Rathus, R.A. (2010) Section 6.2-Operant Conditioning *Psychology: Principles in Practice.* Hold McDougal: Austin, TX. pp.165-171

Operant Conditioning and Reinforcement

In classical conditioning, we learn to associate one stimulus with another. Pavlov's dogs learned to associate a ringing bell with food. Because of classical conditioning, the response made to one stimulus (for example, food) is then made in response to the other (for example, the bell).

Classical conditioning, however, is only one type of learning. Another type of learning is operant conditioning. In **operant conditioning**, people and animals learn to do certain things— and not to do others—because of the results of what they do. In other words, they learn from the consequences of their actions. They may learn to engage in behavior that results in desirable consequences, such as receiving food or social approval. Or they might learn to avoid behaviors that result in negative consequences, such as pain or failure.

In classical conditioning, the conditioned responses are often involuntary biological behaviors, such as salivation or eye blinks. In operant conditioning, however, voluntary responses— behaviors that people and animals have more control over, such as studying— are conditioned. In operant conditioning, an organism learns to do something because of its effects or consequences.

To study operant conditioning, psychologist B. F. Skinner devised an animal cage that has been dubbed the "Skinner box." A Skinner box is ideal for laboratory experimentation. Treatments can be introduced and re moved, and the results can be carefully observed.

In a classic experiment, a rat in a Skinner box was deprived of food. The box was designed so that when a lever inside was pressed, some food pellets would drop into the box. At first, the rat sniffed its way around the box and engaged in random behavior. The rat's first pressing of the lever was accidental. But food appeared.

Soon the rat began to press the lever more frequently. It had learned that pressing the lever would make the food pellets appear. The pellets reinforced the lever-pressing behavior. **Reinforcement** is the process by which a stimulus (in this case, the food) increases the chances that the preceding behavior (in this case, the lever pressing) will occur again. After several reinforced responses, the rat pressed the lever quickly and frequently, until it was no longer hungry.

In operant conditioning, it matters little why the person or animal makes the first response that is reinforced. It can be by chance, as with the rat in the Skinner box, or the person or animal can be physically guided into the response. In training a dog to sit on command, the dog's owner may

say, “Sit!” and then push the dog’s rear end down. Once the dog is sitting, its response can be reinforced with a pat on the head or a food treat.

People can simply be told what they need to do when they are learning how to do things such as boot up a computer. In order for the behavior to be reinforced, however, people need to know whether they have made the correct response. If the computer does not start up the learner will probably think he or she has made a mistake and will not repeat the response. But if the computer does start up, the response will appear to be correct, and the learner will repeat it. Knowledge of results is often all the reinforcement that people need to learn new skills.

Types of Reinforcers

The stimulus that encourages a behavior to occur again is called a reinforcer. There are several different types of reinforcers. Reinforcers can be primary or secondary. They can also be positive or negative.

Primary and Secondary Reinforcers Reinforcers that function due to the biological makeup of an organism are called **primary reinforcers**. For example, food, water, and adequate warmth are all primary reinforcers. People and animals do not need to be taught to value food, water, and warmth.

The value of **secondary reinforcers**, however, must be learned. Secondary reinforcers initially acquire their value through being paired with established reinforcers. Money, attention, and social approval are usually all secondary reinforcers. Money, for example, is a secondary reinforcer because we have learned that it may be exchanged for primary reinforcers such as food and shelter.

Secondary reinforcers may acquire their value through a long chain of associations. For example, good grades can lead to a good college, which can lead to a good job, which can lead to more money and social approval. As a result, good grades may come to be desired in and of themselves.

Positive and Negative Reinforcers Reinforcers can also be positive or negative. **Positive reinforcers** increase the frequency of the behavior they follow when they are applied. Food, fun activities, and social approval are good examples of positive reinforcers. In positive reinforcement, a behavior is reinforced because a person (or an animal) receives something he or she wants following the behavior.

Different reinforcers work with different people. For people who enjoy sports, for example, the opportunity to participate in a sport is a positive reinforcer. For people who do not enjoy sports, however, the opportunity to participate in a sport would not be an effective reinforcer. Similarly, what serves as a reinforcer at one time for a person may not be as effective at another time for that

same person. When a person is hungry, food will work well as a positive reinforcer. But once the person has eaten and is full, food will no longer have an effect.

Unlike positive reinforcement, with negative reinforcement a behavior is reinforced because something unwanted *stops* happening or is removed following the behavior. **Negativereinforcers** increase the frequency of the behavior that follows when they are removed. Negative reinforcers are unpleasant in some way. Discomfort, fear, and social disapproval are negative reinforcers.

Daily life is filled with examples of negative reinforcement. When we become too warm in the sun, we move into the shade. When a food particle is stuck between our teeth, we floss to remove it. Both of these situations involve some uncomfortable stimulus—a negative reinforcer—that we act on to make the discomfort disappears. When a specific behavior reduces or removes the discomfort, that behavior is reinforced, or strengthened. For example, if a child does not want to perform a certain activity, he or she may scream, yell, or whine until the parent relents. The child's action has been reinforced and likely the behavior will be repeated when the child refuses to do something else demanded by the parent.

Rewards and Punishments

Many people believe that being positively reinforced is the same as being rewarded and that being negatively reinforced is the same as being punished. There are some differences, however, particularly between negative reinforcement and punishment.

Rewards Rewards, like reinforcers, increase the frequency of a behavior, and some psychologists do use the term reward interchangeably with the term positive reinforcement.But Skinner preferred the concept of reinforcement to that of reward because the concept of reinforcement can be explained without trying to "get inside the head" of an organism to guess what it will find rewarding. A list of reinforcers is arrived at by observing what kinds of stimuli increase the frequency of a behavior.

Punishments While rewards and positive reinforcers are similar, punishments are quite different from negative reinforcers. Both negative reinforcers and punishments are usually unpleasant. But negative reinforcers increase the frequency of a behavior by being removed. Punishments, on the other hand, are unwanted events that, when they are applied, decrease the frequency of the behavior they follow.

Some school districts tie participation in athletic programs to academic grades, and both punishment and negative reinforcement are involved. To the athlete on the team who does not achieve the required grades, being removed from the team is a punishment. But once the student is off the team, the disappointment over being banned from participation is a negative reinforcer. The student may work harder to raise his or her class grades in order to rejoin the team, thus ending the disappointment.

Strong punishment can rapidly end undesirable behavior. But punishment tends to work only when it is guaranteed. If a behavior is punished some of the time but goes unnoticed the rest of the time, the behavior will probably continue.

Most psychologists feel it is preferable to reward children for desirable behavior rather than punish them for unwanted behavior.

For example, parents and other authority figures should pay attention to children and praise them when the children are behaving well. If good behavior is taken for granted, and only misbehavior receives attention, misbehavior may get reinforced.

Psychologists also point out that children need to be aware of the desired behavior. In addition, they need to be capable of performing it. For example, consider a situation in which parents punish a child for not listening to directions only to find out much later that the child has a hearing problem and could not hear the directions.

Schedules of Reinforcement

A major factor in determining just how effective a reinforcement will be in bringing about a behavior has to do with the **schedule of reinforcement**—when and how often the reinforcement occurs.

Continuous and Partial Reinforcement

We have been discussing **continuous reinforcement**, or the reinforcement of a behavior every time the behavior occurs. For example, the rats in the Skinner box received food every time they pressed the lever. If you go to a friend's house and your friend is there every time, you will probably continue to go to that same location each time you want to see your friend because you have always been reinforced for going there. New behaviors are usually learned most rapidly through continuous reinforcement.

It is not, however, always practical or even possible to reinforce a person or an animal for a behavior every single time the behavior occurs. Moreover, a person or animal who is continuously reinforced for a behavior tends to maintain that behavior only as long as the reinforcement is still there. If for some reason the reinforcement stops occurring, the behavior disappears very quickly. For example, if you go to your friend's house only to be told that your friend no longer lives there, you almost certainly will not return to that house again in search of your friend.

The alternative to continuous reinforcement is called partial reinforcement. In **partial reinforcement**, a behavior is not reinforced every time it occurs. People who regularly go to the movies, for example, may not enjoy every movie they see, but they continue to go because they enjoy at least some of the movies. Behaviors learned through partial reinforcement tend to last longer after they are no longer being reinforced at all than do behaviors learned through continuous reinforcement.

There are two basic categories of partial reinforcement schedules. The first category concerns the amount of time (or interval) that must occur between the reinforcements of a behavior. The second category concerns the number of correct responses that must be made before reinforcement occurs (the ratio of responses to reinforcers).

Interval Schedules If the amount of time— the interval—that must elapse between reinforcements of a behavior is greater than zero seconds, the behavior is on an interval schedule of reinforcement. There are two different types of interval schedules: fixed-interval schedules and variable-interval schedules. These schedules affect how people allocate the persistence and effort they apply to certain tasks.

In a fixed-interval schedule, a fixed amount of time—say, five minutes—must elapse between reinforcements. Suppose a behavior is reinforced at 10:00. If the behavior is performed at 10:02, it will not be reinforced at that time. However, at 10:05, reinforcement again becomes available and will occur as soon as the behavior is performed. Then the next reinforcement is not available until five minutes later, and so on. Regardless of whether or how often the desired behavior is performed during the interval, it will not be reinforced again until five minutes have elapsed.

The response rate falls off after each reinforcement on a fixed-interval schedule. It then picks up as the time when reinforcement will be dispensed draws near. For example, in a one-minute fixed-interval schedule, a rat may be reinforced with food the first time it presses the lever after a minute has elapsed since the previous reinforcement. After each reinforcement, the rat's rate of lever pressing slows down, but as a minute approaches, lever pressing increases in frequency. It is as if the rat has learned that it must wait a while before reinforcement is available.

Similarly, if you know that your teacher gives a quiz every Friday, you might study only on Thursday nights. After each quiz, you might not study again until the following Thursday. You are on a one-week fixed-interval schedule.

Farmers and gardeners are quite familiar with one-year fixed-interval schedules. If a particular type of fruit ripens only in the spring, for example, the farmer probably will not check to see if the fruit is ripe in the autumn or winter. However, as spring begins, the farmer will probably check more and more frequently to see if the fruit is ripe. Once all the fruit has ripened and been picked, the farmer will stop checking until the next spring.

In a variable-interval schedule, varying amounts of time go by between reinforcements. For example, a reinforcement may occur at 10:00, then not again until 10:07 (7-minute

interval), then not again until 10:08 (1-minute interval), and then not again until 10:20 (12-minute interval).

In variable-interval schedules, the timing of the next reinforcement is unpredictable. Therefore, the response rate is steadier than with fixed-interval schedules. For example, if your teacher gives unpredictable pop quizzes, you are likely to do at least some studying fairly regularly because you do not know when the next quiz will be. And since there is always the chance that it could be tomorrow, you want to be prepared.

Ratio Schedules If a desired response is reinforced every time the response occurs, there is a one-to-one (1:1) **ratio** of response to reinforcement (one response, one reinforcement). If, however, the response must occur more than once in order to be reinforced, there is a higher response-to-reinforcement ratio. For example, if a response must occur five times before being reinforced, the ratio is 5:1. As with interval schedules, there are fixed-ratio schedules and variable-ratio schedules.

In a fixed-ratio schedule, reinforcement is provided after a fixed number of correct responses have been made. The rat in the box would have to press the lever, say, five times, and always exactly five times, in order to receive the food. Some stores use fixed-ratio schedules to encourage people to buy more. A video rental store, for instance, may promise customers a free video rental after payment for five rentals.

With a fixed-ratio schedule, the person or animal tends to try to get its fixed number of responses "out of the way" as quickly as it can to get to the reward. With the free video rental offer, for example, a customer may rent the five required videos as soon as possible to get the free one sooner. If the ratio is very high, however, it is often less effective, particularly with people.

With a variable-ratio schedule, reinforcement can come at any time. Sometimes the rat might have to press the lever 5 times to get the food; at other times, 8 or even 14 times.

The rat cannot predict how many times the lever must be pressed because the number changes each time.

This unpredictability maintains a high response rate. Slot machines tend to work on variable-ratio schedules. Even though the players do not know when (or even if) they will win, they continue to drop coins into the machines. And when the players do win, they often continue to play because the next winnings might be just a few lever-pulls away.

Extinction in Operant Conditioning In operant conditioning, as in classical conditioning, extinction sometimes occurs. In both types of conditioning, extinction occurs because the events that had previously followed a stimulus no longer occur.

In operant conditioning, the extinction of a learned response results from repeated performance of the response without reinforcement. In Skinner's experiment with the rats, lever pressing was followed by—and reinforced by—food. But if a rat presses a lever repeatedly and no food follows, it will eventually stop pressing the lever. The lever-pressing behavior will have been extinguished.

Applications of Operant Conditioning

As we have seen, even people who have never had a course in psychology use operant conditioning every day to influence other people. For example, parents frequently use rewards, such as a trip to the park or ice cream, to encourage children to perform certain tasks, such as cleaning their rooms. Some specific applications of operant conditioning in education include shaping, chaining, and programmed learning.

Shaping and Chaining If you have ever tried to teach someone how to do a complex or difficult task, you probably know that the best way to teach the task is to break it up into parts and teach each part separately. When all the parts have been mastered, they can be put together to form the whole task. Psychologists call this shaping. **Shaping** is a way of teaching complex behaviors in which one first reinforces small steps in the total activity.

Learning to ride a bicycle, for example, involves the learning of a complex sequence of behaviors and can be accomplished through shaping and chaining. In **chaining**, chaining, each step of a sequence must be learned and must lead to the next until the final action is achieved. The steps create a response chain. Sometimes several response chains must be learned to complete a desired action. For example, in learning to ride a bike a person must learn to move the bike forward by using the pedals. You may have seen a parent help a young child by holding the seat as the child learns to pedal. Then they must learn to balance the bicycle, and then to steer it. At first, each of these steps seems difficult, and people must pay close attention to each one. After many repetitions, though, and much praise and reassurance from the instructor, each step— and eventually bicycle riding itself—becomes habitual.

Chaining can occur either in a forward or backward response chain. In a forward chain, each step leads to a final goal. If you are learning to tie your shoes you eventually reach the point at which you have performed all the tasks in the chain and your shoes are tied. In backward chaining, you start with the final action and dissect each step it takes to get to that point. For example, suppose you want to know the steps in putting on a jacket. In backward chaining, you start with the jacket on and go backward through the steps necessary to put on the jacket.

Psychologists have used chaining to teach complex behavior patterns to animals. For example, they have trained rats to pedal toy cars by first reinforcing the rats' behavior of turning toward the cars. Next they wait until the rats approach the cars before providing further reinforcement. Then they wait until the rats touch the cars, and so on. In this way, rats have been trained to run up ramps, cross bridges, and climb ladders. This type of learning has also been used to train service animals to help people with disabilities.

Programmed Learning B. F. Skinner developed an educational method called programmed learning that is based on shaping and chaining. Programmed learning assumes that any task, no matter how complex, can be broken down into small steps. Each step can be shaped individually and combined to form the more complicated whole.

In programmed learning, a device called a teaching machine presents the student with the subject matter in a series of steps, each of which is called a frame. Each frame requires the student to make some kind of response, such as answering a question. The student is immediately informed whether the response is correct. If it is correct, the student goes on to the next frame. If the response is incorrect, however, the student goes back over that step until he or she learns it correctly.

These days, teaching machines are most likely to be computers that are programmed so that the material can branch off in several different directions, depending on where the student needs instruction and practice. The use of computers in learning is called computer-assisted instruction.

Programmed learning does not punish students for making errors. Instead, it reinforces correct responses. Teaching machines are infinitely patient with the learner. They are also highly efficient. Eventually, all students who finish the program earn "100 percent"—but they do so by learning in small steps at their own pace.