Operant Conditioning: An Overview

Classical, or Pavlovian, conditioning is a process by which new emotional and glandular reactions develop in response to previously neutral stimuli in the environment. But classical conditioning doesn't explain how we develop new skills or behaviors instrumental in changing our external environments. That learning process involves what is typically referred to as instrumental, or operant, conditioning. Operant conditioning describes how we develop behaviors that "operate upon the environment" to bring about behavioral consequences in that environment. Operant conditioning applies many techniques and procedures first investigated by E. L. Thorndike (1898) but later refined and extended by B. F. Skinner (Skinner, 1938).

Thorndike was an American psychologist who was one of the first to investigate the effects of behavioral consequences on learning. His work led him to emphasize both the effects of positive as well as negative behavioral consequences. Because behaviors are instrumental in bringing about such consequences by operating upon the environment in some way, this process for developing new skilled behaviors was first called instrumental conditioning. In subsequent literatures, especially in those inspired by the work of Skinner, the term "instrumental conditioning" was replaced by the term "operant conditioning." Nevertheless it was Thorndike who first concluded that positive consequences strengthen behaviors to make them more likely in similar situations in the future; a phenomenon he labeled the Law of Effect.

Inspired by the much earlier work of both Pavlov and Thorndike, another American Psychologist, B.F. Skinner, went on to develop the principles of operant conditioning. Skinner formalized these principles and identified many variables involved in this form of learning. For example, Skinner revised Thorndike's concept of "reward" by emphasizing that it has "positive reinforcement" effects which result in the increased likelihood of a behavior's future occurrences. Even painful consequences can increase the future likelihood of behaviors that eliminate or avoid such consequences, and thus Skinner emphasized their function as "negative reinforcements." According to Skinner, reinforcement, whether positive or negative, is the process of increasing future behavioral probabilities; meaning any response that is followed by a reinforcer will increase in its frequency of occurrence across time (a concept emphasizing the rate of specific ways of behaving). Skinner also discovered that such reinforcing events don't have to happen each and every time. Instead, intermittent reinforcement is also effective, and Skinner described the effects that different "schedules" of reinforcement (the timing or frequency of reinforcement) have on behavior. He also identified the process of punishment whereby behavioral probabilities may be decreased by consequences. Any behavior that is followed by punishment decreases in frequency.

Using the variables controlling operant conditioning as a foundation, B.F. Skinner also investigated several alternative operant procedures. For example, shaping (Peterson, n.d.) is a process of operantly conditioning a new form of behavior by reinforcing successive approximations to the ultimately desired form of behavior, known as the "target" behavior. Shaping is simply a method for conditioning an organism to perform a new behavior by reinforcing small and gradual steps toward the desired form of behavior. The experimenter starts the shaping process by reinforcing what the individual already does, then by gradually reinforcing only selective variations of that behavior that lead to the form of the target behavior, the experiment may gradually reach the point where only the target behavior is the one reinforced.

Behaviors come to occur only within certain antecedent environmental circumstances through a process called stimulus discrimination conditioning, and this process was also investigated in detail by Skinner. Both stimulus discrimination and generalization exist in classical conditioning processes as well, but we will currently focus only on these processes as they are employed in operant conditioning. Likewise, procedures exist for the extinction of operant behaviors as well as the parallel process of extinction of classically conditioned responses. In operant conditioning, extinction is the process wherein experimenters stop reinforcing or punishing a specific behavior, thus resulting in that behavior's return to pre-conditioning rates, or probabilities, of occurrence.

Operant conditioning techniques and procedures have many applications across various circumstances and problem areas. They have been utilized in the classroom environment with a great deal of success. Early forms of teaching machines first elaborated by Skinner have evolved into today's computer assisted instruction programs. Such programs allow students to receive feedback on their progress in mastering various types of subject matter while simultaneously shaping students to become more skilled in more generalized behaviors such as reading comprehension (Ray, 2004). Appropriate and learning-supportive classroom behaviors may also be developed and maintained with operant principles. In a process called the token economy, students are rewarded for good classroom behaviors or even independent study management programs using points or tokens that may be traded for more primary forms of reinforcement. Token economies are also used in psychiatric hospitals and other institutions to maintain and teach appropriate and adaptive behaviors (Ayllon |_2 Azrin, 1968).

Operant conditioning also has made significant contributions in therapeutic settings. For example, anxiety and other similar physiological disorders can be treated with a technique known as biofeedback. Many of the earliest demonstrations of biofeedback came from Neal Miller's laboratory (Miller, 1969), and typically involved the monitoring of a patient's vital body functions (heart and breathing rate, blood pressure, etc.) while also displaying their status via some type of feedback device, such as a computer display. Patients may be trained to relax or otherwise behave in ways that keep their vital physiological processes within a more healthy range of functioning. Feedback telling individuals that they are being successful reinforces their efforts to control their own physiological functions. Patients may, for example, wear portable monitoring and feedback devices to learn to relax in usually anxiety provoking situations.

Procedures like shaping and chaining are also applied in the training of performance animals, in obedience classes for family pet, and in the training of animals as personal assistants for blind or paralyzed humans (Pryor, 1985). Performance animals like these, as well as the ones seen in marine parks and circuses, learn complex chains of behaviors through operant conditioning procedures involving reinforcement and antecedent stimulus discrimination. For example dogs in obedience classes are taught to behave to simple commands that offer visual and auditory cues antecedent to desired behaviors as well as positive reinforcement as consequences for performing those behaviors in response to those cues.

Origins of Operant Conditioning: Instrumental Learning and the Law of Effect

Edward Lee Thorndike was an American psychologist studying animal learning while a graduate student at Harvard University in the late 1890's. He was especially interested in how animals learn to engage in new behaviors that are instrumental in solving problems, such as escaping from a confined space. The instrumental character of behavior in changing an animals' circumstances led subsequent authors to refer to Thorndike's form of learning as instrumental learning, although Thorndike preferred to describe it as "trial and success" learning (Chance, 1999). Pretty much these same behavioral processes were renamed operant conditioning by a much later researcher, B.F. Skinner (Skinner, 1938), who was also interested in how such skills "operate upon environments" (hence his more descriptive term "operant") to bring about significant consequences for the individual.

Thorndike designed many ingenious experiments into study such behavior. In one series of investigations Thorndike placed hungry cats into an apparatus called a puzzle box, from which the animals learned to escape to obtain rewards of food. At first Thorndike's cats seemed to behave almost randomly, using trial and error to find their way out of the puzzle box. Thorndike graphed the time it took an animal to escape from the puzzle box for each successive trial he gave the animal. He quickly discovered that the time for escape gradually declined over several repeated trials, with each successive trial typically taking less and less time. He called this a learning curve and proposed that the slope of this curve reflected the rate at which learning occurred (Chance, 1999). From such studies Thorndike proposed his Law of Effect, which states that if successful behaviors in a trial and error situation are followed by pleasurable consequences, those behaviors become strengthened, or "stamped in" and will thus be more quickly performed in future trials (Thorndike, 1898).

As noted above, in order to study the problem-solving behavior of cats using trial and error procedures Thorndike developed a special puzzle box apparatus. Various forms of puzzle boxes were constructed, but a typical one was a wooden cage equipped with a door held by a weighted loop of string holding, and a pedal, and a bar. A cat had to press the pedal, pull the string, then push the bar to unlatch the door to the box. This allowed the animal to then escape from the box and obtain food as a consequence.

The term instrumental conditioning is used to describe Thorndike's procedures for animal learning because the term ties behaviors to the generation of their consequences in learning-that is, the behavior is instrumental in obtaining important consequential outcomes in the environment. Thorndike's procedures involved what many refer to as "trial and error" procedures. For example, when Thorndike placed a hungry cat into his puzzle box, the cat would produce many behaviors in its attempts to escape the confinement. Eventually, the animal would produce the correct behavior quite by chance, usually clawing a string and then stepping on a pedal to open the door. This correct behavior had consequences because Thorndike would leave a plate of food just outside the box that the cat would eat from once it escaped. Thorndike's Law of Effect proposed that such rewards strengthened the behaviors that obtained the reward, thus making that behavior more quickly performed with fewer errors on future trials.

Thorndike's Law of Effect took two forms: the "strong" form and the "weak" form. Food as consequences represented the strong, or behavioral strengthening, form. The "weak" side of the Law of Effect describes what happens when a behavior fails to accomplish such pleasurable consequences, thus leading to a weakened, or "stamped out" impulse to behave in a similar fashion in similar situations in the future. Thorndike's studies were among the first to demonstrate and precisely measure the power of consequences in the environment (especially rewards) and their ability to control behavior, and thus Thorndike's work laid the foundation for the subsequent development of a more behavioral perspective on the learning process.

Operant Conditioning Principles

Another American Psychologist working at Harvard, B. F. Skinner, also studied the behavior of animals with a focus on consequences. Although Skinner's work came much later than that of Thorndike (Skinner began his work on operant conditioning in the 1930s), his research was based on the principles Thorndike had identified. Skinner (1938) believed that in order to understand psychology you had to focus only on observable behaviors.

Because observable behaviors and the role environments play in developing and controlling those behaviors are the focus of operant conditioning, Skinner and the field of operant conditioning is often considered to represent the most radical form of the perspective on learning called "behaviorism." Thorndike's work anticipates this movement as well, but Thorndike predated the philosophical emphasis on observable behaviors as the exclusive outcomes in learning. Throughout our discussion of operant conditioning, you will read terms such as "behavioral," "behaviorism," and "behaviorist." These terms typically refer to the work of Pavlov and Thorndike as the foundations of the perspective, but it was John Watson (1913) who described the perspective in most detail, and Skinner (1938) who most completely illustrated the power of the approach in what he called "radical behaviorism". These researchers emphasized the importance of observable behavior in understanding psychology and generally excluded mental activities in their studies. Because of this focus on behavior, their work is deemed "behavioral" and their conceptualization of learning is labeled "behaviorism." Keep in mind that this term does not include the cognitive or ecological perspectives.

Through his research, Skinner's radical behaviorism (1938) identified variables and formalized procedures using those variables in a conceptualization to learning called "operant conditioning." This term comes from Skinner's emphasis on the fact that behaviors operate (thus being an "operant") on the environment in order to gain certain consequential stimuli and to avoid others. Unlike classical conditioning, which Skinner called Respondent Conditioning because it focuses on the processes of learning in reflexive responses, operant conditioning focuses on how organisms learn totally new behaviors through experience with consequences in the environment. Skinner's operant conditioning is founded on Thorndikes' instrumental conditioning, but Operant Conditioning involves a wider variety of processes and labels consequences quite differently.

Skinner used rats as subjects for much of his work, but he is even more famous for his later work with pigeons. Dissatisfied with mazes or Thorndike's puzzle box, Skinner designed an apparatus to study animal behavior in a slightly different fashion. The operant chamber, or Skinner box as it came to be known, was designed to prevent human interruption of the experimental session and to allow the study of behavior as a continuous process, rather than in separated trial-by-trial procedures.

In Thorndike's puzzle box, the animal would have to be physically placed back into the box after each rewarded escape trial. Skinner felt that such procedures interfere with behavior as a "stream of events". For rats an operant chamber has a lever (technically called an operandum) that can be pressed over and over to deliver food pellets, with each press counting as a single occurrence of the behavior. For pigeons, one or more disks can be pecked as the operanda to deliver reinforcement for this behavior, usually in the form of food grain. The disks are often lighted for stimulus discrimination and generalization training. After an animal receives a reinforcement for pressing a bar or pecking a disk, there is no need to reset the system; the chamber is ready to deliver more reinforcements as soon as the animal responds again.

The cumulative recorder was another innovation introduced by Skinner to automatically graph response rates (that is, it shows an accumulation of the number of total responses as this total is distributed across time). in its original form, this machine recorded the number and timing of an operant behavior by using a continuously rolling piece of paper with a fixed ink-pen to mark time across a continuous X axis, as well as another pen that advanced one step up the Y axis each time a bar was pressed or key was pecked. Skinner was able to study animal behavior for as long as he deemed necessary without ever having to interfere with or even observe his animal.

Almost all of what Skinner (1938) discovered about operant conditioning principles came from his use of the operant conditioning chamber and its cumulative recorder-produced data. One procedure and its associated variables that Skinner identified was that of reinforcement. According to Skinner reinforcement involves the presentation or removal of stimulus consequences that increase the future rate of any specific class of operant behaviors, such as bar pressing or key pecking. The consequential stimulus variable is considered to be a reinforcer only if it's presentation or removal as a consequence for a behavior increases the future rate or probability of that form of behavior.

Skinner felt that when the presentation of a stimulus results in an increase in behavioral probability, positive reinforcement has occurred. Skinner also identified two types of positive reinforcers; primary (usually biological) and secondary or conditioned (must be classically conditioned to acquire reinforcing functions like the primary stimulus has). When the removal of a stimulus as a consequence for a behavior increases the likelihood of that form of behavior, negative reinforcement has occurred. Escape and avoidance learning are how we often describe changes in behavior rates that increase because of negative reinforcement.

Skinner also studied the procedure of punishment. Punishment is the opposite of reinforcement. It occurs when the probability of a behavior decreases with the presentation or removal of a stimulus. If presentation of an aversive stimulus decreases the likelihood of behavior occurring again, positive punishment has taken place. If the removal of a positively valued stimulus decreases the chances that a behavior will occur again, negative punishment, also called time out, has occurred. Skinner noted that punishment is often an inefficient way of controlling behavior, and in order to work at all it must be applied immediately after the behavior, it must be consistent and follow after every instance of the behavior, and it must be fairly strong.

So Skinner (1938) developed his procedures for operant conditioning through the manipulation of the operant variables of reinforcing and punishing consequences. But Skinner noticed that when he presented a reinforcement every time a behavior occurred, the rat or pigeon would become satiated quickly and would stop producing a certain behavior in high rates. Skinner labeled this procedure continuous reinforcement. Eventually he tried reinforcing behaviors using a non-continuous procedure -- a process he called partial, or intermittent, reinforcement. There are several types of partial reinforcement, each with different rules for applying one or another of type of consequences (reinforcements or punishments). These rules for "scheduling" reinforcement intermittently either rely on counting behaviors, such as fixed or variable ratio rules, or adding a time interval to the behavioral rule, such as fixed or variable interval schedules. Each type of schedule rule effects behavior in different and unique ways.

Skinner was also one of the first to seriously consider a fundamental flaw in Thorndike's trial and error learning procedure. Instead of using Thorndike's vocabulary which described an animal as random "trying" to solve a problem, Skinner preferred to talk about different activities as alternative forms of emitted behaviors. And all responses or behaviors that look alike or act upon the environment in a similar fashion form a "class" of related emitted responses, or an "operant class" of behavior. This contrasts with Pavlov's elicited behaviors, such as salivation, where known unconditional stimuli are used to bring about the "respondent" behaviors, as Skinner referred to them. But what if the animal never emits the correct behavior in a trial and error situation? Having made only errors, nothing could be reinforced and thus no learning (relatively permanent change in behavior) would take place either.

Skinner believed that by manipulating consequences in a certain systematic way, an organism could be led to the correct behavior much faster than if one simply waited for the animal to happen upon the response by chance. The procedure he developed for accomplishing this step-by-step process is called shaping, and it's purpose it to reinforce behavior in gradual steps that begin with only rough approximations to the eventual "target" that one has as the goal of learning. In Skinner's research a behavior he often shaped was a lever press by a rat in an operant chamber (Skinner, 1951; Peterson, n.d.). He would first reinforce the animal for being in the vicinity of the bar, then for sniffing the bar, then touching the bar with a paw, then standing over the bar and eventually pressing the bar, all in successive approximations or gradual steps to the final bar pressing he wanted the animal to learn. Chaining is yet another procedure that is based on shaping, but it is used to condition a whole complex series of different responses, not just one.

Extinction, stimulus discrimination and stimulus generalization also exist in operant conditioning just as in classical conditioning. Extinction occurs when reinforcement or punishment no longer occurs as a consequence for a given behavior. Spontaneous recovery can also occur in operant conditioning if extinction is tested again later, and rapid reacquisition occurs if reinforcement or punishment is again the consequence for behavior. Stimulus discrimination involves presenting reinforcement or punishment only under certain antecedent stimulus conditions and not others until the organism only produces the behavior under the given antecedent settings. Generalization is the opposite: reinforcement or punishment is the consequence of behavior in many antecedent settings and the organism produces the behavior across these many different circumstances. The procedures developed by Skinner have been tested in many different applied settings and are very commonly used today.

Reinforcement in Operant Conditioning

Thorndike's studies of instrumental learning where cats learned to escape from puzzle boxes led to his conclusion that behaviors are controlled by their consequences, which was stated as his Law of Effect (Thorndike, 1898). In his studies of operant conditioning Skinner (1938) also stressed the importance of behavioral consequences, which he referred to as reinforcement and punishment. Reinforcement occurs when the probability of a certain behavior increases due to the presentation of a stimulus as a behavioral consequence (positive reinforcement) or the removal of a stimulus as a behavioral consequence (negative reinforcement).

It is important to keep in mind that reinforcement is a process and occurs only if behavioral probability increases. Thus a consequential stimulus is not a reinforcer if it's presentation (positive reinforcement) or removal (negative reinforcement) does not increase the likelihood that the behavior will occur again. We often assume that something will reinforce behavior, but until the behavior has shown an increase in probability, you cannot be sure. For instance, you may think candy would reinforce a child for studying, but if the child doesn't study more often when given candy upon doing so, candy is not a positive reinforcer.

There are two kinds of positively reinforcing stimuli (stimuli that are generally reinforcing when presented to an individual) known as primary reinforcers and secondary (or conditioned) reinforcers. Both types can be delivered following various rules for delivery, thus defining various schedules of reinforcement. Often some type of procedure, such as deprivation, is required to establish that a certain stimulus will function as an effective reinforcer. According to Skinner (1938), reinforcement is much better at controlling behavior than punishment, which is defined by a decrease in the probability of any behavior that causes the punishing stimulus to be presented (negative punishment) or removed (positive punishment, or time-out).

Another way to positively reinforce behavior is to rely upon Premack's Principle (Premack, 1959, 1971). According to the Premack Principle, a normally higher frequency behavior can be used to positively reinforce a desired behavior that is normally lower in frequency. A parent is more likely to positively reinforce a child for studying (a low frequency behavior without intervention) by allowing the child to watch TV (a high frequency behavior without intervention) after studying for some specified time. In this case, allowing the consequential behavior of watching TV causes the probability of studying to increase. The Premack Principle has also been utilized in operant conditioning research on rats. Rats can be successfully reinforced for bar pressing (very low frequency behavior without intervention) by allowing the rat to run in a running wheel (normally high frequency behavior in rats).

Skinner (1938) also found that consequences resulting in the removing an aversive (painful, uncomfortable, or undesired) stimulus that was already present could also increase the probability that a certain behavior would occur. He called this process negative reinforcement. Crucial to negative reinforcement is: 1) the pre-existing presence of an aversive stimulus, 2) then a specific form of behavior that 3) has the consequence of terminating or removing that aversive stimulus. A parent who wants to reinforce the studying behavior of the child could use negative reinforcement by removing normally required chores for a week. It is important to remember that negative reinforcement is labeled "negative" because it relies upon the removal of an aversive stimulus, not because it is a "negative" way to reinforce behavior. And it is reinforcement because the behavior that removes the stimulus increases in probability.

Frequent use of negative reinforcement, inside or outside of the laboratory, will lead to what is often referred to as escape and/or avoidance behavior, as when you have an increased probability of taking an aspirin to escape a headache or to avoid developing muscle pain after strenuous exercise. Escape is the first of two phases of behavioral development involving the use of negative reinforcement. Avoidance is the second phase. If the floor of an operant conditioning chamber is electrified to deliver a mild electrical shock, a rat's bar press in the presence of this shock is negatively reinforced when the bar pressing turns off the shock. The rat will always experience the shock, but through negative reinforcement it learns to escape this aversive stimulus by pressing the lever that terminates the shock.

A child is negatively reinforced for whining about doing chores when someone reduces the time the child spends doing those chores. In this case whining becomes a means for escaping chores. But the child still has to come into contact with the aversive event (chores) before he/she can escape them by whining. As noted above, taking aspirin for a headache is a classic example of escape learning. The reduction or elimination of pain negatively reinforces taking the medicine. The headache must be experienced for this to occur, but the individual escapes the pain through pill taking behavior.

Avoidance is also a term that refers to increasing the likelihood of behaviors by the use of negative reinforcement. Avoidance typically appears as a second phase of development following the phase of escape. If a rat learns to press a lever by escaping a brief shock, eventually that rat begins to press even before the shock is delivered if pressing delays the next onset of shock (i.e., keeps the shock from occurring for a while). In this case, the rat may never again come into contact with shock, but bar pressing continues because it has been negatively reinforced. This is the essence of avoidance learning. A child who's whining is always reinforced by the removal of chores may learn to avoid doing chores altogether by whining even before starting chores. It would be far better to establish studying as a means by which the child can avoid chores!

Sometimes avoidance learning is facilitated by using some sort of antecedent stimulus signal for the impending shock. If, for example, a light in the chamber signals that a bar press by a rat may prevent the occurrence of an electric shock, the rat's bar press will be negatively reinforced by the termination of the light (escape behavior). Of course, at the same time the rat must also be avoiding any contact with shock because shock was prevented from coming on by the bar being pressed. After only a few experiences with actual shocks following such a light signal, the rat will learn to prevent shocks altogether by pressing the bar as soon as the light turns on. The bar press is avoidance behavior that is under the control of a discriminative antecedent stimulus (the warning light), and is thus called discriminative avoidance.

Reinforcement is a naturally occurring process, and doesn't have to be managed by someone. For example, can you think of any superstitions? Many people believe that walking under a ladder will give you bad luck or finding a four-leaf clover will bring you good luck. Well, in operant conditioning, superstitious behavior is a behavior that increases in probability because it happened to be reinforced merely by chance (Skinner, 1948). This happens especially when reinforcement occurs based on rules that are independent of a specific behavior, such as time since last reinforcement, rather than on what behavior was occurring. In pigeons, superstitious behavior may include shaking wings or other unusual behaviors before pecking a disk for reinforcement. The pigeon may have shaken its wings before pecking for food when it was first reinforced. That wing-shaking behavior is said to be superstitious because it has nothing to do with gaining reinforcement, yet it has increased in likelihood none-the-less.

In humans, blowing on dice before rolling them may be a form of superstitious behavior. A gambler may have once blown on a pair of die and then won the jackpot after he/she rolled the right numbers. The gambler may now believe that this blowing behavior led to the winning and will continue to do so on every roll. This behavior is superstitious because blowing on dice has nothing to do with the numbers you roll or the winnings you obtain. This can happen in a punishment situation as well. If blowing on the die resulted in a bad roll and the gambler lost everything, blowing on die will greatly decrease in frequency if it even occurs again. This decrease in behavior is superstitious because again, blowing on the die did not determine the result of the roll or the loss of money.

Conditioned Reinforcement and Operant Conditioning

Skinner (1938) described two types of reinforcing stimuli: primary and conditioned (or secondary) reinforcers. A primary reinforcer is anything that has the power to increase behavioral probabilities because it is involved with a biological need of the organism. Food, sex and temperature stabilities are often used as reinforcement because we need them as a species. Escape from pain and social acceptance/contact can also be considered as primary reinforcers due to their evolutionary importance to humans and certain other organisms. Primary reinforcers also provide a powerful source of motivation when an organism is deprived of them.

In operant conditioning, primary reinforcers are often used because of their immediate power to modify behavior. This power comes from the fact that they are of direct biological importance to the organism. Food, water, exercise, and escape from pain can be considered as primary reinforcers because a lack of these can be physiologically harmful and/or painful. Skinner used primary reinforcers, usually food, in most of his laboratory studies with pigeons and rats.

A conditioned or secondary reinforcer is anything that can increase the probability of behavior because of its reliable association with primary reinforcers. Classical (Pavlovian) conditioning is at work here, as can be seen in the case of money. Money can be a powerful reinforcer, although it has no real use to us unless it can buy the things that meet more primary needs, such as food, shelter, and entertaining stimulation. This is classical conditioning with money as the CS and food, etc. as the UCS. A small piece of metal (such as coins) or piece of paper (such as dollar bills) has little value per se, as illustrated by play money. However, food, social contact, relief from pain, and even relief from boredom (all primary reinforcers) can be obtained with a sufficient amount of money that is legal tender for things we need. Therefore, because money has been paired with these primary reinforcers so often, it takes on the power to increase the probability of behavior in and of itself. This is why it is labeled as a secondary, or conditioned, reinforcer.

Skinner's work illustrated that deprivation is a common procedure for effectively changing the nature of a reinforcer in operant conditioning (Skinner, 1938). Such procedures are described by some researchers as establishing motivation, and are thus referred to as "establishing operations" (Michael, 1982, 1993; Dougher |_2 Hackbert, 2000). A pigeon will not press a lever very frequently for food if it is satiated (full). Depriving the animal of this primary reinforcer (usually experimental deprivation involves food or water) will motivate the animal to perform, because now the reinforcer satisfies a biological need. In operant research, animals are not deprived of food or water to a point that is dangerous or very distressing (all research with animals must follow strict ethical guidelines in any discipline). The animals are usually made just hungry or thirsty enough so that food or water works as an effective primary reinforcer. But almost anyone who has eaten too much thanksgiving dinner can relate to the fact that food may eventually turn into an aversive stimulus when too much has been consumed!

Deprivation of a secondary reinforcer, like money, works much in the same way. A person who has |_817,000,000 is not going to be highly motivated to work for money. Someone with only |_83 to his or her name will do almost any kind of work for money if they have no other means for eating or staying warm. You may have noted, however, that the person with only |_83 is probably also deprived of primary reinforcers (like food, shelter, or social contact) as well as money. And the person with |_817,000,000 may work because he/she is deprived of certain social stimuli that money may not buy.

Punishment in Operant Conditioning

Thorndike's earliest studies of cats escaping from puzzle boxes led him to distinguish between two forms of his famous Law of Effect. Thorndike (1898) held that behaviors could be "stamped in" by satisfying consequences or "stamped out" by annoying consequences. This became the basis for his distinguishing between a Strong Law of Effect and his Weak Law of Effect. When behavior is stamped out by annoyers Thorndike felt that a "strong" Law of Effect was at work. He later withdrew this punishment element of his theory, eventually leaving only the "weak" Law of Effect that resulted in a "stamping in" of behavior. His work with human subjects learning verbal behaviors had convinced him that saying "wrong" had less effect than saying nothing, and the most effective response was saying "right" to the learner's responses. Thus Thorndike interpreted these results as arguing against the effectiveness of punishment (Catania, 1998).

Likewise, in his studies of operant conditioning Skinner (1938) described the phenomenon of punishment as well as reinforcement. Positive punishment involves a decrease in the probability of a behavior through presentation of (addition of, and thus the term "positive") an aversive stimulus as a behavioral consequence. Negative punishment describes the removal of a positive stimulus as a behavioral consequence. It is worth re-emphasizing that the stimulus that is presented in positive punishment is usually a painful or otherwise aversive stimulus, while those stimuli that act as negative (removed) punishers are usually sought-after or appetitive stimuli. Sound confusing? Then let's consider these distinctions in more detail.

As noted, punishment is an operant process of decreasing the probability that a particular behavior will occur. According to Skinner (1938) a stimulus cannot be considered a "punisher" if it's presentation (positive punishment) or removal (negative punishment) does not decrease the likelihood of a behavior. For instance, it may seem intuitive that giving extra chores will be a good punishment for a child having drawn on the wall. If, however, the child continues to draw on the wall with the same frequency despite the extra chores, the chores are not punishers and punishment has not occurred.

It is easy to confuse the use of positive and negative to describe types of punishment. As in the case of positive reinforcement, positive punishment refers to the presentation of a stimulus, only now it decreases behavioral probabilities instead of the increase probability that defines reinforcement. But aversive stimuli are used often in operant conditioning procedures. Anything that causes pain, discomfort, high levels of physical and/or mental stress or anything that is undesired is classified as an aversive stimulus. In successful negative reinforcement, their removal results in an increase in the probability of a certain behavior. In positive punishment, however, the presentation of an aversive stimulus results in the decrease of the probability of a certain behavior. But because the stimulus being presented in positive punishment is usually aversive to, in other words, unpleasant for, the organism, some people are inclined to speak of such aversive stimuli as being negative for attracting/repelling the individual. See the potential for confusion? Positive punishment adds negative stimulus consequences.

As with negative reinforcement, negative punishment involves the removal of a stimulus. In this case, the goal is to decrease the probability of a behavior, so the stimulus removed is a desired, pleasant, or "positive" stimulus. Punishing a teenager for missing a curfew by taking away use of the car for a period of time is an example of negative punishment. It is the time during which the stimulus is not available that negative punishment gets its other, more common, name of "time out." Remember, though, "negative" refers to the removal of some stimulus, it is not a value judgment of this type of punishment.

Punishment is generally not a very effective means of behavior control, but there are several punishment factors that will modify how effective it is for decreasing behavioral probabilities. As Skinner noticed, every behavior serves some purpose for the organism (i.e. some children misbehave for attention) and if you decrease the likelihood of a behavior, it will appear again unless you shape a new behavior that will have the same purpose for the organism. Punishment doesn't make behaviors disappear; it just reduces the likelihood that they will appear. Another issue with the use of punishment is what happens to the status of the punisher. A dog (or a child for that matter) may come to find the person who continually punishes it as itself aversive, and it will avoid the individual as it comes to associate him or her with punishment.

Sometimes, punishment is necessary. However, in order to be effective at all, the following factors in punishment must be present. Punishment must occur immediately after the behavior, it must be strong, but not overwhelming and it must consistently follow every instance of the behavior to be reduced. How many times did you hear "Just wait until your father comes home," after you were caught misbehaving as a child? While this is meant to scare you, the punishment is still a long way off if it even comes at all. This type of behavior control doesn't work, much to the frustration of many mothers. In order for punishment to decrease the occurrence of behavior, it must occur immediately after the inappropriate behavior. Too much delay makes any future punishment random and not tied close enough to the behavior that needs to be decreased.

Punishment must not only be immediate, but also must be strong in order to be effective. Telling a child "Stop that!" when he/she is caught hitting another child will not be enough to decrease the behavior. However, the child does not need to have his/her toys taken away for a week for the transgression either. The punishment must fit the crime. A "time out" of about 5 minutes and a lecture of why hitting is wrong is usually aversive enough to a young child to greatly decrease the behavior. As Skinner noted, punishment should not be strong enough to cause harm, but it should be strong enough to be aversive.

Another very important issue in the effectiveness of punishment is consistency. As Skinner noted in his research, punishing behavior only occasionally is not an efficient way to decrease the likelihood of behavior. A child who is punished for hitting needs to be punished every time he/she is caught doing it, otherwise the punishment does not work to reduce the occurrence of this behavior.

Operant Conditioning Procedures

B. F. Skinner's (1938) investigations of operant conditioning introduced a variety of unique experimental procedures as well as demonstrations that various processes observed in Pavlov's classical conditioning also have counterparts in operant conditioning. Skinner's operant conditioning procedures introduce alternative manipulations of operant conditioning variables, such as antecedent stimuli and reinforcement contingency rules. These various operant procedures include extinction, generalization, discrimination, shaping, chaining, and a variety of different schedules of reinforcement.

The processes of extinction, generalization and discrimination that were discussed in the classical conditioning section have counterparts in operant conditioning. In extinction, reinforcement that has been a reliable consequence of a behavior is no longer presented. That is, the behavior no longer generates reinforcing consequences. Skinner noticed that when a behavior is first put on extinction, the organism displays a burst of the behavior and then begins to produce new, but related, behaviors -- a phenomenon called response induction. But eventually the behavior decreases in frequency to the point that it is very rarely emitted. If an instance of the behavior is reinforced again, however, spontaneous recovery will occur.

The operant procedure of discrimination training requires a stimulus be presented before the behavior even occurs, leading to its description as an antecedent to behavior. This antecedent stimulus serves to "set the occasion" that any lever press occurring in the presence of this antecedent will be reinforced. Experimentally such a stimulus may be auditory (i.e. a tone) or visual (usually a light of a certain color) or any other modality.

Skinner illustrated discrimination by reinforcing a rat's lever presses in the presence of an antecedent discriminative stimulus (also called an S|_3 or Sd) and not in it's absence (a condition called S- or Sdelta) . Thus behavior in S- is on an extinction schedule in the absence of the discriminative stimulus. Eventually, rats only pressed the lever in the presence of the stimulus, hence completing the discrimination process. Stimulus discrimination is also used in the process of chaining, where one behavior signals that a different behavior will be subsequently reinforced.

Like stimulus discrimination, stimulus generalization in operant conditioning is only slightly different than its counterpart in classical conditioning. Let's say, for example, that a rat's lever press has been reinforced in the presence of a red light but not it the presence of a green light. The rat will come to press only in the presence of the red light, hence demonstrating stimulus discrimination. If a pink or orange light is shown and the rat presses the lever, stimulus generalization has been demonstrated. As an operant conditioning procedure developed by Skinner, stimulus generalization occurs when an organism performs a behavior under antecedent conditions similar to conditions under which it was reinforced.

Schedules of reinforcement involve procedures whereby not every occurrence of a given form of behavior is followed by a reinforcer. Skinner (cf. Ferster |_2 Skinner, 1957) noted that when every instance of a behavior is reinforced, the animal quickly becomes satiated (has enough of the reinforcer that the stimulus loses reinforcing power) and stops engaging in the behavior. To create more steady and long lasting rates of behavior, Skinner would only reinforce a behavior some of the time. This is called a partial, or intermittent, reinforcement schedule (rather than a continuous reinforcement schedule, or CRF) and there are four major types of partial reinforcement procedural rules: fixed ratio (FR), fixed interval (FI), variable ratio (VR) and variable interval (VI). Each procedure calls for presentation of the reinforcement based on either the number of behaviors produced (ratio) or the timing between behaviors (interval). These schedules of reinforcement each have different effects on behavior and we will see (after discussion of other operant procedures) examples of these schedules in everyday situations.

Skinner eventually became dissatisfied with Thorndike's trial and error learning procedures. Skinner felt that by appropriate manipulation of behavioral consequences an experimenter could lead an individual to a correct or desired behavior much more quickly than it would be discovered by chance occurrences. He was thus interested in finding a much more efficient form of learning than trial and error. Skinner described his alternative process as one of shaping a desired, or target, response through reinforcement of successive approximations to the target behavior (Peterson, n.d.).

In shaping, reinforcement is presented for varying successive approximations in forms of behavior as they approximate the eventual behavior to be learned. Step-by-step, the organism comes to engage in behaviors that more and more closely approximate the target behavior. Eventually only the target behavior is the one reinforced. Shaping usually takes much less time than trial and error learning, where an experimenter must wait for the organism to produce the target behavior and subsequently reward it. Related to shaping is a process called chaining. Chaining is used to condition an individual to produce a specific series, or sequence, of different behaviors before the final behavior is reinforced. The chaining process uses discriminative stimuli presented after each step to "link" the chain of behaviors together.

Extinction in Operant Conditioning

Extinction is as much as an operant conditioning procedure as it is aclassical conditioning one. Extinction is sometimes considered a schedule of reinforcement as it is the process of withholding reinforcement for a previously reinforced behavior. Skinner (1938) noticed that this procedure brings about interesting results in and out of the laboratory. When a rat that has been reinforced for lever pressing is put on extinction, two things will occur: bursts of lever pressing and the appearance of new behaviors. The rat will show and increase in response rate immediately after extinction has begun. The rat will then emit new behaviors that may have been infrequent or not recorded. Each of these are dimensions of what is called response induction. As we have seen, the new behaviors that often follow extinction are key to the shaping procedure.

If a lever press that has been put on extinction is reinforced again, it usually only takes one or two reinforcements before lever pressing returns to it's pre-extinction frequency. This occurs even if extinction lasts days or weeks. This phenomenon (the rapid return of lever pressing) is called spontaneous recovery. As in the case of classical conditioning, the existence of spontaneous recovery suggests that, after extinction, behavior is not extinguished, it is somehow suppressed. The lever pressing did not need to be re-shaped; it emerged quickly after extinction. A human example of extinction can be demonstrated when a soda machine does not give a soft drink even after a person has deposited money into it. Usually, you get response burst, (person pushes the button many times and may deposit more money) and the emergence of new behaviors (kicking, swearing, calling the vendor, etc.)

It is important to note that following the extinction of a reinforced behavior an organism will often display an early increase in the rate of that behavior and then the emergence of new behavior. Skinner called this process of increased response rate and variation "response induction" and it is one effect of extinction. Behavior does not instantly stop as soon as extinction is implemented.

As noted, new behaviors often follow the extinction of a reinforced behavior. Skinner capitalized on this phenomenon when he was developing the operant conditioning procedure of shaping (reinforcing successive approximations and then putting them on extinction in order to draw out new behaviors that would more closely approximate a lever press). This phenomenon may also have some survival value, because if new behaviors were not emitted when reinforcement (especially in the form of food or water) no longer follows a particular behavior, an organism would perish if it simply continued producing the same response over and over again.

After the operant procedure of extinction has been implemented for a previously reinforced behavior and the rate of the behavior jumps initially (bursts) due to response induction, response rates then gradually decline to very low rates. If, however, (even after days of extinction) the behavior is reinforced, the response rate jumps back to near pre-extinction rates. This may happen in only one or two reinforcements. This phenomenon is called spontaneous recovery.

Operant Response Shaping and Chaining

Response shaping is an operant procedure developed by B. F. Skinner to bring about new behaviors in an organism (Peterson, n.d.). This procedure is often used in animal training and usually, but not always, involves positive reinforcement. Shaping procedures also include elements of extinction and is a process whereby the form or function of a behavior is developed into a targeted response. Training a rat to press a lever (target behavior) for food in an operant conditioning chamber is a common example of a shaping procedure. A rat generally does not press a bar very often, if at all, when it is first placed into an operant conditioning chamber (also known as a Skinner box). So how do we get it to do so?

Skinner used the ideas of operant conditioning to find an answer this question. Why not reinforce the rat's behaviors that approximate a bar press? Beginning with what the animal does relatively frequently, say looking at, going over to, and even just sniffing the bar (a behavior that occurs often when a rat is placed into an operant chamber), Skinner reinforced each of these to increase their probability. Then, as each became more likely, Skinner changed the rules of reinforcement to include only those behaviors that more closely resembled or actually were bar presses.

It is important to remember that following the extinction of a reinforced behavior an organism will typically increase the probability of that behavior and also engage in a wider variation of that form of behavior, often resulting in the emergence of new, but related, behaviors. Behavior does not instantly disappear as soon as extinction is implemented but rather reflects this typical "burst" in probability and variability as an early effect of extinction. The appearance of new related forms of behaviors is thus another early effect of extinction.

So after the rat consistently emitted one of the "approximate" behaviors, such as first looking at, or later approaching, and even later for sniffing the bar, it was reinforced (usually with food) for doing so. But soon Skinner would no longer reward the behaviors that least approximated actual bar presses, hence initiating extinction for that behavioral approximation. As soon as that behavior was no longer reinforced, the rat engaged in the behavior even more and emitted variations of the behavior. One variation of sniffing a bar, for example, might be rearing up and placing paws on the bar. When this occurred, Skinner began to reinforce this new behavior. When placing paws on the bar reached a fairly high probability, Skinner would then stop reinforcing paws on the bar and the rat would again begin to emit new variations of such behaviors, one of which typically involves actually scratching at and even pressing down on the bar. Skinner would reinforce this and the shaping procedure would be complete. A bar press behavior had been taught through reinforced successive behavioral approximations to a behavior that might begin with a zero probability of ever occurring.

The shaping process, because of its use of alternating use of reinforcement and extinction, is often called differential reinforcement of successive approximations in behavior. Successive approximations refer to the different behaviors that lead, step-by-step, to the target behavior (looking at the bar, approaching the bar, the bar sniff, paws on bar, and finally the bar press in this case). Differential reinforcement refers to the fact that we reinforce these approximations until the behaviors are produced reliably and then reinforcement is withheld so that new and different (hence the word differential) behaviors appear that better approximate the target response to be shaped.

The process of shaping also incorporates the creation and use of secondary reinforcers. If you were to shape a dog to "shake hands", you may not want to have to give it food (a primary reinforcer) every time it emits the correct behavior. By the time shaping is half-completed, the dog may be satiated, and food may not work as a reinforcer anymore. Different schedules of reinforcement may not be appropriate in this case, either. What many people do is say "Good, dog!" right before giving it a treat. Eventually, because of the pairing of the praise and food, the praise takes on reinforcing properties (it increases the probability of behavior). Through this classical conditioning procedure of pairing praise with food, you can reinforce the dog less with food and more with praise (now a conditioned reinforcer) and hence complete the shaping process.

In the case of operant chambers rather than dog training, the delivery of food is typically accomplished by a revolving magazine mechanism, much like those that deliver bubble gum one ball at a time from glass vending machines. The sound of this magazine shifting to deliver, in this case, a food pellet serves as a secondary reinforcer much like the praise example above. This allows for behaviors that take place at quite a distance from the actual food dispenser to be reinforced via secondary reinforcers. The establishment of such secondary, or conditioned, reinforcement functions is often referred to as magazine training and the process involves a conditional stimulus (CS is magazine sound) pairing with an unconditional stimulus (UCS is food) relation which is the same as Pavlov's metronome and food in classical conditioning.

Shaping is not limited to use on animals for simple training. Skinner demonstrated the technique had wide applications with his teaching machine, a device that shaped the skills of human students in correctly answering questions in many subjects. Skinner broke down the complex tasks of learning a new subject into small successive units that gradually built into much more complex systems of knowledge. This technique was called programmed instruction and was the basis for how the teaching machine worked. Skinner's teaching machines served as the prototypes for many modern computer-assisted instructional and training programs.

In order to shape very complex behaviors, as is often seen in animal performance shows, an operant conditioning procedure known as chaining must be implemented. In chaining, one behavior is "bridged" or linked to another by use of a discriminative stimulus that is always associated with the next behavior being reinforced. This process can be used to allow many behaviors to follow one another before reinforcement is actually delivered. In certain animals, the "chains" can be very long while in others they are short and reinforcement must be delivered more often. Eventually the discriminative stimuli that bridge each behavior to the next may be gradually "faded" to generalize the discrimination to the behavioral act itself, thus generating a sequence where one behavior sets the occasion for the next behavior, with the eventual end of the chain of different behaviors being the one reinforced.

If a dog trainer wants a dog to learn to "shake hands," then jump through a hoop and then stand on two feet, begging, that trainer will first shape the begging behavior in the presence of some hand signal, such as "hand raised in air." Once this is reliable, the trainer will present hoops (a second discriminative stimulus) and the dog will only be reinforced when it jumps through hoops and sees a hand raised to signal the begging. Finally, the dog will be shaped to shake, which will bring about the hoops, which signal that jumping and then begging will be reinforced. If the dog doesn't shake, the hoops will not appear, and no reinforcement will be given. Eventually, through this chaining procedure, the dog will shake, jump through the hoop and beg in smooth succession (reinforcement being given after the beg only).

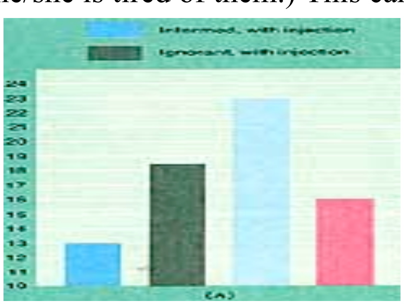
Users of the CyberRat laboratory simulation may wish to read a step-by-step description of how best to shape an animal with no prior experimental history. There is such a collection of topics available in the Appendix. These include:

Shaping a New Behavior.
Getting Ready for Shaping.
Understanding the Experimental Chamber.
Getting Your Subject Ready for Shaping: Habituation
Getting Your Subject Ready for Shaping: Magazine Training
Getting Your Subject Ready for Shaping: Observe Behavior Carefully
Begin Shaping (If Operant Level is Low)
Shaping: Not Too Slow, Not Too Fast
Other Factors Involved in Creating New Behavior: Prompting
Other Factors Involved in Creating New Behavior: Discrimination
Factors Involved in Creating New Behavior: Intermittent Reinforcement

Schedules of Reinforcement

One group of procedures Skinner developed in his work on operant conditioning is that involving reinforcement schedules (Ferster |_2 Skinner, 1957). Schedules of reinforcement are simple rules for when reinforcement should be given following a specific behavior. The two main schedule rules are continuous and partial reinforcement. Another word for partial reinforcement is intermittent (less than continuous) reinforcement. The most common intermittent reinforcement rules include four specific types of schedules: fixed ratio , variable ratio, fixed interval and variable interval. Skinner observed that these different schedules have different effects on rates of responding, each of which will be illustrated by the graphics that accompany our more detailed descriptions of each schedule in this or subsequent topical discussions.

In a continuous reinforcement schedule every occurrence of a behavior is reinforced. If a rat is on a continuous schedule of reinforcement (often abbreviated as CRF) for lever pressing, every lever press is reinforced. A child, for example, who gets some dessert every time he or she finishes dinner is on a continuous schedule of reinforcement. As Skinner noted, this schedule produces a relatively moderate and steady rate of responding until the organism becomes satiated (an animal gets so much food as reinforcement that it is no longer hungry or the child has received desserts so often, he/she is tired of them.) This can occur relatively quickly, depending on the size of the reinforcer, and thus is not an efficient means for maintaining a steady rate of responding over sustained periods of time. The accompanying figure is a simulated graphic illustrating both the relatively steady rate of responding and the slowing or elimination effects of satiation to the reinforcer within a single session.



Both to avoid having to use so much food and to counteract the satiation effects of continuous reinforcement, Skinner used intermittent schedules of reinforcement (Ferster |_2 Skinner, 1957). In intermittent schedules of reinforcement, only certain occurrences of a class of behaviors are reinforced. Sometimes the rule defining which behavioral occurrence should be reinforced is based on time elapsed plus the required response. Thus in what are called the interval schedules a predetermined amount of time must go by before reinforcement is delivered for the first response occurring after the interval of non-reinforcement for responding. Such interval schedules exist as either fixed or variable interval schedules. That is, the amount of time that reinforcement is not delivered for any behaviors is either the same interval following actual reinforcement, or time intervals are randomized durations around some average interval length.

Alternatively, delivery of reinforcement may be based on the number of times a specific class of behavior occurs (called ratio schedules because a particular type of response must occur a certain number of times before reinforcement is given). Such rules include fixed ratio schedules, where the required number of responses stay the same from one reinforced behavior to another, and variable ratio schedules where the number required between reinforcement delivery is some random number around a specific "average" of responses, such a an average of 10-to-1 or 20-to-1 (that is, on average one of 10 or 20 responses will be reinforced, but will randomly vary from 1 to any number, so long as in the long-term, the average of 10 or 20 is maintained).

In laboratory studies using either rats or pigeons, Skinner (Skinner |_2 Ferster, 1957) found that the rates of behavior are different for the various partial schedules of reinforcement and that the schedule chosen is often a function of what type of responding a researcher, or employer for that matter, might desire. Both the interval schedules of reinforcement and the ratio schedules of reinforcement and how they effect the rate of responding effects each type of schedule are covered in more detail in those respective topical discussions.

While conducting research on schedules of reinforcement as variations in operant conditioning procedures, Skinner noticed an interesting phenomenon surrounding the use of partial reinforcement. When a pigeon on continuous reinforcement is subsequently put on extinction (no reinforcement is delivered), the animal emits a burst of responses at first, but then gradually stops responding. In contrast, a pigeon that has been gradually moved to a partial schedule of reinforcement (especially if it is "lean" meaning reinforced rarely in the face of producing lots of responses) will continue responding for a very long time when moved to extinction; often taking multiple sessions before slowing down after extinction is started. This resistance to extinction follows any type of partial reinforcement schedule as long as the schedule is brought on gradually and is a relatively lean schedule, This resistance to extinction phenomenon is thus one of the primary partial reinforcement effects.

Ratio Schedules of Reinforcement

Skinner's research on operant conditioning procedures eventually led him to investigate intermittent, as opposed to continuous, reinforcement schedules (Ferster |_2 Skinner, 1957). Intermittent schedules of reinforcement are simple rules for delivering single reinforcements for multiple occurrences of a specific type of behavior, such as lever pressing. Skinner's original investigations used continuous reinforcement, where each and every lever press was reinforced. But in subsequent research he began to investigate what would happen if not every lever press was reinforced, a practice known as applying rules of intermittent (less than continuous) reinforcement.

One of the simplest, and thus most common intermittent reinforcement rules involves using some ratio of some number of required lever presses for each delivery of one reinforcement. Such ratios may use a "fixed" number, such as FR-10 where every 10th response would be reinforced, or a variable number, such as VR-10, where any constantly varying and random number of responses is used as the criterion for delivering reinforcement for the criterion lever press, so long as a large sample of these ratios average the reference ratio number (in our case, 10).

A special case of the ratio schedule, known as CRF or FR-1, is actually a continuous reinforcement schedule where every occurrence of a behavior is reinforced. If a rat is on a continuous schedule of reinforcement for lever pressing, every lever press is reinforced. A child, for example, who gets some dessert every time he or she finishes dinner is on a continuous schedule of reinforcement. Skinner's earliest work (Skinner, 1938) investigated this schedule almost exclusively, and he observed this schedule to produce a relatively moderate and steady rate of responding until the organism becomes satiated (an animal gets so much food as reinforcement that it is no longer hungry or the child has received desserts so often he/she is tired of them.) This can occur relatively quickly, depending on the size of the reinforcer, and thus is not an efficient means for maintaining a steady rate of responding over sustained periods of time. The accompanying figure is a simulated graphic illustrating both the relatively steady rate of responding and the slowing or elimination effects of satiation to the reinforcer within a single session.

To avoid using so many food pellets, which in his early research he had to hand-manufacture, Skinner eventually investigated intermittent schedules of reinforcement (Ferster |_2 Skinner, 1957). In intermittent schedules of reinforcement, only certain occurrences of a class of behaviors are reinforced. Sometimes the rule defining which behavioral occurrence should be reinforced is based on some interval of time elapsed plus the required response, thus generating what is known as the interval schedules of reinforcement.

Alternatively, delivery of reinforcement may be based on the number of times a specific class of behavior occurs (a particular type of response, such as a lever press or a key peck, must occur a certain number of times before reinforcement is given). Such rules include fixed ratio schedules, where the required number of responses stay the same from one reinforced behavior to another, and variable ratio schedules where the number required between reinforcement delivery is some random number around a specific average number of responses, such a an average of 10-to-1 or 20-to-1 (that is, on average one of 10 or 20 responses will be reinforced, but will randomly vary from 1 to any number, so long as in the long-term, the average of 10 or 20 is maintained).

In laboratory studies using either rats or pigeons, Skinner (Skinner |_2 Ferster, 1957) found that the rates of behavior are different for the various partial schedules of reinforcement and that the schedule chosen is often a function of what type of responding a researcher, or employer for that matter, might desire. A rat which gets reinforced every twentieth lever press is operating under a fixed ratio schedule of reinforcement (actually, an FR-20). As Skinner's research illustrated, this version of partial reinforcement produces very steady rates of responding, but only after a brief break after the reinforcement is delivered-a pattern often referred to as break run (see the accompanying illustration). Factory workers who get a certain amount of money for, say, every 5th completed unit of product are working under a fixed ration schedule of reinforcement. In the workplace fixed ratio schedules are known as piecework schedules wherein pay is based on a fixed number of components produced (an example of measuring behavior by its effects on environment, which is the defining feature of operant behavior -- it "operates upon environments" to produce some effect -- in this case, a part of some sort). In his laboratory, Skinner discovered that this schedule produced a fairly predictable brief break followed by a steady rate of subsequent responding. Employers often use piecework schedules because they usually result in relatively high productivity.

The variable ratio schedule is also a partial schedule of reinforcement. A variable ratio schedule of reinforcement, like a fixed ratio, involves the delivery of reinforcement based on the number of behavior occurrences. In a variable schedule, however, it is an average number, not a fixed number, of responses that are reinforced. A rat on this schedule may get reinforced, on average, for every ten responses. But because it is an average, reinforcement may come after two responses or after twenty. Reinforcement is not delivered every ten responses, although there may be a time when a tenth response is reinforced. Skinner noted that this is a very powerful schedule and it produces very high and quite constant rates of responding.

Gambling on slot machines is a clear example of a variable ratio schedule of reinforcement. For instance, on average, every 50th hit on the machine will being about a jackpot (reinforcement). This means that the jackpot could occur in two hits or two hundred hits. Gamblers find this possibility (two hits until jackpot) irresistible and many develop gambling problems because this type of reinforcement schedule produces high and consistent rates of responding. Many individuals continue gambling despite increasing debt because the indeterminate predictability of reinforcement and also because of the unique resistance to extinction caused by intermittent reinforcement.

As with interval schedules of reinforcement, a pigeon or a person that has been gradually moved to a ratio schedule of intermittent reinforcement (especially if it is a lean schedule, meaning reinforced on average or on every, say, 50th response rather than, say, every 5th response) will continue responding for a very long time when moved to extinction; often taking multiple sessions before slowing down after extinction is started. This resistance to extinction follows any type of partial reinforcement schedule as long as the schedule is brought on gradually and is relatively lean. This phenomenon is thus called the resistance to extinction effect of intermittent schedules of reinforcement.

Interval Schedules of Reinforcement

One group of procedures Skinner developed in his work on operant conditioning is that involving reinforcement schedules (Ferster |_2 Skinner, 1957), or simple rules for which occurrences of a behavior will be reinforced. If reinforcement is not continuous --

delivered for every occurrence of a type of behavior -- then a partial, or intermittent, reinforcement schedule is in effect. Common rules for scheduling intermittent reinforcement include fixed ratio, variable ratio, fixed interval and variable interval. In this section we will consider the unique response rate patterns generated by either fixed or variable interval schedules of reinforcement.

In interval schedules of reinforcement, the amount of time that reinforcement is not delivered for any behaviors is either the same duration following a reinforcement (fixed interval schedules), or time intervals are randomized durations around some average interval length (variable interval schedules, which are sometimes also called random interval schedules). If, as a researcher, you were to reinforce a rat for the first lever press that occurred after one minute following a previous reinforcement, you would be using a fixed interval schedule of reinforcement called an FI-60 (for the 60 seconds that elapses after the last reinforcement before the next response will generate another reinforcer). As a partial schedule, only the lever press that follows a one-minute interval is reinforced, and then a new one-minute interval is reset. Any lever press occurring during this one minute interval fails to bring reinforcement.

In laboratory studies using either rats or pigeons, Skinner (Skinner |_2 Ferster, 1957) found that the rates of behavior generated by interval schedules of reinforcement are uniquely but predictably variable across time, thereby creating a predictable "pattern" in the rate of responding for fixed vs variable interval schedules. A fixed interval schedule, as can be seen in the accompanying graphic illustration, produces one of the more unique response patterns. This pattern is typically described as a scalloping change in the rate of responding. What happens is that immediately after being reinforced, the rat stops responding. But as time passes, responding begins, at first slowly, and then the rate increases until it is very high near the end of the minute. This means that when the interval expires, the animal is very likely to respond, and the first such response is reinforced. As you can see, because of the scalloping effect, this schedule is not efficient at producing steady rates of responding.

Many teachers give weekly exams to test what students have learned in a given course. The opportunity to be reinforced for studying, by receiving an "A", is on a fixed (because it is every week) interval schedule of reinforcement. As Skinner noted in the laboratory (and as many parents and teachers know) the rate of responding under this schedule is scalloped. This means that right after a quiz, responding (studying) drops to nothing. Responding then slowly increases three nights away from the next quiz, until students are "cramming" the night before. This schedule is ineffective at maintaining a steady rate of responding (studying) both in and out of the laboratory.

The variable interval schedule uses rules similar to the fixed interval schedules, except the duration of the interval constantly changes in a random fashion around some average. Thus a VI-60 schedule would result in an average non-reinforcement availability period of 60 seconds, but each specific period could be any duration so long as a large sample of these intervals results in an average of 60 seconds. So under a variable, or random, interval schedule, a rat would be reinforced, on average, for the first response occurring after one minute from the last response. This means that the rat may receive reinforcement after, say, twenty seconds between behaviors or after four minutes between behaviors.

Has a teacher/professor ever given pop quizzes in a course you have taken? If he/she did, they would be using a variable interval schedule of reinforcement. What was the effect on the way members of the class prepared for the exam? As demonstrated by the accompanying graphic illustration, Skinner observed that the rate of responding under a variable interval schedule, although not as high in ratio schedules, is very steady and consistent.

So your instructor who gives pop quizzes is using a variable interval schedule of reinforcement in an effort to maintain high and consistent rates of studying. Perhaps the quizzes are given weekly, on average, but a subsequent quiz could come, one or two days or even two weeks after the last quiz. Since a student has no idea when the next quiz will come, studying is much more consistent than with weekly (fixed interval) exams.

While conducting research on schedules of reinforcement in operant conditioning, Skinner noticed an interesting phenomenon accompanying the use of partial reinforcement. When a pigeon on continuous reinforcement is subsequently put on extinction (reinforcement is no longer delivered for responding a specific way), the animal emits a burst of responses at first (a phenomenon known as response induction), but then gradually stops responding. In contrast, a pigeon that has been gradually moved to a partial schedule of reinforcement (especially if it is "lean," meaning, for example, the animal is reinforced for the first response following on average an interval of 180 seconds) will continue responding for a very long time when moved to extinction. In fact, it often takes multiple sessions before responding begins slowing down after extinction has started. This is known as resistance to extinction and it follows any type of partial reinforcement schedule as long as the schedule is brought on gradually and is relatively lean. Resistance to extinction explains why very lean schedules of reinforcement result in behaviors that are very persistent as well as occurring at a high rate. So both high rates of responding and resistance to extinction are often known as partial reinforcement effects.

Applications of Operant Conditioning

B. F. Skinner's operant conditioning principles have been applied in many areas. In education operant procedures have been used to develop programmed instruction, a teaching technique based on elements of shaping and chaining (Skinner, 1968). Programmed instruction was used extensively in Skinner's early development of teaching machines, which today have been replaced by computers. In fact, the text you are reading right now incorporates mainly operant principles in its approach to computer-assisted instructional design. Its reliance on adaptive adjustments to changing learner skills is a direct application of shaping procedures applied to higher-level reading comprehension and learning skills (Ray, 2004).

Operant principles are also used in various therapies. Miller introduced operant procedures as the defining technique in biofeedback (Miller, 1969); a therapy designed to reduce stress and physical reactions to stress. Alternatively, behavior modification for misbehaving children is often implemented in the home as well as in the school. Through secondary reinforcement offered through token economies (Ayllon |_2 Azrin, 1968) a client's good behavior may be positively reinforced and inappropriate behaviors ignored.

Animal trainers use operant conditioning procedures when training performance animals. Circuses, marine parks and zoos use shaping, chaining, and other procedures in order to teach animals to perform both for display and entertainment as well as for routine animal care, such as presenting a leg for drawing a blood sample. People also use operant procedures, sometimes unwittingly, when training their own pets (Pryor, 1985). Giving a dog a treat when it "shakes hands" is using the principle of positive reinforcement.

Programmed Instruction

Whether teachers recognize it or not, operant conditioning principles are often incorporated into the classroom environment (Skinner, 1968). In traditional teaching, instructors lecture and students take notes. Or perhaps students work either independently or in small groups on what has been previously presented by the teacher. But because there is only one teacher in a classroom, students do not get immediate feedback as to the accuracy of their work. Skinner's approach to this problem was to develop what is called programmed instruction.

Skinner also designed an apparatus called a teaching machine (Skinner, 1989) as a key element in the delivery of programmed instruction. The machine was a box with a window and a scrolling knob. The student was presented with some introductory material and then questions about that material for the student to answer. The student then scrolled the knob to reveal the answer to confirm whether they had learned the material presented. The student was to continue this process until he/she reached mastery of a series of such programmed sequences, or "frames" of material.

As the student progressed, the machine could be fitted to present more difficult material. Alternatively, it could also be taken down a level of difficulty if extra review was necessary. Through reinforcement (feedback indicating that the student answered correctly) and successive approximations (increasing difficulty of materials) students are shaped and taught mastery in any subject the instructor uses the machine to teach.

Does this sound familiar? It should, the operations in the teaching machine are the forerunners of computer aided instruction, and are the very basis for how the tutoring component of the software you are currently using works (Ray, 2004). Teaching machines never really caught on in mainstream classrooms, mostly because people feared that they were impersonal and lacked the warmth of teachers. Nevertheless, many students progressed far more rapidly and with far fewer errors in their learning using such teaching designs. Thus programmed instruction has evolved into the increasingly popular computer assisted instruction of today and research has found such instruction to be highly effective.

As noted, Skinner's teaching machine used the procedure of shaping, or successive approximations, to assist students in their learning process. Successive approximations to a final learning goal is often the foundation in programmed instruction. But not all forms of computer assisted instruction are based on this operant principle. Most of computerized teaching programs may look similar to programmed instruction, in that they ask students questions after readings and then give immediate feedback as to the accuracy of answers. But few such programs require that the student progress toward less and less dependency upon the programmed strategy. This is a shortcoming of all but the most sophisticated of programs that incorporate what is called "adaptive instruction" designs (Ray, 2004). That is, the learning goals are changed to adapt to the individual learner's changing skills and knowledge.

All computer-assisted instruction allows the teacher to spend time with students who are having difficulty while allowing more advanced students to continue and excel with immediate feedback. But few are designed to give and then fade supportive prompts and to present successively more difficult questions as the adaptive instructional software you are currently using does when used in "Tutor" mode of presentation.

Thus the computer assisted instruction material you are currently studying uses principles and procedures based originally on Skinner's programmed instruction (Ray, 2004). The tutor mode of the MediaMatrix software program turns your computer into a more modern and sophisticated version of Skinner's teaching machine. When you are in the tutor mode, the system helps prompt you as to the most important concepts and properties of those concepts, then asks you questions at the end of each segment of material presented. The system begins with the highest density of prompting, the smallest frame of content, and the easiest form of question, multiple choice. As you progress, the prompts are gradually faded, the unit or frame of content presented gets larger, and the questions become more difficult if you answer a series of questions accurately.

The MediaMatrix adaptive instructional system gradually moves from multiple choice questions to the less prompted fill in the blank, association recognition and, finally, minimally prompted verbal associates questions as you become more proficient in learning the material with less help and greater accuracy. If you begin to have difficulty with specific content or at a current level of difficulty, the program will successively drop levels until you are succeeding again. Because all lower level (multiple choice and fill-blank) questions are also represented in the association form of questioning, the system shapes the user into being able to answer accurately the more challenging association questions that depend upon total recall, as opposed to mere recognition, of the material.

This programmed instructional format, called adaptive instruction, relies upon artificial intelligence to compare the students' growing verbal or semantic networks of terms to an expert's network to adjust all of its varieties of presentations, including which questions you are asked. Such adaptive programmed instruction is designed to eventually wean the student from the need for programmed formats, thus teaching the student how to learn more traditionally presented materials through a shaping of that reading comprehension skill.

Such computer-based adaptive instruction can assist instructors and students alike (Ray, 2004). Students who have tutored on materials assigned prior to a class that intends to cover much the same topic of material find they are well prepared in the fundamentals, which allows the instructor to take more time teaching other dimensions, such as ethics, applications, or research foundations, and less time on simple definitional, conceptual, and review of fundamentals. The goal is to create prepared learners and to allow everyone to be at much the same level of understanding when the class begins. These were the goals of Skinner when he first designed programmed instruction and teaching machines. But it took the development of modern personal computers and sophisticated software development to achieve the real aspirations of Skinner's inventions based on the application of operant principles in and out of the classroom.

Therapeutic Applications of Operant Conditioning

Skinner's operant conditioning principles also are the foundation of various therapeutic applications (Skinner, 1972). Behavior modification, or the process of changing responses through stimulus control and token reinforcement economies, is one such application. Token economies (Ayllon |_2 Azrin, 1968) work to reinforce positive behaviors while simultaneously placing inappropriate behaviors on extinction. In psychiatric institutions token economies help maintain appropriate behaviors by reinforcing those behaviors directly. Likewise, disruptive classroom behaviors can be reduced through extinction and the reinforcement of more appropriate behaviors as well (Swiezy, Matson, |_2 Box, 1992). Even maladaptive physiological responses, such as anxiety or migraine headaches, can be addressed by operant techniques that use additional feedback, known as biofeedback (Miller, 1969), to help an individual know the state of normally unconscious bodily processes. It is thus worth considering each of these types of therapeutic applications of operant conditioning in a bit more detail.

Behavior modification is an operant approach to overt behavioral therapy and education (Bellack, Hersen, |_2 Kazdin, 1982).. The responses targeted for change are usually maladaptive for the individual and/or are inappropriate is given situations. For example behavior modification processes are often used in schools to help children whose behaviors have become disruptive and harmful to themselves or others. Behavior modification is also used in psychiatric institutions or institutions for the severely mentally challenged. The goal of behavior modification is to teach appropriate behaviors that serve the same function as maladaptive or absent behaviors (i.e. functions such as getting attention, help, praise, food, relief from boredom, etc.) This is usually accomplished with token economies being the specific type of consequential stimulus (reinforcement)

Token economies rely upon conditioned reinforcers, such as poker chips, points, or stars on a chart, used in a rule-guided process. A token economy utilizes both positive reinforcement and secondary reinforcers. Behavior modification is accomplished through the use of tokens to reward certain behaviors that occur during certain situations (stimulus discrimination) and thus represent a popular method of both antecedent and consequential stimulus control (Ayllon |_2 Azrin, 1968). Once an individual has accumulated a certain amount of these tokens (conditioned reinforcers), the tokens can be traded in for more direct and tangible reinforcers in the form of toys, favorite snacks, time with a computer game or anything that can be presented that will increase the probability of a behavior (positive reinforcement).

For example, consider a first grade classroom where playing with blocks is not a maladaptive or harmful behavior unless it occurs during the class time when a teacher is trying to instruct her class in preparation for some activity. Some children may start or continue playing with a toy during a time or circumstance such as this when it is inappropriate to do so. Instead of punishing the child to eliminate such undesirable behaviors, token economies can be used to create more desirable alternatives. For example, the child can receive a gold star every time they pay attention when it is appropriate in class. Then, at recess they can play freely and safely for even more gold stars. After so many stars, the child may then receive a favorite snack or time with a favorite game or computers. If the right reinforcement is used for the right behavior in the right setting, the child will begin to play with blocks only during play time and to pay attention during the class time that requires attending. This application of Skinner's operant procedures has been found to be very effective across many different situations, both educational and institutional.

Stimulus control in behavior modification may refer not only to the process of reinforcement, but also the process of controlling for antecedent stimulus discriminations in such a way that maladaptive behaviors become more appropriate and acceptable responses in specific situations. We cannot completely rid a person (or any organism) of a particular behavior. It may reappear anytime circumstances permit. But by controlling reinforcement and the discriminative settings where behaviors are appropriate, it is possible to create environments where maladaptive behaviors have no function and where new and more acceptable behaviors do. Eventually, acceptable responses replace those that are inappropriate. This is the essence of antecedent and consequential stimulus control, and it represents an application of Skinner's operant procedures.

Quite a different form of application of operant conditioning was developed largely by Neal Miller (1969) and is called biofeedback. Biofeedback is an operant approach to therapy that uses visual and/or auditory signals to reflect some internal state of the patient -- states that he or she would otherwise not be aware of. These signals, or feedback, serve as positive reinforcement when they indicate that the individual has successfully changed his or her internal responses in some target direction or amount. For example, someone who gets anxious in crowds may wear a heart rate monitor in a crowded situation. They then may read the monitor and use relaxation techniques to keep their heart rate under a certain level. The same can occur with high blood pressure. An individual can wear a blood pressure monitor at work and learn to keep it under a certain level by relaxing when a stressful situation presents itself. Since successful readings serve as positive reinforcers, people learn to relax in anxiety or stress provoking situations. The research on the effectiveness of Miller's biofeedback therapy is mixed, but it has been shown to be useful under various conditions, including control of migraine headaches (c.f., Sturgis, Tollison, |_2 Adams, 1978).

Operant Procedures in Animal Training

Response shaping is an operant procedure developed by B. F. Skinner to bring about new behaviors in an organism (Peterson, n.d.). This procedure is often used in animal training and usually, but not always, involves positive reinforcement (Skinner, 1951). Shaping procedures also include elements of extinction and is a process whereby the form or function of a behavior is gradually developed into the desired (target) response. Training a rat to press a bar (the target behavior) for food in an operant chamber is a common example of a shaping procedure. A rat generally does not press a bar very often, if at all, when it is first placed into an operant conditioning chamber (also known as a Skinner box). So how do we get it to do so?

Skinner used the processes of operant conditioning to find an answer this question. Why not begin by reinforcing the rat's behaviors that approximate a bar press, even if they are remote from actual bar presses, and then gradually shift the criteria for reinforcement to only those behaviors that more closely resemble bar pressing? Beginning with what the animal does relatively frequently, say looking at, going over to, and even just sniffing the bar (a behavior that occurs often when a rat is placed into an operant chamber), Skinner reinforced each of these to increase their probability. Then, as each of these behaviors became more likely, Skinner changed the rules of reinforcement to include only those behaviors that more closely resembled or actually were bar presses.

It is important to remember that following the extinction of a reinforced behavior an organism will typically increase the probability of that behavior and also engage in a wider variation of that form of behavior, often resulting in the emergence of new, but related, behaviors. Behavior does not instantly disappear as soon as extinction is implemented but rather reflects response induction, which is an increase in probability and variability as an early effect of extinction. The appearance of new, but somewhat similar or related forms of behaviors is thus another early effect of extinction.

So after the rat consistently emitted one of the "approximate" behaviors, such as first looking at, or later approaching, and even later for sniffing the bar, it was reinforced (usually with food) for doing so. But soon Skinner would no longer reward the behaviors that least approximated actual bar presses, hence initiating extinction for that behavioral approximation. As soon as that behavior was no longer reinforced, the rat engaged in response induction by emitting the behavior even more frequently and engaging in variations on that behavior.

One variation of sniffing a bar, for example, might be rearing up and placing paws on the bar. When this occurred, Skinner reinforced this new behavior. When placing paws on the bar reached a fairly high probability, Skinner would then stop reinforcing paws on the bar and the rat would again begin to emit new variations of such behaviors, one of which typically involves actually scratching at and even pressing down on the bar. Skinner would reinforce this and the shaping procedure would be complete. A bar press behavior had been taught through reinforced successive behavioral approximations to a behavior that might begin with a zero probability of ever occurring.

The response shaping process, because of its use of alternating use of reinforcement and extinction, is often called differential reinforcement of successive approximations in behavior. Successive approximations refer to the different behaviors that lead, step-by-step, to the target behavior (steps such as looking at the bar, then approaching the bar, then bar sniffs, paws on bar, and finally the bar press in our example). Differential reinforcement refers to the fact that we, at first, will reinforce any or all of these variations until one of the behaviors is produced reliably and then reinforcement is withheld so that new and different (hence the word differential) behaviors appear that more closely approximate the target response being shaped.

The process of shaping also incorporates the creation and use of secondary reinforcers. If you were to shape a dog to "shake hands", you may not want to have to give it food (a primary reinforcer) every time it emits the correct behavior. By the time shaping is half-completed, the dog may be satiated, and food may not work as a reinforcer anymore. Different schedules of reinforcement may not be appropriate in this case, either. What many people do is to use a child's toy "cricket" to produce a click, or to say "Good, dog!" right before giving it a treat (Pryor, 1985). Eventually, because of the pairing of the click or praise and food, the sound takes on reinforcing properties (it increases the probability of behavior). This is a classical conditioning, or stimulus contingency, procedure involving the pairing of the previously neutral sound (NS/CS) of the clicker or praise with food (UCS). This allows you to be able to reinforce the dog less with food and more with clicks or praise (now conditioned reinforcers) and hence to complete the shaping process without the animal becoming satiated on food.

In the case of operant chambers rather than dog training, the delivery of food is typically accomplished by a revolving magazine mechanism, much like those that deliver bubble gum one ball at a time from glass ball vending machines. The sound of this magazine shifting to deliver, in the rat's case, a food pellet serves as a secondary reinforcer much like the clicker or praise example above. This allows for behaviors that take place at quite a distance from the actual food dispenser to be reinforced via the secondary or conditioned reinforcer of the sound. The establishment of such secondary conditioned reinforcement functions is often referred to as magazine training.

There is a highly sophisticated computer simulation program available (at www.cyberrat.net) called CyberRat, that allows students who do not have access to live animal laboratories to experience both magazine training and shaping dynamics via simulation. This simulator uses an extremely large array (over 1800) of very brief digital video clips of live animals in a traditional operant chamber to produce the highly realistic illusion of a seamless real-time video feed showing a live animal being placed into an operant conditioning chamber and behaving exactly as real animals behave in this environment.

Through artificially intelligent algorithms the sequences of these clips may be altered through a student's reinforcement button by delivering simulated "water reinforcements" to any selected individual animal for its successive approximations to bar pressing. If you reward the animal appropriately, the video clips alter their sequences to simulate actual changes in behavior that simulate the entire operant response shaping process with highly realistic results. If you have access to CyberRat, visit the Appendix Elaboration on Shaping linked here for more details on how you can learn to shape a laboratory rat just as Skinner did! To access any or all of the Appendix Topics on how to shape a naive rat in CyberRat, click on the topic of interest:

Of course, shaping techniques are not limited to use on animals for simple training. Skinner (1989) demonstrated that the technique has wide applications with his teaching machine, a device that shaped the skills of human students in correctly answering questions in many subjects. Skinner broke down the complex tasks of learning a new subject into small successive units that gradually built into much more complex systems of knowledge. This technique was called programmed instruction and was the basis for how the teaching machine worked. Skinner's teaching machines served as the prototypes for many modern computer-assisted instructional and training programs.