything. One practical application of generalized reinforcers is known as a *token economy*. In a token economy, every time people perform a desired behavior, they are given a token. Periodically, they are allowed to trade their tokens for any one of a variety of reinforcers. Token economies have been used in prisons, mental institutions, and even schools.

HINT

Students sometimes intuit that if there is no consequence to a behavior, its likelihood will be unchanged; remember, unless behaviors are reinforced, the likelihood of their recurrence decreases.

Intuitively, you probably realize that what functions as a reinforcer for some may not have the same effect on others. Even primary reinforcers, like food, will affect different animals in different ways depending, most notably, on how hungry they are. This idea, that the reinforcing properties of something depend on the situation, is expressed in the *Premack principle*. It explains that whichever of two activities is preferred can be used to reinforce the activity that is not preferred. For instance, if Peter likes apples but does not like to practice for his piano lesson, his mother could use apples to reinforce practicing the piano. In this case, eating an apple is the preferred activity. However, Peter's brother Mitchell does not like fruit, including apples, but he loves to play the piano. In his case, playing the piano is the preferred activity, and his mother can use it to reinforce him for eating an apple.

Reinforcement Schedules

When you are first teaching a new behavior, rewarding the behavior each time is best. This process is known as *continuous reinforcement*. However, once the behavior is learned, higher response rates can be obtained using certain partial-reinforcement schedules. In addition, according to the *partial-reinforcement effect*, behaviors will be more resistant to extinction if the animal has not been reinforced continuously.

Reinforcement schedules differ in two ways:

- What determines when reinforcement is delivered—the number of responses made (ratio schedules) or the passage of time (interval schedules).
- The pattern of reinforcement—either constant (fixed schedules) or changing (variable schedules).

A fixed-ratio (FR) schedule provides reinforcement after a set number of responses. For example, if a rat is on an FR-5 schedule, it will be rewarded after the fifth bar press. A variable-ratio (VR) schedule also provides reinforcement based on the number of bar presses, but that number varies. A rat on a VR-5 schedule might be rewarded after the second press, the ninth press, the third press, the sixth press, and so on; the average number of presses required to receive a reward will be five.

A fixed-interval (FI) schedule requires that a certain amount of time elapse before a bar press will result in a reward. In an FI-3 minute schedule, for instance, the rat will be reinforced for the first bar press that occurs after three minutes have passed. A variable-interval (VI) schedule varies the amount of time required to elapse before a response will result in reinforcement. In a VI-3 minute schedule, the rat will be reinforced for the first response made after an average time of three minutes.

Variable schedules are more resistant to extinction than fixed schedules. Once an animal becomes accustomed to a fixed schedule (being reinforced after x amount of time or y number of responses), a break in the pattern will quickly lead to extinction. However, if the reinforcement schedule has been variable, noticing a break in the pattern is much more difficult. In effect, variable schedules encourage continued responding on the chance that just one more response is needed to get the reward.

> **HINT**
>
> Variable schedules are more resistant to extinction than fixed schedules, and all partial-reinforcement schedules are more resistant to extinction than continuous reinforcement.

Sometimes one is more concerned with encouraging high rates of responding rather than resistance to extinction. For instance, someone who employs factory workers to make widgets wants the workers to produce as many widgets as possible. Ratio schedules promote higher rates of responding than interval schedules. It makes sense that when people are reinforced based on the number of responses they make, they will make more responses than if the passage of time is also a necessary precondition for reinforcement as it is in interval schedules. Factory owners historically paid for piece work; that is, the more workers produced, the more they were paid.

Biology and Operant Conditioning

Just as limits seem to exist concerning what one can classically condition animals to learn, limits seem to exist concerning what various animals can learn to do through operant conditioning. Researchers have found that animals will not perform certain behaviors that go against their natural inclinations. For instance, rats will not walk backward. In addition, pigs refuse to put disks into a banklike object and tend, instead, to bury the disks in the ground. The tendency for animals to forgo rewards to pursue their typical patterns of behavior is called *instinctive drift.*

COGNITIVE LEARNING

Radical behaviorists like Skinner assert that learning occurs without thought. However, cognitive theorists argue that even classical and operant conditioning have a cognitive component. In classical conditioning, such theorists argue that the subjects respond to the CS because they develop the expectation that it will be followed by the US. In operant conditioning, cognitive psychologists suggest that the subject is cognizant that its responses have certain consequences and can therefore act to maximize their reinforcement.

The Contingency Model of Classical Conditioning

The Pavlovian model of classical conditioning is known as the *contiguity model* because it postulates that the more times two things are paired, the greater the learning that will take place. Contiguity (togetherness) determines the strength of the response. Robert Rescorla revised the Pavlovian model to take into account a more complex set of circumstances. Suppose that dog 1, Rocco, is presented with a bell paired with food ten times in a row. Dog 2, Sparky, also experiences ten pairings of bell and food. However, intermixed with those ten trials are five trials in which food is presented without the bell and five more trials in which the bell is rung but no food is presented. Once these training periods are over, which dog will have a stronger salivation response to the bell? Intuitively, you will probably see that Rocco will, even though a model based purely on contiguity would hypothesize that the two dogs would respond the same since each has experienced ten pairings of bell and food.

Rescorla's model is known as the *contingency model* of classical conditioning and clearly rests upon a cognitive view of classical conditioning. *A* is contingent upon *B* when *A* depends upon *B* and vice versa. In such a case, the presence of one event reliably predicts the presence of the other. In Rocco's case, the food is contingent upon the presentation of the bell; one does not appear without the other. In

HINT

Pavlov's contiguity model of classical conditioning holds that the strength of an association between two events is closely linked to the number of times they have been paired in time. Rescorla's contingency model of classical conditioning reflects more of a cognitive spin, positing that it is necessary for one event to reliably predict another for a strong association between the two to result.

Sparky's experience, sometimes the bell rings and no snacks are served, other times snacks appear without the annoying bell, and sometimes they appear together. Sparky learns less because, in her case, the relationship between the CS and US is not as clear. The difference in Rocco's and Sparky's responses strongly suggest that their expectations or thoughts influence their learning.

In addition to operant and classical conditioning, cognitive theorists have described a number of additional kinds of learning. These include observational learning, latent learning, abstract learning, and insight learning.

Observational Learning

As you are no doubt aware, people and animals learn many things simply by observing others. Watching children play house, for example, gives us an indication of all they have learned from watching their families and the families of others. Such *observational learning* is also known as *modeling* and was studied a great deal by Albert Bandura in formulating his social-learning theory. This type of learning is said to be species-specific; it only occurs between members of the same species.

Modeling has two basic components: observation and imitation. By watching his older sister, a young boy may learn how to hit a baseball. First, he observes her playing baseball with the neighborhood children in his backyard. Next, he picks up a bat and tries to imitate her behavior. Observational learning has a clear cognitive component in that a mental representation of the observed behavior must exist in order to enable the person or animal to imitate it.

A significant body of research indicates that children learn violent behaviors from watching violent television programs and violent adult models. Bandura, Ross, and Ross's (1963) classic Bobo doll experiment illustrated this connection. Children were exposed to adults who modeled either aggressive or nonaggressive play with, among other things, an inflatable Bobo doll that would bounce back up after being hit. Later, given the chance to play alone in a room full of toys including poor Bobo, the children who had witnessed the aggressive adult models exhibited strikingly similar aggressive behavior to that which they had observed. The children in the control group were much less likely to aggress against Bobo, particularly in the ways modeled by the adults in the experimental condition.

Latent Learning

Latent learning was studied extensively by Edward Tolman. Latent means hidden, and latent learning is learning that becomes obvious only once a reinforcement is given for demonstrating it. Behaviorists had asserted that learning is evidenced by gradual changes in behavior, but Tolman conducted a famous experiment illustrating that sometimes learning occurs but is not immediately evidenced. Tolman had three groups of rats run through a maze on a series of trials. One group got a reward each time it completed the maze, and the performance of these rats improved steadily over the trials. Another group of rats never got a reward, and their performance improved only slightly over the course of the trials. A third group of rats was not rewarded during the first half of the trials but was given a reward during the second half of the trials. Not surprisingly, during the first half of the trials, this group's performance was very similar to the group that never got a reward. The interesting finding, however, was that the third group's performance improved dramatically and suddenly once it began to be rewarded for finishing the maze.

Tolman reasoned that these rats must have learned their way around the maze during the first set of trials. Their performance did not improve because they had no reason to run the maze quickly. Tolman credited their dramatic improvement in maze-running time to latent learning. He suggested they had made a mental representation, or cognitive map, of the maze during the first half of the trials and evidenced this knowledge once it would earn them a reward.

Abstract Learning

Abstract learning involves understanding concepts such as *tree* or *same* rather than learning simply to press a bar or peck a disk in order to secure a reward. Some researchers have shown that animals in Skinner boxes seem to be able to understand such concepts. For instance, pigeons have learned to peck pictures they had never seen before if those pictures were of chairs. In other studies, pigeons have been shown a particular shape (for example, square or triangle) and rewarded in one series of trials when they picked the same shape out of two choices and in another set of trials when they pecked at the different shapes. Such studies suggest that pigeons can understand concepts and are not simply forming S-R connections, as Thorndike and Skinner had argued.

Insight Learning

Wolfgang Kohler is well known for his studies of *insight learning* in chimpanzees. Insight learning occurs when one suddenly realizes how to solve a problem. You have probably had the experience of skipping over a problem on a test only to realize later, in an instant, (we hope before you handed the test in) how to solve it.

Kohler argued that learning often happened in this sudden way due to insight rather than because of the gradual strengthening of the S-R connection suggested by the behaviorists. He put chimpanzees into situations and watched how they solved problems. In one study, Kohler suspended a banana from the ceiling well out of reach of the chimpanzees. In the room were several boxes, none of which was high enough to enable the chimpanzees to reach the banana. Kohler found the chimpanzees spent most of their time unproductively rather than slowly working toward a solution. They would run around, jump, and be generally upset about their inability to snag the snack until, all of a sudden, they would pile the boxes on top of each other, climb up, and grab the banana. Kohler believed that the solution could not occur until the chimpanzees had a cognitive insight about how to solve the problem.