

03 - 2.3 Learning Theory

2.3 Learning Theory

effects. Patients whose pathology stems from exaggerated early attachments may attempt to replicate them in therapy. Therapists must enable such patients to recognize the ways their early experiences have interfered with their ability to achieve independence. For patients who are children and whose attachment difficulties may be more apparent than those of adults, therapists represent consistent and trusted figures who can engender a sense of warmth and self-esteem in children, often for the first time. Relationship Disorders A person's psychological health and sense of well-being depend significantly on the quality of his or her relationships and attachment to others, and a core issue in all close personal relationships is establishing and regulating that connection. In a typical attachment interaction, one person seeks more proximity and affection, and the other either reciprocates, rejects, or disqualifies the request. A pattern is shaped through repeated exchanges. Distinct attachment styles have been observed. Adults with an anxious-ambivalent attachment style tend to be obsessed with romantic partners, suffer from extreme jealousy, and have a high divorce rate. Persons with an avoidant attachment style are relatively uninvested in close relationships, although they often feel lonely. They seem afraid of intimacy and tend to withdraw when there is stress or conflict in the relationship. Break-up rates are high. Persons with a secure attachment style are highly invested in relationships and tend to behave without much possessiveness or fear of rejection. REFERENCES Freud S. The Standard Edition of the Complete Psychological Works of Sigmund Freud. 24 vols. London: Hogarth Press; 1953- Greenberg JR, Mitchell SA. Object Relations in Psychoanalytic Theory. Cambridge, MA: Harvard University Press; 1983. Laplanche J, Pontalis J-B. The Language of Psycho-analysis. New York: Norton; 1973. Mahler MS, Pine F, Bergman A. The Psychological Birth of the Human Infant. New York: Basic Books; 1975. Pallini S, Baiocco R, Schneider BH, Madigan S, Atkinson L. Early child-parent attachment and peer relations: A metaanalysis of recent research. J Fam Psychol. 2014;28:118. Stern D. The Interpersonal World of the Infant. New York: Basic Books; 1985. Meissner, W.W. Theories of Personality in Psychotherapy. In: Sadock BJ, Sadock VA, eds. Kaplan & Sadock's Comprehensive Textbook of Psychiatry. 9th ed. Vol. 1. Philadelphia: Lippincott Williams & Wilkins; 2009:788. 2.3 Learning Theory Learning is defined as a change in behavior resulting from repeated practice. The principles of learning are always operating and always influencing human activity. Learning principles are often deeply involved in the etiology and maintenance of psychiatric disorders because so much of human behavior (including overt behavior, thought patterns, and emotion) is acquired through learning. Learning processes also

strongly influence psychotherapy because human behavior changes. Thus, learning principles can influence the effectiveness of therapy. In fact, no method of therapy can be said to be immune to the effects of learning. Even the simple prescription of a medication can bring learning processes

into play because the patient will have opportunities to learn about the drug's effects and side effects, will need to learn to comply with the instructions and directions for taking it, and will need to learn to overcome any resistance to compliance. BASIC CONCEPTS AND CONSIDERATIONS A great deal of modern research on learning still focuses on Pavlovian (classical) and operant learning. Pavlovian conditioning, developed by Ivan Petrovich Pavlov (1849– 1936), occurs when neutral stimuli are associated with a psychologically significant event. The main result is that the stimuli come to evoke a set of responses or emotions that may contribute to many clinical disorders, including (but not limited to) anxiety disorders and drug dependence. The events in Pavlov's experiment are often described using terms designed to make the experiment applicable to any situation. The food is the unconditional stimulus (US) because it unconditionally elicits salivation before the experiment begins. The bell is known as the conditional stimulus (CS) because it only elicits the salivary response conditional on the bell-food pairings. The new response to the bell is correspondingly called the conditional response (CR), and the natural response to the food itself is the unconditional response (UR). Modern laboratory studies of conditioning use a very wide range of CSs and USs and measure a wide range of conditioned responses. Operant conditioning, developed by B.F. Skinner (1904–1990), occurs when a behavior (instead of a stimulus) is associated with a psychologically significant event. In the laboratory, the most famous experimental arrangement is the one in which a rat presses a lever to earn food pellets. In this case, as opposed to Pavlov's, the behavior is said to be an operant because it operates on the environment. The food pellet is a reinforcer—an event that increases the strength of the behavior of which it is made a consequence. A major idea behind this method is that the rat's behavior is "voluntary" in the sense that the animal is not compelled to make the response (it can perform it whenever it "wants" to). In this sense, it is similar to the thousands of operant behaviors that humans choose to commit—freely—in any day. Of course, the even larger idea is that even though the rat's behavior appears as though it is voluntary, it is lawfully controlled by its consequences: If the experimenter were to stop delivering the food pellet, the rat would stop pressing the lever, and if the experimenter were to allow the lever press to produce larger pellets, or perhaps pellets at a higher probability or rate, then the rate of the behavior might increase. The point of operant conditioning experiments, then, is largely to understand the relation of behavior to its payoff. Pavlovian and operant conditioning differ in several ways. One of the most fundamental differences is that the responses observed in Pavlov's experiment are elicited and thus controlled by the presentation of an antecedent stimulus. In contrast,

the "response" observed in Skinner's experiment is not elicited or compelled by an antecedent stimulus in any obvious way—it is instead controlled by its consequences. This distinction between operants and respondents is important in clinical settings. If a young patient is referred to the clinic for acting out in the classroom, an initial goal of the clinician will be to determine whether the behavior is a respondent or an operant, and then the clinician will go about changing either its antecedents or its consequences, respectively, to reduce its probability of occurrence. Despite the academic separation of operant and respondent conditioning, they have an important common function: Both learning processes are designed by evolution to allow organisms to adapt to the environment. The idea is illustrated by considering the law of effect (Fig. 2.3-1), which says that whether an operant behavior increases or decreases in strength depends on the effect it has on the environment. When the action leads to a positive outcome, the action is strengthened; conversely, when the action leads to a negative outcome, we have punishment, and the action is weakened. In a similar manner, when an action decreases the probability of a positive event, behavior also

declines. (Such a procedure is now widely known as time-out from reinforcement.) When an action terminates or prevents the occurrence of a negative event, the behavior will strengthen. By thus enabling the organism to maximize its interaction with positive events and minimize its interaction with negative ones, operant conditioning allows the organism to optimize its interaction with the environment. Of course, events that were once positive in the human's earlier evolutionary history are so prevalent in modern society that they do not always seem adaptive today. Thus, reward learning also provides a framework for understanding the development of rather maladaptive behaviors like overeating (in which behavior is reinforced by food) and drug taking (in which behaviors are reinforced by the pharmacological effects of drugs)—cases in which reward principles lead to psychopathology. FIGURE 2.3-1 The law of effect in instrumental/operant learning. Actions either produce or prevent good or bad events, and the strength of the action changes accordingly (arrow). "Reinforcement" refers to a strengthening of behavior. Positive reinforcement occurs

when an action produces a positive event, whereas negative reinforcement occurs when an action prevents or eliminates a negative event. (Courtesy of Mark E. Bouton, PhD.) A parallel to Figure 2.3-1 exists in Pavlovian conditioning, in which one can likewise think of whether the CS is associated with positive or negative events (Fig. 2.3-2). Although such learning can lead to a wide constellation or system of behaviors, in a very general way, it also leads to behavioral tendencies of approach or withdrawal. Thus, when a CS signals a positive US, the CS will tend to evoke approach behaviors— called sign tracking. For example, an organism will approach a signal for food. Analogously, when a CS signals a negative US, it will evoke behaviors that tend to move the organism away from the CS. Conversely, CSs associated with a decrease in the probability of a good thing will elicit withdrawal behaviors, whereas CSs associated with the decrease in the probability of a bad thing can elicit approach. An example of the latter case might be a stimulus that signals safety or the decrease in probability of an aversive event, which evokes approach in a frightened organism. In the end, these very basic behavioral effects of both operant (see Fig. 2.3-1) and Pavlovian (see Fig. 2.3-2) learning serve to maximize the organism's contact with good things and minimize contact with bad things. FIGURE 2.3-2 Sign tracking in Pavlovian learning. Conditional stimuli (CSs) signal either an increase or a decrease in the probability of good or bad events, and the CS generally engages approach or withdrawal behaviors accordingly. (Courtesy of Mark E. Bouton, PhD.) Perhaps because they have such similar functions, Pavlovian learning and operant learning are both influenced by similar variables. For example, in either case, behavior is especially strong if the magnitude of the US or reinforcer is large, or if the US or reinforcer occurs relatively close to the CS or operant response in time. In either case, the learned behavior decreases if the US or reinforcer that was once paired with the CS or the response is eliminated from the situation. This phenomenon, called extinction, provides a means of eliminating unwanted behaviors that were learned through either

form of conditioning and has led to a number of very effective cognitive-behavioral therapies. PAVLOVIAN CONDITIONING Effects of Conditioning on Behavior Many lay people have the mistaken impression that Pavlovian learning is a rigid affair in which a fixed stimulus comes to elicit a fixed response. In fact, conditioning is considerably more complex and dynamic than that. For example, signals for food may evoke a large set of responses that function to prepare the organism to digest food: They can elicit the secretion of gastric acid, pancreatic enzymes, and insulin in addition to Pavlov's famous salivary response. The CS can also elicit approach behavior (as described earlier), an increase in body temperature, and a state of arousal and excitement. When a signal for food is

presented to a satiated animal or human, he or she may eat more food. Some of these effects may be motivational; for example, an additional effect of presenting a CS for food is that it can invigorate ongoing operant behaviors that have been reinforced with food. CSs thus have powerful behavioral potential. Signals for food evoke a whole behavior system that is functionally organized to find, procure, and consume food. Pavlovian conditioning is also involved in other aspects of eating. Through conditioning, humans and other animals may learn to like or dislike certain foods. In animals like rats, flavors associated with nutrients (sugars, starches, calories, proteins, or fats) come to be preferred. Flavors associated with sweet tastes are also preferred, whereas flavors associated with bitter tastes are avoided. At least as important, flavors associated with illness become disliked, as illustrated by the person who gets sick drinking an alcoholic beverage and consequently learns to hate the flavor. The fact that flavor CSs can be associated with such a range of biological consequences (USs) is important for omnivorous animals that need to learn about new foods. It also has some clinical implications. For example, chemotherapy can make cancer patients sick, and it can therefore cause the conditioning of an aversion to a food that was eaten recently (or to the clinic itself). Other evidence suggests that animals may learn to dislike food that is associated with becoming sick with cancer. On the flip side, conditioning can enable external cues to trigger food consumption and craving, a potential influence on overeating and obesity. Pavlovian conditioning also occurs when organisms ingest drugs. Whenever a drug is taken, in addition to reinforcing the behaviors that lead to its ingestion, the drug constitutes a US and may be associated with potential CSs that are present at the time (e.g., rooms, odors, injection paraphernalia,). CSs that are associated with drug USs can sometimes have an interesting property: They often elicit a conditioned response that seems opposite to the unconditional effect of the drug. For example, although morphine causes a rat to feel less pain, a CS associated with morphine elicits an opposite increase, not a decrease in pain sensitivity. Similarly, although alcohol can cause a drop in body temperature, a conditioned response to a CS associated with alcohol is typically an

increase in body temperature. In these cases, the conditioned response is said to be compensatory because it counteracts the drug effect. Compensatory responses are another example of how classical (Pavlovian) conditioning helps organisms prepare for a biologically significant US. Compensatory conditioned responses have implications for drug abuse. First, they can cause drug tolerance, in which repeated administration of a drug reduces its effectiveness. As a drug and a CS are repeatedly paired, the compensatory response to the CS becomes stronger and more effective at counteracting the effect of the drug. The drug therefore has less impact. One implication is that tolerance will be lost if the drug is taken without being signaled by the usual CS. Consistent with this idea, administering a drug in a new environment can cause a loss of drug tolerance and make drug overdose more likely. A second implication stems from the fact that compensatory responses may be unpleasant or aversive. A CS associated with an opiate may elicit several compensatory responses—it may cause the drug user to be more sensitive to pain, undergo a change in body temperature, and perhaps become hyperactive (the opposite of another unconditional opiate effect). The unpleasantness of these responses may motivate the user to take the drug again to get rid of them, an example of escape learning, or negative reinforcement, and a classic example of how Pavlovian and operant learning processes might readily interact. The idea is that the urge to take drugs may be strongest in the presence of CSs that have been associated with the drug. The hypothesis is consistent with self-reports of abusers, who, after a period of abstinence, are tempted to take the drug again when they are reexposed to drug-associated cues. Pavlovian learning may

potentially be involved in anxiety disorders. CSs associated with frightening USs can elicit a whole system of conditioned fear responses, broadly designed to help the organism cope. In animals, cues associated with frightening events (such as a brief foot shock) elicit changes in respiration, heart rate, and blood pressure, and even a (compensatory) decrease in sensitivity to pain. Brief CSs that occur close to the US in time can also elicit adaptively timed protective reflexes. For example, the rabbit blinks in response to a brief signal that predicts a mild electric shock near the eye. The same CS, when lengthened in duration and paired with the same US, elicits mainly fear responses, and fear elicited by a CS may potentiate the conditioned eyeblink response elicited by another CS or a startle response to a sudden noise. Once again, CSs do not merely elicit a simple reflex, but also evoke a complex and interactive set of responses. Classical fear conditioning can contribute to phobias (in which specific objects may be associated with a traumatic US), as well as other anxiety disorders, such as panic disorder and posttraumatic stress disorder (PTSD). In panic disorder, people who have unexpected panic attacks can become anxious about having another one. In this case, the panic attack (the US or UR) may condition anxiety to the external situation in which it occurs (e.g., a crowded bus) and also internal (“interoceptive”) CSs created by early symptoms of the attack (e.g., dizziness or a sudden pounding of the heart). These CSs may then evoke anxiety or panic responses. Panic disorder may begin because external cues associated with panic can arouse anxiety, which may then exacerbate the next unconditional panic attack and/or panic response elicited by an interoceptive CS. It is possible that the emotional reactions elicited by CSs may not require conscious awareness for their occurrence or development. Indeed, fear conditioning may be independent of conscious awareness. In addition to eliciting conditioned responses, CSs also motivate ongoing operant behavior. For example, presenting a CS that elicits anxiety can increase the vigor of operant behaviors that have been learned to avoid or escape the frightening US. Thus,

an individual with an anxiety disorder will be more likely to express avoidance in the presence of anxiety or fear cues. Similar effects may occur with CSs that predict other USs (such as drugs or food)—as already mentioned, a drug-associated CS may motivate the drug abuser to take more drugs. The motivating effects of CSs may stem from the fact that CSs may be associated with both the sensory and emotional properties of USs. For example, the survivor of a traumatic train derailment might associate stimuli that occur immediately before derailment (such as the blue flash that occurs when the train separates from its overhead power supply) with both the emotional and the sensory aspects of the crash. Consequently, when the survivor later encounters another flashing blue light (e.g., the lights on a police car), the CS might evoke both emotional responses (mediated by association with the trauma’s emotional qualities) and sensory associations (mediated by association with the trauma’s sensory qualities). Both might play a role in the nightmares and “re-experiencing” phenomena that are characteristic of PTSD. The Nature of the Learning Process Research beginning in the late 1960s began to uncover some important details about the learning process behind Pavlovian conditioning. Several findings proved especially important. It was shown, for example, that conditioning is not an inevitable consequence of pairing a CS with a US. Such pairings will not cause conditioning if there is a second CS present that already predicts the US. This finding (known as blocking) suggests that a CS must provide new information about the US if learning is to occur. The importance of the CS’s information value is also suggested by the fact that a CS will not be treated as a signal for the US if the US occurs equally often (or is equally probable) in the presence and the absence of the CS. Instead, the organism treats the CS as a signal for the US if the probability of the US is greater in the presence

of the CS than in its absence. In addition, the organism will treat the CS as a signal for “no US” if the probability of the US is less in the presence of the CS than in its absence. In the latter case, the signal is called a conditioned inhibitor because it will inhibit performance elicited by other CSs. The conditioned inhibition phenomenon is clinically relevant because inhibitory CSs may hold pathological CRs like fear or anxiety at bay. A loss of the inhibition would allow the anxiety response to emerge. There are also important variants of classical conditioning. In sensory preconditioning, two stimuli (A and B) are first paired, and then one of them (A) is later paired with the US. Stimulus A evokes conditioned responding, of course, but so does stimulus B—indirectly, through its association with A. One implication is that exposure to a potent US like a panic attack may influence reactions to stimuli that have never been paired with the US directly; the sudden anxiety to stimulus B might seem spontaneous and mysterious. A related finding is second-order conditioning. Here, A is paired with a US first and then subsequently paired with stimulus B. Once again, both A and B will evoke responding. Sensory preconditioning and second-order conditioning increase the range of stimuli that can control the conditioned response. A third variant worth mentioning occurs, as indicated previously, when the onset of a stimulus becomes associated with the rest of that stimulus, as when a sudden increase in heart rate caused by the onset of a panic attack comes to predict the rest of the

panic or feeling, or when the onset of a drug may predict the rest of the drug effect. Such intraevent associations may play a role in many of the body’s regulatory functions, such that an initial change in some variable (e.g., blood pressure or blood glucose level) may come to signal a further increase in that variable and therefore initiate a conditioned compensatory response. Emotional responses can also be conditioned through observation. For example, a monkey that merely observes another monkey being frightened by a snake can learn to be afraid of the snake. The observer learns to associate the snake (CS) with its emotional reaction (US/UR) to the other monkey being afraid. Although monkeys readily learn to fear snakes, they are less likely to associate other salient cues (such as colorful flowers) with fear in the same way. This is an example of preparedness in classical conditioning—some stimuli are especially effective signals for some USs because evolution has made them that way. Another example is the fact that tastes are easily associated with illness but not shock, whereas auditory and visual cues are easily associated with shock but not illness. Preparedness may explain why human phobias tend to be for certain objects (snakes or spiders) and not others (knives or electric sockets) that may as often be paired with pain or trauma. Erasing Pavlovian Learning If Pavlovian learning plays a role in the etiology of behavioral and emotional disorders, a natural question concerns how to eliminate it or undo it. Pavlov studied extinction: Conditioned responding decreases if the CS is presented repeatedly without the US after conditioning. Extinction is the basis of many behavioral or cognitive-behavioral therapies designed to reduce pathological conditioned responding through repeated exposure to the CS (exposure therapy), and it is presumably a consequence of any form of therapy in the course of which the patient learns that previous harmful cues are no longer harmful. Another elimination procedure is counterconditioning, in which the CS is paired with a very different US/UR. Counterconditioning was the inspiration for systematic desensitization, a behavior therapy technique in which frightening CSs are deliberately associated with relaxation during therapy. Although extinction and counterconditioning reduce unwanted conditioned responses, they do not destroy the original learning, which remains in the brain, ready to return to behavior under the right circumstances. For example, conditioned responses that have been eliminated by extinction or counterconditioning can recover if time passes before the CS is presented again (spontaneous recovery). Conditioned

responses can also return if the patient returns to the context of conditioning after extinction in another context, or if the CS is encountered in a context that differs from the one in which extinction has occurred (all are examples of the renewal effect). The renewal effect is important because it illustrates the principle that extinction performance depends on the organism being in the context in which extinction was learned. If the CS is encountered in a different context, the extinguished behavior may relapse or return. Recovery and relapse can also occur if the current context is associated again with the US (“reinstatement”) or if the CS is paired with the US again (“rapid reacquisition”). One theoretical approach assumes that extinction and counterconditioning do not destroy the original learning but instead entail new learning that gives the CS a second meaning (e.g., “the CS is safe” in addition to “the CS is dangerous”). As with an ambiguous word,

which has more than one meaning, responding evoked by an extinguished or counterconditioned CS depends fundamentally on the current context. Research on context effects in both animal and human learning and memory suggests that a wide variety of stimuli can play the role of context (Table 2.3-1). Drugs, for example, can be very salient in this regard. When rats are given fear extinction while under the influence of a benzodiazepine tranquilizer or alcohol, fear is renewed when the CS is tested in the absence of the context provided by the drug. This is an example of state-dependent learning, in which the retention of information is best when tested in the same state in which it was originally learned. State-dependent fear extinction has obvious implications for combining therapy with drugs. It also has implications for the administration of drugs more generally. For example, if a person were to take a drug to reduce anxiety, the anxiety reduction would reinforce drug taking. State-dependent extinction might further preserve any anxiety that might otherwise be extinguished during natural exposure to the anxiety-eliciting cues. Thus, drug use could paradoxically preserve the original anxiety, creating a self-perpetuating cycle that could provide a possible explanation for the link between anxiety disorders and substance abuse. One point of this discussion is that drugs can play multiple roles in learning: They can be USs or reinforcers on one hand, and CSs or contexts on the other. The possible complex behavioral effects of drugs are worth bearing in mind. Table 2.3-1 Effective Contextual Stimuli Studied In Animal and Human Research Laboratories Another general message is that contemporary theory emphasizes the fact that extinction (and other processes, such as counterconditioning) entails new learning rather than a destruction of the old. Recent psychopharmacological research has built on this idea: If extinction and therapy constitute new learning, then drugs that might facilitate new learning might also facilitate the therapy process. For example, there has been considerable recent interest in D-cycloserine, a partial agonist of the N-methyl-D-aspartate (NMDA) glutamate receptor. The NMDA receptor is involved in long-term potentiation, a synaptic facilitation phenomenon that has been implicated in several examples of learning. Of interest, there is evidence that the administration of D-cycloserine can facilitate extinction learning in rats and possibly in humans undergoing exposure therapy for anxiety disorders. In the studies supporting this conclusion, the administration of the drug increased the amount of extinction that was apparent after a

small (and incomplete) number of extinction trials. Although such findings are promising, it is important to remember that the context dependence of extinction, and thus the possibility of relapse with a change of context, may easily remain. Consistent with this possibility, although D-cycloserine allows fear extinction to be learned in fewer trials, it does not appear to prevent or reduce the strength of the renewal effect. Such results further underscore the importance of behavioral research—and behavioral theory—in understanding the effects of drugs on therapy.

Nonetheless, the search for drugs that might enhance the learning that occurs in therapy situations will continue to be an important area of research. Another process that might theoretically modify or erase a memory is illustrated by a phenomenon called reconsolidation. Newly learned memories are temporarily labile and easy to disrupt before they are consolidated into a more stable form in the brain. The consolidation of memory requires the synthesis of new proteins and can be blocked by the administration of protein synthesis inhibitors (e.g., anisomycin). Animal research suggests that consolidated memories that have recently been reactivated might also return briefly to a similarly vulnerable state; their “reconsolidation” can likewise be blocked by protein synthesis inhibitors. For example, several studies have shown that the reactivation of a conditioned fear by one or two presentations of the CS after a brief fear conditioning experience can allow it to be disrupted by anisomycin. When the CS is tested later, there is little evidence of fear—as if reactivation and then drug administration diminished the strength of the original memory. However, like the effects of extinction, these fear-diminishing effects do not necessarily mean that the original learning has been destroyed or erased. There is some evidence that fear of the CS that has been diminished in this way can still return over time (i.e., spontaneously recover) or with reminder treatments. This sort of result suggests that the effect of the drug is somehow able to interfere with retrieval or access to the memory rather than to be an actual “reconsolidation.” Generally speaking, the elimination of a behavior after therapy should not be interpreted as erasure of the underlying knowledge. For the time being, it may be safest to assume that after any therapeutic treatment, a part of the original learning may remain in the brain, ready to produce relapse if retrieved. Instead of trying to find treatments that destroy the original memory, another therapeutic strategy might be to accept the possible retention of the original learning and build therapies that allow the organism to prevent or cope with its retrieval. One possibility is to conduct extinction exposure in the contexts in which relapse might be most problematic to the patient and to encourage retrieval strategies (such as the use of retrieval cues like reminder cards) that might help to remind the patient of the therapy experience.

OPERANT/INSTRUMENTAL LEARNING

The Relation Between Behavior and Payoff Operant learning has many parallels with Pavlovian learning. As one example,

extinction also occurs in operant learning if the reinforcer is omitted following training. Although extinction is once again a useful technique for eliminating unwanted behaviors, just as we saw with Pavlovian learning, it does not destroy the original learning—spontaneous recovery, renewal, reinstatement, and rapid reacquisition effects still obtain. Although early accounts of instrumental learning, beginning with Edward Thorndike, emphasized the role of the reinforcer as “stamping in” the instrumental action, more-modern approaches tend to view the reinforcer as a sort of guide or motivator of behavior. A modern, “synthetic” view of operant conditioning (see later discussion) holds that the organism associates the action with the outcome in much the way that stimulus–outcome learning is believed to be involved in Pavlovian learning. Human behavior is influenced by a wide variety of reinforcers, including social ones. For example, simple attention from teachers or hospital staff members has been shown to reinforce disruptive or problematic behavior in students or patients. In either case, when the attention is withdrawn and redirected toward other activities, the problematic behaviors can decrease (i.e., undergo extinction). Human behavior is also influenced by verbal reinforcers, like praise, and, more generally, by conditioned reinforcers, such as money, that have no intrinsic value except for the value derived through association with more basic, “primary” rewards. Conditioned reinforcers have been used in schools and institutional settings in so-called token economies in which positive behaviors are reinforced

with tokens that can be used to purchase valued items. In more natural settings, reinforcers are always delivered in social relationships, in which their effects are dynamic and reciprocal. For example, the relationship between a parent and a child is full of interacting and reciprocating operant contingencies in which the delivery (and withholding) of reinforcers and punishers shapes the behavior of each. Like Pavlovian learning, operant learning is always operating and always influencing behavior. Research on operant conditioning in the laboratory has offered many insights into how action relates to its payoff. In the natural world, few actions are reinforced every time they are performed; instead, most actions are reinforced only intermittently. In a ratio reinforcement schedule, the reinforcer is directly related to the amount of work or responding that the organism emits. That is, there is some work requirement that determines when the next reinforcer will be presented. In a “fixed ratio schedule,” every xth action is reinforced; in a “variable ratio schedule,” there is an average ratio requirement, but the number of responses required for each successive reinforcer varies. Ratio schedules, especially variable ratio schedules, can generate high rates of behavior, as seen in the behavior directed at a casino slot machine. In an interval reinforcement schedule, the presentation of each reinforcer depends on the organism emitting the response after some period of time has also elapsed. In a “fixed interval schedule,” the first response after x seconds have elapsed is reinforced. In a “variable interval schedule,” there is an interval requirement for each reinforcer, but the length of that interval varies. A person checking e-mail throughout the day is being reinforced on a variable interval schedule—a new message is not present to reinforce each checking response, but the presence of a new message becomes available after variable time points throughout the day. Of interest, on interval schedules, the response rate can vary substantially without influencing the overall rate of reinforcement. (On ratio schedules, there is a more direct relationship between behavior rate and reinforcement rate.) In part because of this, interval schedules tend to generate slower response rates than ratio schedules. Classic research on operant behavior underscores the fact that the performance of any action always involves choice. That is, whenever the individual performs a particular behavior, he or she chooses to engage in that action over many other possible alternatives. When choice has been studied by allowing the organism to perform either of two different operant behaviors (paying off with their own separate schedules of reinforcement), the rate of operant behavior depends not only on the behavior’s rate of reinforcement, but also on the rate of reinforcement of all other behaviors in the

situation. Put most generally, the strength of Behavior 1 (e.g., the rate at which Behavior 1 is performed) is given by $B_1 = KR_1/(R_1 + R_0)$ where B_1 can be seen as the strength of Behavior 1, R_1 is the rate at which B_1 has been reinforced, and R_0 is the rate at which all alternative (or “other”) behaviors in the environment have been reinforced; K is a constant that corresponds to all behavior in the situation and may have a different value for different individuals. This principle, known as the quantitative law of effect, captures several ideas that are relevant to psychiatrists and clinical psychologists. It indicates that an action can be strengthened either by increasing its rate of reinforcement (R_1) or by decreasing the rate of reinforcement for alternative behaviors (R_0). Conversely, an action can be weakened either by reducing its rate of reinforcement (R_1) or by increasing the rate of reinforcement for alternative behaviors (R_0). The latter point has an especially important implication: In principle, one can slow the strengthening of new, undesirable behavior by providing an environment that is otherwise rich in reinforcement (high R_0). Thus, an adolescent who experiments with drugs or alcohol would be less likely to engage in this behavior at a high rate (high B_1) if his or her environment were otherwise rich with reinforcers (e.g., provided

by extracurricular activities, outside interests, and so forth). Choice among actions is also influenced by the size of their corresponding reinforcers and how soon the reinforcers will occur. For example, individuals sometimes have to choose between an action that yields a small but immediate reward (e.g., taking a hit of a drug) versus another that yields a larger but delayed reward (e.g., going to a class and earning credit toward a general educational development certificate). Individuals who choose the more immediate reward are often said to be “impulsive,” whereas those who choose the delayed reward are said to exercise “self-control.” Of interest, organisms often choose immediate small rewards over delayed larger ones, even though it may be maladaptive to do so in the long run. Such “impulsive” choices are especially difficult to resist when the reward is imminent. Choice is believed to be determined by the relative value of the two rewards, with that value being influenced by both the reinforcer’s size and its delay. The bigger the reinforcer, the better is the value, and the more immediate the reinforcer, the better too: When a reward is delayed, its value decreases or is “discounted” over time. When offered a choice, the organism will always choose the action that leads to the reward whose value is currently higher.

Theories of Reinforcement It is possible to use the foregoing principles of operant conditioning without knowing in advance what kind of event or stimulus will be reinforcing for the individual patient. None of the reinforcement rules say much about what sorts of events in an organism’s world will play the role of reinforcer. Skinner defined a reinforcer empirically, by considering the effect it had on an operant behavior. A reinforcer was defined as any event that could be shown to increase the strength of an operant if it was made a consequence of the operant. This empirical (some would say “atheoretical”) view can be valuable because it allows idiosyncratic reinforcers for idiosyncratic individuals. For instance, if a therapist works with a child who is injuring himself, the approach advises

the therapist merely to search for the consequences of the behavior and then manipulate them to bring the behavior under control. So if, for example, the child’s self-injurious behavior decreases when the parent stops scolding the child for doing it, then the scold is the reinforcer, which might seem counterintuitive to everyone (including the parent who thinks that the scold should function as a punisher). On the other hand, it would also be useful to know what kind of event will reinforce an individual before the therapist has to try everything. This void is filled by several approaches to reinforcement that allow predictions ahead of time. Perhaps the most useful is the Premack principle (named for researcher David Premack), which claims that, for any individual, reinforcers can be identified by giving the individual a preference test in which she or he is free to engage in any number of activities. The individual might spend the most time engaging in activity A, the secondmost time engaged in activity B, and the third-most time engaged in activity C. Behavior A can thus be said to be preferred to B and C, and B is preferred to C. The Premack principle asserts that access to a preferred action will reinforce any action that is less preferred. In the present example, if doing activity C allowed access to doing A or B, activity C will be reinforced—it will increase in strength or probability. Similarly, activity B will be reinforced by activity A (but not C). The principle accepts large individual differences. For example, in an early study, some children given a choice spent more time eating candy than playing pinball, whereas others spent more time playing pinball than eating candy. Candy eating reinforced pinball playing in the former group. In contrast, pinball playing reinforced candy eating in the latter group. There is nothing particularly special about food (eating) or any particular kind of activity as a possible reinforcer. Any behavior that is preferred to a second behavior will theoretically reinforce the second behavior. The principle has been refined over the years. It is now recognized that even a lesspreferred behavior can

reinforce a more-preferred behavior if the organism has been deprived of doing the low-preferred behavior below its ordinary level. In the foregoing example, even the low-preference activity C could reinforce A or B if it were suppressed for a while below its baseline level of preference. However, the main implication is that in the long run, a person's reinforcers can be discovered by simply looking at how he or she allocates his or her activities when access to them is free and unconstrained. Motivational Factors Instrumental action is often said to be goal oriented. As Edward Tolman illustrated in many experiments conducted in the 1930s and 1940s, organisms may flexibly perform any of several actions to get to a goal; instrumental learning thus provides a variable means to a fixed end. Tolman's perspective on the effects of reinforcers has returned to favor. He argued that reinforcers are not necessary for learning, but instead are important for motivating instrumental behavior. The classic illustration of this point is the latent learning experiment. Rats received several trials in a complex maze in which they were removed from the maze without reward once they got to a particular goal

location. When arriving at the goal was suddenly rewarded, the rats suddenly began working through the maze with very few errors. Thus, they had learned about the maze without the benefit of the food reinforcer, but the reinforcer was nonetheless important for motivating them to get through the maze efficiently. The reinforcer was not necessary for learning, but it gave the organism a reason to translate its knowledge into action. Subsequent research has identified many motivating effects of reward. For example, organisms that have had experience receiving a small reward may show positive contrast when they are suddenly reinforced with a larger reward. That is, their instrumental behavior may become more vigorous than that in control subjects who have received the larger reward all along. Conversely, organisms show negative contrast when they are switched from a high reward to a lower reward—their behavior becomes weaker than control subjects who have received the same smaller reward all along. Negative contrast involves frustration and emotionality. Both types of contrast are consistent with the idea that the current effectiveness of a reinforcer depends on what the organism has learned to expect; an increase from expectation causes elation, whereas a decrease from expectation causes frustration. There is a sense in which receiving a reward that is smaller than expected might actually seem punishing. Negative contrast is an example of a paradoxical reward effect—a set of behavioral phenomena given the name because they show that reward can sometimes weaken behavior and that nonreward can sometimes strengthen it. The best known is the partial reinforcement extinction effect, in which actions that have been intermittently (or “partially”) reinforced persist longer when reinforcers are completely withdrawn than actions that have been continuously reinforced. The finding is considered paradoxical because an action that has been reinforced (say) half as often as another action may nonetheless be more persistent. One explanation is that the action that has been partially reinforced has been reinforced in the presence of some frustration—and is thus persistent in new adversity or sources of frustration. Other evidence suggests that effortfulness is a dimension of behavior that can be reinforced. That is, human and animal participants that have been reinforced for performing effortful responses learn a sort of “industriousness” that transfers to new behaviors. One implication is that new behaviors learned in therapy will be more persistent over time if high effort has been deliberately reinforced. The effectiveness of a reinforcer is also influenced by the organism's current motivational state. For example, food is more reinforcing for a hungry organism, and water is more reinforcing for a thirsty one. Such results are consistent with many theories of reinforcement (e.g., the Premack principle) because the presence of hunger or thirst would undoubtedly increase the organism's preference ranking for food or water. Recent

research, however, indicates that the effects of motivational states on instrumental actions are not this automatic. Specifically, if a motivational state is going to influence an instrumental action, the individual needs first to learn how the action's reinforcer will influence the motivational state. The process of learning about the effects the reinforcer has on the motivational state is called incentive learning. Incentive learning is best illustrated by an experimental example. In 1992, Bernard Balleine reported a study that taught trained rats that were not hungry to lever press to earn a novel food pellet. The animals were then food deprived and tested for their lever pressing under conditions in which the lever press no longer produced the pellet. The hunger state had no effect on lever-press rate; that is, hungry rats did not lever press

any more than rats that were not food deprived. On the other hand, if the rat had been given separate experience eating the pellets while it was food deprived, during the test it lever pressed at a high rate. Thus, hunger invigorated the instrumental action only if the animal had previously experienced the reinforcer in that state—which allowed it to learn that the specific substance influenced the state (incentive learning). The interpretation of this result, and others like it, will be developed further later in this section. The main idea is that individuals will perform an instrumental action when they know that it produces an outcome that is desirable in the current motivational state. The clinical implications are underexplored but could be significant. For example, persons who abuse drugs will need to learn that the drug makes them feel better in the withdrawal state before withdrawal will motivate drug seeking. Persons with anxiety might not be motivated to take a beneficial medication while anxious until they have actually had the opportunity to learn how the medication makes them feel when they are in the anxious state, and a person with depression may need to learn what natural reinforcers really make them feel better while they are depressed. According to theory, direct experience with a reinforcer's effect on depressed mood might be necessary before the person will be interested in performing actions that help to ameliorate the depressed state.

PAVLOVIAN AND OPERANT LEARNING TOGETHER

Avoidance Learning Theories of the motivating effects of reinforcers have usually emphasized that Pavlovian CSs in the background are also associated with the reinforcer, and that the expectancy of the reinforcer (or conditioned motivational state) the CSs arouse increases the vigor of the operant response. This is two-factor or two-process theory: Pavlovian learning occurs simultaneously and motivates behavior during operant learning. The interaction of Pavlovian and instrumental factors is especially important in understanding avoidance learning (see Fig. 2.3-1). In avoidance situations, organisms learn to perform actions that prevent the delivery or presentation of an aversive event. The explanation of avoidance learning is subtle because it is difficult to identify an obvious reinforcer. Although preventing the occurrence of the aversive event is obviously important, how can the nonoccurrence of that event reinforce? The answer is that cues in the environment (Pavlovian CSs) come to predict the occurrence of the aversive event, and consequently they arouse anxiety or fear. The avoidance response can therefore be reinforced if it escapes from or reduces that fear. Pavlovian and operant factors are thus both important: Pavlovian fear conditioning motivates and allows reinforcement of an instrumental action through its reduction. Escape from fear or anxiety is believed to play a significant role in many human behavior disorders, including the anxiety disorders. Thus, the obsessive-compulsive patient checks or washes his or her hands repeatedly to reduce anxiety, the agoraphobic stays home to escape fear of places associated with panic attacks, and the bulimic learns to vomit after a meal to reduce the learned anxiety evoked by eating the meal.

Although two-factor theory remains an important view of avoidance learning, excellent avoidance can be obtained in the laboratory without reinforcement: for example, if an animal is required to perform an action that resembles one of its natural and prepared fear responses—so-called species-specific defensive reactions (SSDRs). Rats will readily learn to freeze (remain motionless) or flee (run to another environment) to avoid shock, two behaviors that have evolved to escape or avoid predation. Freezing and fleeing are also respondents rather than operants; they are controlled by their antecedents (Pavlovian CSs that predict shock) rather than being reinforced by their consequences (escape from fear). Thus, when the rat can use an SSDR for avoidance, the only necessary learning is Pavlovian—the rat learns about environmental cues associated with danger, and these arouse fear and evoke natural defensive behaviors including withdrawal (negative sign tracking; Fig. 2.3-2). To learn to perform an action that is not similar to a natural SSDR requires more feedback or reinforcement through fear reduction. A good example is lever pressing, which is easy for the rat to learn when the reinforcer is a food pellet but difficult to learn when the same action avoids shock. More recent work with avoidance in humans suggests an important role for CS-aversive event and response-no aversive event expectancies. The larger point is that Pavlovian learning is important in avoidance learning; when the animal can avoid with an SSDR, it is the only learning necessary; when the required action is not an SSDR, Pavlovian learning permits the expectation of something bad. A cognitive perspective on aversive learning is also encouraged by studies of learned helplessness. In this phenomenon, organisms exposed to either controllable or uncontrollable aversive events differ in their reactivity to later aversive events. For example, the typical finding is that a subject exposed to inescapable shock in one phase of an experiment is less successful at learning to escape shock with an altogether new behavior in a second phase, whereas subjects exposed to escapable shock are normal. Both types of subject are exposed to the same shock, but its psychological dimension (its controllability) creates a difference, perhaps because subjects exposed to inescapable shock learn the independence of their actions and the outcome. Although this finding (and interpretation) was once seen as a model of depression, the current view is that the controllability of stressors mainly modulates their stressfulness and negative impact. At a theoretical level, the result also implies that organisms receiving instrumental contingencies in which their actions lead to outcomes might learn something about the controllability of those outcomes. One of the main conclusions of work on aversive learning is that there are both biological (i.e., evolutionary) and cognitive dimensions to instrumental learning. The possibility that much instrumental learning can be controlled by Pavlovian contingencies is also consistent with research in which animals have learned to respond to positive reinforcers. For example, pigeons have been widely used in operant learning experiments since the 1940s. In the typical experiment, the bird learns to peck at a plastic disk on a wall of the chamber (a response “key”) to earn food. Although pecking seems to be an operant response, it turns out that the pigeon’s peck can be entrained by merely illuminating the key for a few seconds before presenting the reinforcer on a number of trials. Although there is no requirement for the bird to peck the key, the bird will begin to peck at the illuminated key—a Pavlovian predictor of food—anyway. The pecking response is only weakly controlled by its consequences; if the experimenter arranges things so that pecks actually prevent the delivery of food (which is otherwise delivered on trials without pecks), the birds will continue to peck almost indefinitely on many trials. (Although the peck has a negative correlation with food, key illumination remains a weakly positive predictor of food.) Thus, this classic “operant” behavior is at least partly a Pavlovian one. Pavlovian contingencies cannot be ignored. When rats are punished with mild foot shock for pressing a lever that otherwise produces

food, they stop lever pressing at least partly (and perhaps predominantly) because they learn that the lever now predicts shock and they withdraw from it. A child might likewise learn to stay away from the parent who delivers punishment rather than refrain from performing the punished behavior. A great deal of behavior in operant learning settings may actually be controlled by Pavlovian learning and sign tracking rather than true operant learning. A Synthetic View of Instrumental Action The idea, then, is that behavior in any instrumental learning situation is controlled by several hypothetical associations, as illustrated in Figure 2.3-3. Much behavior in an instrumental learning arrangement can be controlled by a Pavlovian factor in which the organism associates background cues (CSs) with the reinforcer (S^* , denoting a biologically significant event). As discussed earlier, this type of learning can allow the CS to evoke a variety of behaviors and emotional reactions (and motivational states) that can additionally motivate instrumental action. FIGURE 2.3-3 Any instrumental/operant learning situation permits a number of types of learning, which are always occurring all the time. R, operant behavior or instrumental action; S, stimulus in the background; S^* , biologically significant event (e.g., reinforcer, US). (Courtesy of Mark E. Bouton, PhD.) In modern terms, the instrumental factor is represented by the organism learning a direct, and similar, association between the instrumental action (R) and the reinforcer (S^*). Evidence for this sort of learning comes from experiments on reinforcer devaluation (Fig. 2.3-4). In such experiments, the organism can first be trained to perform two instrumental actions (e.g., pressing a lever and pulling a chain), each paired with a

different reinforcer (e.g., food pellet versus a liquid sucrose solution). In a separate second phase, one of the reinforcers (e.g., pellet) is paired with illness, which creates the conditioning of a powerful taste aversion to the reinforcer. In a final test, the organism is returned to the instrumental situation and is allowed to perform either instrumental action. No reinforcers are presented during the test. The result is that the organism no longer performs the action that produced the reinforcer that is now aversive. To perform in this way, the organism must have (1) learned which action produced which reinforcer and (2) combined this knowledge with the knowledge that it no longer likes or values that reinforcer. The result cannot be explained by the simpler, more traditional view that reinforcers merely stamp in or strengthen instrumental actions. FIGURE 2.3-4 The reinforcer devaluation effect. Results of the test session. The result indicates the importance of the response–reinforcer association in operant learning. For the organism to perform in the way that it does during testing, it must learn which action leads to which reinforcer and then choose to perform the action that leads to the outcome it currently liked or valued. R1, R2, operant behaviors or instrumental actions. (Data from Colwill and Rescorla [1986]. From Bouton ME: Learning and Behavior: A Contemporary Synthesis. Sunderland, MA: Sinauer; 2007.) Organisms also need to learn how reinforcers influence a particular motivational state—the process called “incentive learning.” Incentive learning is crucially involved in instrumental learning as a process through which the animal learns the value of the reinforcer. Thus, in the reinforcer devaluation experiment shown in Figure 2.3-4, an important thing that occurs in the second phase is that the organism must actually contact the reinforcer and learn that it does not like it. As described earlier, incentive learning is probably always involved in making outcomes (and the associated actions

that produce them) more or less desirable. Other experiments have illustrated the other associations to the stimulus that are represented in Figure 2.3-3. In addition to being directly associated with the reinforcer, a stimulus can signal a relation between an action and an outcome. This is called occasion setting: Instead of eliciting a response directly, stimuli in operant situations can set the occasion for the operant response. There is good evidence that they do so by signaling

a specific response-reinforcer relationship. For example, in one experiment, rats learned to lever press and chain pull in the presence of a background noise and a background light. When the noise was present, lever pressing yielded a pellet reinforcer, and chain pulling yielded sucrose. In contrast, when the light was present, the relations were reversed: Lever pressing yielded sucrose, and chain pulling yielded pellet. There was evidence that the rats learned corresponding relationships. In a second phase, pellets were associated with illness, so the rat no longer valued the pellet. In a final test, rats were allowed to lever press or chain pull in extinction in the presence of noise or light present during separate tests. In the presence of the noise, the animals chain pulled rather than lever pressed. When the light was present, the animals lever pressed rather than chain pulled. Thus, the noise informed the rat that lever press yielded pellet, and the light informed the rat that the chain pull did. This is the occasion-setting function illustrated in Figure 2.3-3. It is worth noting that other stimuli besides lights and noises set the occasion for operant behavior. Modern research on learning in animals has underscored the importance of other stimuli, such as temporal and spatial cues, and of certain perception and memory processes. A particularly interesting example of research on the stimulus control of operant behavior is categorization. Pigeons can be shown images of cars, chairs, flowers, and cats on a computer screen positioned on the wall of a Skinner box. Pecking one of four keys in the presence of these images is reinforced in the presence of any picture containing a car, a chair, a flower, or a cat. Of interest, as the number of exemplars in each category increases, the pigeon makes more errors as it learns the discrimination. However, more exemplars create better learning in the sense that it is more ready to transfer to new test images—after many examples of each category, the pigeon is more accurate at categorizing (and responding accurately to) new stimuli. One implication is that training new behaviors in a variety of settings or ways will enhance generalization to new situations. The final association in Fig. 2.3-3 is simple habit learning, or a direct association between the stimulus and the response. Through this association, the background may elicit the instrumental action directly, without the intervening cognition of $R-S^*$ and the valuation of S^* . Although $S-R$ learning was once believed to dominate learning, the current view sees it as developing only after extensive and consistent instrumental training. In effect, actions that have been performed repeatedly (and repeatedly associated with the reinforcer) become automatic and routine. One source of evidence is the fact that the reinforcer devaluation effect—which implies a kind of cognitive mediation of operant behavior—no longer occurs after extensive instrumental training, as if the animal reflexively engages in the response without remembering the actual

Revision #1

Created 2026-01-04 19:50:31 UTC by Omar Ayman

Updated 2026-01-04 19:50:31 UTC by Omar Ayman