Chapter 9

ADVANCES IN RESEARCH

There have been many advances in conditioning and learning research during the past decades and many of them are discussed in the preceding chapters. Others have not been included, either because they were not uniquely relevant to a particular theorist's ideas, or because they would have required an unreasonable lengthening of a chapter. This would have been the case for the Pavlov and Skinner chapters in particular.

This chapter provides a selection of important research related to operant and classical conditioning. It includes only research that has important practical applications or that has real significance in furthering our understanding of ourselves. I have purposely left out two kinds of research: that which already appears in textbooks so frequently that most students cannot help but be exposed to it and that which has no obvious practical relevance, though it may nevertheless be important. An example of the first category is the work of Piaget, and an example of the second is modern connectionist (neural network) theory.

I have been at pains to include findings that are of great importance and that are not widely known, though they already may be appreciated by specialists. Examples are found in the novel interpretation of conditioned reinforcement by Neuringer and Chung, in reinterpretations of what were formerly called "pleasure centers," and in research on behavioral contrast.

1. INSTRUMENTAL CONDITIONING AND THE LAW OF EFFECT

The Law of Effect and Hedonism

We have seen that Thorndike (e.g., 1913), Skinner (e.g., 1938, 1953) and other advocates of the law of effect repeatedly emphasized that satisfiers and annoyers, as well as reinforcers and punishers, had nothing necessarily to do with sensory pleasures and pains. Yet, there seems little doubt that the law of effect will continue to be viewed as a hedonistic principle, both by its critics and by some of it adherents. Needless to say, the same interpretation appears in less authoritative sources, and things would no doubt be clearer were it not for Thorndike's sad choice of terms, which invite such interpretations, as well as Skinner's common remark to the effect that we find this or that "reinforcing" (e.g., 1953, 1983b). The virtue of the law of effect, as demonstrated by Thorndike, Skinner, and many others, lies in its application as an empirical generalization referring to the important effects that consequences have on our behavior and our experience. Naturally, some of these consequences are accompanied by pleasures and pains, but often they are not. Being told that one's answer was "right" acts as a reinforcer for the answer that produced it, but is such a consequence really that pleasant? Is it pleasant to touch a wound to see if it still hurts? Yet, the slight pain that ensues acts as a strong reinforcer for touching. If we treat such mild pain as pleasant, we lose whatever meaning pleasure and pain have. They are important in our lives but utterly irrelevant to the law of effect.

The reason that *hedonism* is brought up in a chapter devoted to research developments is that we now have abundant evidence that almost anything can act as a reinforcer or as a punisher under the proper conditions. To understand some of the following developments, we must realize that reinforcement and punishment are not synonyms for pleasure and pain.

With this view, it will be instructive to examine recent research concerned with so-called reward and punishment centers in the brain. This topic has received a lot of popular attention, and the belief in the existence of such centers reflects the influence of the old hedonistic view. If such centers exist, it must be that pleasures and pains are fundamental causes of thought and action. But are things that simple? Could stimulating electrodes provide us with pleasure whenever we wished? The press suggests that this is the case, but consider the data.

Pleasure Centers in the Brain

A well-publicized discovery was made through experimental error. James Olds, a postdoctoral student of Donald Hebb at McGill University in Montreal, was learning to perform electrode implants in the brains of rats. His purpose was to insert a stimulating electrode in the reticular formation and determine whether stimulation led to faster maze learning. Hebb believed that this area of the upper spinal cord and medulla was responsible for arousal of the entire forebrain and that stimulation there might speed up learning. Moruzzi and Magoun (1949) had named the area the Ascending Reticular Activating System, or ARAS.

As it happened, the electrode bent on the way to the ARAS and lodged far forward, in the medial forebrain bundle. Both Olds and Peter Milner, who was working with him, noted the odd effects of stimulation produced by this implant. When stimulated, the rat returned to the place where it was standing when the stimulation was delivered; when a lever was added and pressing produced bursts of stimulation, the rat pressed!

Is this a significant finding? According to Hebb (1976), who wrote a eulogy after Olds's death, little attention was paid to it at the time. But a newspaper reporter for the *Montreal Star* came around to see what was new around McGill University and took some notes on Olds's and Milner's finding. As Hebb put it, the reporter "laid his story away in a drawer until one day when he was short of copy and turned it in as current news on the university scene." A higher-up in the newspaper was anxious for local news for the front page and ran a headline something like, "McGill Scientists Open Great New Field of Research," along with a lengthy story. This sensationalism embarrassed the university in general and Hebb in particular.

Whether Olds had actually discovered pleasure centers in the brain is another matter, which we will discuss later. But, given the publicity, it was difficult for him to believe that he had not done so, and he spent years defending the fact that he had (e.g., Olds, 1956). It is true that some placements of electrodes in the brains of rats (and humans) *appear* to produce pleasure and there is no question that such stimulation can act as a potent reinforcer.

The limbic system, where brain stimulation is effective, was named by the neuroanatomist James Papez in 1936. It is located deep in the center of the brain. Rats will press a lever over long periods of time when the only payoff is brain stimulation; Valenstein (1973) reported a rat pressing a lever 850,000 times over 21 days for such a reward. Others have reported that a rat may starve in the presence of food because it is "chained" to the lever that provides electrical stimulation of the brain (ESB), or intracranial stimulation (ICS), as it is also called. During the 1960s, science fiction writers used ICS as a theme in a number of pieces, in which evil "electrogarths" enslaved populations with stimulating electrodes implanted in their brains.

Olds did show that ICS may act much as other reinforcers, such as food. A rat will press a lever for it and will cross a shock grid to get it. Additionally, the discovery of Olds's so-called reward center was only a reasonable extension of what José Delgado had found a few years earlier; he had reported the existence of so-called punishment centers in the brain.

What Delgado found at that time (published years later) was that cats and other animals appear to dislike stimulation in some parts of the limbic system (1955). Not only will they learn to press a bar to avoid or escape it, but such stimulation seemed to punish other behaviors. In an obviously theatrical demonstration, Delgado (1969) showed that ICS can stop the charge of a "brave bull." The demonstration occurred in a rented bull ring in California, where Delgado stopped the bull as it was about to trample him! He also showed that ICS could punish eating, as well as interrupt the care of infant rhesus monkeys by the normally attentive mother. Little doubt remained in the minds of many that Delgado had demonstrated the existence of punishment centers in the brain. Olds's discovery just supplied the other side of the coin.

We have noticed that the theory of hedonism seems inadequate to explain the behavior of ourselves or of animals. Nonetheless, this theory has such a long history and is so ingrained in the minds of both laypersons and brain researchers that we all are prepared to believe that there might be brain centers specifically devoted to the experiences of pleasure and pain. Larry Stein, a prominent researcher, even suggested that reward and punishment circuits are permanently wired in the brain and that they are served by two different neurotransmitters. He suggested that punishment circuits pass through the midline of the limbic system, just adjacent to the cerebral ventricles (Stein, 1964). This periventricular system is supposed to be a cholinergic network; the neurotransmitter involved is acetylcholine. The lateral portions of the limbic system are assumed to be adrenergic; the neurotransmitter is norepinephrine and this network mediates reward—the pleasure center.

Olds tried very hard to show that ICS did indeed tap the biological bases activated by normal rewards and punishments. Olds and Milner (1954) showed that rats would press bars for ICS, would cross charged grids for it, and that its effects increased when animals were hungry. Similarly, Neal Miller (1958) argued that such stimulation produced actual hunger and thirst. ICS makes sated animals eat food as well as inedible objects! Isn't that real hunger?

A Reappraisal Since the early work of Delgado and of Olds and Milner, countless studies have been conducted to investigate the nature of rewarding and punishing ICS, as well as merely to use it in the course of other investigations. At the same time, it has become increasingly evident that we may have been hasty in describing the effects of ICS as the activation of reward/pleasure and punishment/pain centers.

ICS and Hedonism We now have clear evidence that the rewarding effects of ICS have nothing to do with pleasure and pain. Even before the accident that led Olds to the discovery of the effects of ICS in rats, a surgeon at Columbia University named Robert Heath had stimulated the limbic system in humans. His purpose was to provoke some sort of effect in schizophrenic patients, who are often characterized by the lack of emotional expression. Heath knew that the limbic system had been linked with the experience and expression of emotion since the 1930s and he decided to try electrical stimulation of limbic structures in patients for whom nothing else seemed to work.

This work was begun as early as 1950 (Valenstein, 1973), and continued through the 1960s. No patient reported feelings of great pleasure, even though the sites stimulated were in structures that were later defined as part of the reward system (for example, the septum, medial forebrain bundle, and lateral hypothalamus). According to Valenstein, Heath and other researchers found that some patients reported some pleasure, especially if they had been in pain; others reported pleasure and they continued to report it long after the stimulating equipment was turned off! Remember, the subjects were not average persons; they were people in such condition that they had submitted to an operation involving the implanting of an electrode in the middle of their brains!

In addition to a conspicuous lack of pleasure when the so-called pleasure center was stimulated, there were no reports of hunger, thirst, and so on when various parts of the hypothalamus were stimulated. The typical effect of hypothalamic stimulation was "abdominal discomfort, a feeling of warmth and a feeling of fullness of the head, and a pounding heart" (Valenstein, 1973).

Delgado and Punishment Valenstein (1973) was very critical of Delgado as the most casual reader of his book must perceive. In Valenstein's view, Delgado has sometimes promoted his views by using somewhat deceptive evidence. Take the case of the brave bull, which appears in many introductory textbooks. As mentioned earlier, Delgado rented a bull ring in southern California, called the press, and demonstrated that ICS of a punishment center could cause the inhibition of aggression in the animal. Delgado had implanted a stimulating electrode, which could be remotely activated and which would stop the bull's charge. When the big moment came and the cameras were rolling, the bull charged Delgado, who stopped it in its tracks just before the beast reached him. The punishing stimulation was so effective that Delgado could walk up to the bull and pet it!

But Delgado was stimulating the caudate nucleus, an area with motor functions in humans and other animals (excluding birds). If the caudate nucleus on one side is stimulated in a human, perhaps in an attempt to treat Parkinsonism, the effect is paralysis, turning of the head in the direction away from the side stimulated. How does this relate to the bull? Films of the charge show clearly why the bull stopped before reaching Delgado; it was too busy turning and, presumably, becoming momentarily paralyzed. Delgado may have stopped the bull, but describing the effect as activation of a punishment center was clearly misleading.

ICS, Reward, and Punishment The findings reported by investigators using ICS have indeed been very informative. The assumptions of the original researchers have proven false; there are not necessarily reward and punishment centers as such in the brain, and reinforcement and punishment surely are not synonyms for pleasure and pain. What does happen when a rat presses a lever 850,000 times in 21 days for the ICS that each lever press produces? By most definitions, such stimulation is a reinforcer, but our ignorance about the nature of reinforcement means that we do not really understand why such stimulation is reinforcing.

One finding that gives clues in this regard concerns a difference between ICS rewards and conventional reinforcers. Apparently, rats and other subjects that respond to produce ICS often seem to forget how pleasant the stimulation is and therefore must be primed, or given a few free stimulations before they return to the lever that produces it. This might seem odd; equally odd, if we believe that the rat that presses hour after hour enjoys it, is the finding that cats and rats also press a lever to turn off ICS. An experiment by Phillips, Cox, Kakolewski, and Valenstein (1969) illustrates what happens and how these data might be interpreted if one is looking for specific centers in the brain.

Rat subjects were given electrode implants in the so-called lateral hypothalamic reward system. They were then placed in a box containing two photocell assemblies, one at each end of the box. When the rats interrupted one photo beam they received brain stimulation, which continued until the beam at the other end of the box was broken. What did the rats do under such circumstances? They ran back and forth, turning on the stimulation, turning it off, turning it on, and so on.

Then food pellets were placed on the side of the box, where stimulation was turned on. The animals picked up a food pellet, carried it to the opposite end of the box, and dropped it when the stimulation ended. Thus, the stimulation seemed to be eliciting a common rat behavior, hoarding: the reward center also acted as a "food-hoarding" center! But that is not the whole story.

When other objects, such as rubber erasers, were placed at one end of the box, these were also carried and deposited at the other end of the box. Does this mean that the stimulation was activating a "hoarding of office supplies" center? Next, a rat pup was placed at the "stimulation on" end of the box and it was carried across the box and deposited. Since this resembled the retrieval of pups that occurs as a part of rat parental behavior, one could imagine that a "parental behavior" center had been activated.

But the stimulating electrode was always in the same site; only the material available to be carried was varied. After this training the rats' tendency to carry objects across the box was so strong that when there was nothing to carry, the animal occasionally picked up its tail or one of its forelegs, struggled across the box, and "deposited" it. If some specific center were being activated, it is clearly difficult to assign it a name. Also, though all of the subjects were nominally stimulated in the same hypothalamic region, the individual placements varied quite a bit. What was going on?

The effect occurred only when the on and off beams were at opposite ends of the chamber; if the same pattern of stimulation was given independently of the animals' position, no carrying occurred. One can only conclude that when rats are motivated to move back and forth and when this is accompanied by a parallel arousal and calming, they are apt to carry things. Our old conception of specific centers does not seem to help us here. We need a better conception of reward and punishment than that provided by the "brain centers" view, and from the old theory of hedonism. Let us consider some more promising alternatives.

Premack's Principle: The Relativity of Reinforcement

How does the law of effect work? Why is it that some things fairly regularly act as reinforcers and others as punishers? We saw that Thorndike had difficulties with this problem and concluded only that the effects of satisfiers depend on readiness, or the ongoing preparatory behavior that precedes them. Many reinforcers, such as food and water, obviously act to reduce biological *drives*, as Hull emphasized. But many other reinforcers seem to increase drive levels, as is the case when an organism learns to press a bar for the opportunity to explore or when we travel across the city to watch a horror movie.

By the late 1950s it seemed that we would never explain the law of effect, as Hull had tried to do. Yet, it seemed unsatisfying to be left with no more than Skinner's empirical law of effect, which defines a reinforcer as anything that acts to strengthen a behavior. This leaves the law of effect as a circular law: We don't know what behaviors are reinforceable until we apply what we hope is a reinforcer. And we know that it is a reinforcer only when we see its effect. Meehl (1950) attempted to deal with the problem by suggesting that reinforcers are *transsituational;* we may have to try out our reinforcers initially, but once we have found one, we can use it to strengthen any behavior. But what this provides is a list of reinforcers, which is little more than we began with, and we quickly find so many exceptions to transsituationality that we throw out our list.

David Premack, in a number of insightful papers (e.g., 1959, 1965) seemed to solve this problem; the *Premack Principle* could tell us in advance what would act as a reinforcer or punisher. To use it required a change in our basic views of stimuli, behaviors, and consequences.

First, it has been common to treat the world, our behavior, and its consequences as separate things; hence, the distinction between stimuli, responses, and reinforcers seems natural. Premack suggested that a better way of casting things is in *response language*. Our behavior is a set of actual and possible activities that range along a continuum of *value*. Rather than speak of a response (of no intrinsic value) reinforced by food (of high intrinsic value), we constantly experience transitions among activities, each of some value. The value of an activity, such as eating, resting, or watching television, may change from time to time, of course, but that is not a major problem. At any given time we may assess the value of various activities by observing how much time we (or our subject) spend engaged in them.

The essence of the Premack Principle lies in the suggestion that engaging in an activity (such as bar pressing) will be reinforced if it provides access to a higher-valued activity. At first sight this may seem to be merely restating the obvious: food reinforces bar pressing because eating is of higher value than pressing: hardly a startling revelation. Premack's Principle accounts for familiar cases of reinforcement but seems to add little to our understanding of what is going on. But that is not the case. The Premack Principle also accounts for a variety of instances of reinforcement that have escaped our notice because of our preconceptions about pleasure and drive reduction.

For example, a rat will run for the opportunity to eat under some circumstances, which have often been taken for granted. Does this mean that running is just a no-value behavior and that eating is always a reinforcer? What if we feed the rat to bursting while confining it in a little cage? Food will no longer reinforce anything, and running may become a very highly valued activity. It is then possible to train the rat to eat for the opportunity to run. The opportunity to engage in a higher-valued behavior always reinforces a lower-valued behavior, even when the latter appears to be a *real* reinforcer.

The same event may even act as a positive and a negative reinforcer at almost the same time. (Recall that a negative reinforcer strengthens a behavior that removes it; a burn negatively reinforces the removing of our hand from the hot stove.) Consider the following conditions for a rat: In its cage, there is a running wheel with a bar to press and a water tube to lick mounted on the inside axle, so that pressing and licking can occur while the rat is running. Pressing the bar (a low-valued activity) allows the occurrence of a higher-valued activity; the wheel unlocks and is turned by a small motor, allowing easy running. The problem is that the motor continues turning and running becomes less appealing as time goes on. Of course, one need not run, but the alternative is tumbling, as the clothes in a dryer do.

Premack (1965) showed that rats would press a lever, which started the wheel, and would lick a water spout, which then stopped the wheel. Forced running thus acted as a positive reinforcer for lever pressing and shortly after as a negative reinforcer for spout licking. That's not very surprising? Once shown that the same activity may act as a reinforcer under some conditions and as a punisher or a negative reinforcer under others, it is easy to believe that we always knew it. But we actually did not; when we think of food, praise, money, and so on as reinforcers we do not typically bear in mind that such things act as reinforcers only under specific conditions. Virtually anything we do will act as a reinforcer for something else we do if the conditions are right. And what are the right conditions? All that is required is that the value of the two activities differ and that the lower-valued activity produce access to the higher-valued activity.

The only problem with this is the difficulty encountered when we try to assign values. Time spent in different activities may work in some cases, as when we assess the value of a child's playing with different toys, or the value of a rat's running, pressing, or gnawing wood. But what of activities such as eating, watching television, reading, and sleeping? Is thirty minutes spent eating equivalent to the same amount of time watching TV or sleeping? If a person spends an hour a day running and eight hours sleeping, is the value of sleeping eight times that of running? This problem of *commensurability of units* of activities gave Premack a good deal of trouble, and he never really resolved it. Nor did he resolve the problem of changes in the value of activities with time. For example, as a rat is eating for the opportunity to run, the value of running surely decreases with repeated opportunities. Do we reach a point at which the value of running is lower than that of eating and acts only to punish eating? *Value* is nebulous and always changing; it is not as easy to scale as Premack assumed.

Response Deprivation

The problem of dealing with activities differing in value was pointed out first by Eisenberger, Karpman, and Trattner (1967) and later was elaborated upon by Timberlake and Allison (1974). What is really important in Premack's Principle is not differences in the values of activities but the restriction of access to activities used to demonstrate the reinforcement of lower- by higher-valued activities. For example, consider two activities, eating and running, in humans or rats. Assume that running currently is the higher-valued activity. This means that our subject will eat if such activity produces access to running, but it will not run for the opportunity to eat; our subject will simply continue running if given the chance, and any eating will have to be forced. But if we look closely at our demonstration of this fact we find something very interesting.

To demonstrate that our subject will eat for the opportunity to run, we have to restrict its running, allowing only a limited amount of it to occur after so much eating. Could it be that this restriction is the really important factor, whatever the values of the activities? We decided that running was of higher value than eating, since we saw that our subject spent more time running than eating just before our experiment. We then restricted its access to running to a level less than it wanted by requiring so much eating for opportunities to run. But what if we reversed the contingencies here and restricted the opportunity to engage in the lower-valued activity so that eating could not occur even at its low-value level? Suppose we then require an increase in time spent running for opportunities to eat and find that the amount of running increases. The lower-valued activity (eating) acts as a reinforcer for the higher-valued activity (running). Timberlake and Allison report numerous instances of this kind (though not this one), all showing the importance of what they called *response deprivation*. For example, if given a choice between licking drinking spouts that provide three-percent or four-percent sucrose, rats spend more time licking the four-percent solution; in Premack's terms, it is the higher-valued activity. Suppose now that access to the less preferred three-percent solution tube is blocked and an increase in licking the four-percent solution tube provides brief access to the three-percent tube. Licking the four-percent tube increases, reinforced by access to the less favored solution, which remains acting as a reinforcer as long as access is restricted.

Thus, one need only have a measure of the baseline (unrestricted) levels of occurrence of whatever activities are of interest in whatever units seem appropriate. The units could be time spent in one activity and frequency of occurrence of another; the units need not be commensurate. If access to one activity is then restricted so that it cannot occur at its baseline level and a second, unrestricted, activity produces access to it, the latter will increase above its baseline level; it will be reinforced.

Practical implications of this extension of the Premack Principle are easy to imagine. For example, to influence the behavior of schoolchildren, one needs no tokens or other externally imposed reward system. All that is necessary is to monitor current levels of several behaviors, such as sitting quietly, reading, playing, fighting, or whatever else occurs with some frequency. If a restriction is then imposed, so that one (or more) behavior(s) cannot occur at baseline levels and an increase in another behavior allows access to the first, the latter will increase in frequency. Access to the restricted behavior acts to reinforce the behavior that produces access and will continue to do so as long as the restricted behavior is not allowed to occur at its baseline level. This means that any behavior may be used to reinforce any other behavior, given that both originally occur at some baseline level. The opportunity to read may reinforce sitting quietly, just as fighting may be reinforced by the opportunity to read.

Consider a hypothetical example illustrated in Figure 9.1. The left column represents observed levels of four behaviors shown by a child over a seven-day period. Note that talking with friends and solitary play are scaled in average time in minutes per day, while reading and running are recorded as average number of bouts greater than five minutes duration occurring per day. The units in which activities are scaled need not be the same.

Suppose, for illustration's sake, that one wishes to increase the time that such a child spent reading. According to the response deprivation principle, restriction of access to a second activity renders the opportunity to engage in that activity a reinforcer for other behavior that produces access to the restricted behavior. Naturally, the level of restricted behavior must be kept below its unrestricted baseline throughout.

We might easily restrict the opportunity to talk with friends (assuming the presence of other children, as is the case while the child is in school). We then could require that the time spent reading increase from its average of five minutes per hour to fifteen minutes per hour; when such a longer bout occurs we allow a ten-minute break to talk with friends. We continue this procedure, being certain that the time allowed talking with friends remains below the baseline level of 100 minutes per day. Alternatively, what would happen if we decided to reinforce time spent in solitary play with opportunity to read? Could it be done?

This view of reinforcement (and punishment) seems so obvious that one wonders why it was not proposed earlier. In fact, Premack (1971) seemed to show some irritation in his last general attempt to get the point across; it is frustrating to speak to deaf ears. As simple as Premack's principle and Timberlake and Allison's extensions are, they have not been universally received. A reasoned criticism of Premack's Principle appears in Smith (1974); unless you understand the principle well, Smith's criticisms are difficult to counter.

Conservation Theory

A recent derivative of Premack's view appears in the simple conservation theory of Allison, Miller, and Wozny (1979). By the way of background, they cited a number of studies that used both animal and human subjects and showed odd relationships between work requirements and payoffs (Collier, 1972).

1. As work required increases, work output increases, but total payoff decreases. This occurs well before a "ceiling effect" on work output is reached. Subjects will simply work so much, even though this means that payoff decreases.
2. As magnitude of payoff increases, total payoff received increases, but work output decreases.
3. If one behavior is required for access to a second behavior, increases in the former are accompanied by decreases in the latter.

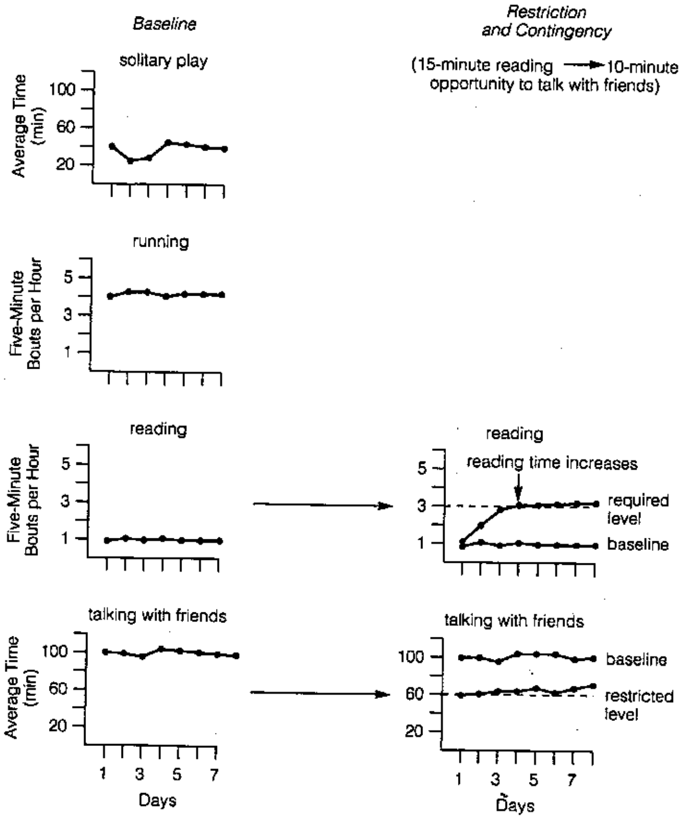

FIGURE 9.1 Results of response deprivation for a child's activities.

These effects suggest a *conservation process*, whereby total behavior (e.g., working and eating) remains constant as work requirements and payoffs change. Allison, et al. cite to "the total amount of a dimension" that is conserved. To illustrate, they describe a hypothetical experiment, beginning with a baseline session in which a rat is free to press a lever or drink from a water spout. They assume, for purposes of illustration, that each lever press requires twice as much work as does a one-second drink. Suppose that during the baseline session the rat drinks for 1,194 seconds and presses three times. In their terms, this reflects the operant level of pressing and drinking, or Oi and Oc, where "*i*" refers to pressing (an instrumental behavior) and "*c*" refers to drinking (a consummatory behavior).

Total work (in drinking units) is $kOi + Oc$. The constant "*k*" is used to express pressing in drink-sec units; in this case, $k = 2$. Thus, the work done during baseline is $kOi + Oc = 2(3) + 1{,}194 = 1{,}200$ drink-sec units. This is the quantity that is conserved when conditions are changed and the change amounts to a reciprocal schedule. Such a schedule requires so much of one activity for access to so much of another activity, for access to so much of the first activity, and so on. This will be clear below.

What happens when this rat is required to press once for a ten-second opportunity to drink, which is required for a second opportunity to press, which is required for a second ten-second drink, and so on? How many times will the rat go through this cycle? If 1,200 units occurred in baseline, that is what will occur under this contingency. One press for ten seconds of drinking translates to $2(1) + 10 = 12$ units; the rat will therefore go through the press/drink sequence $1{,}200/12 = 100$ times if conservation holds.

In general, $N(kI + C) = kOi + Oc$, which means that the subject will perform the two responses $(kI + C)$ required by the reciprocal schedule N times. N is the number of times necessary to equal the total amount of behavior during baseline $(kOi + Oc)$. Thus, $N = kOi + Oc/kI + C$. Such conservation holds, the authors tell us, in studies using rats drinking saccharine solutions, as well as when humans study Russian for the opportunity to sew or knit!

As an example of their evidence, the first experiment reported baseline levels of drinking and pressing for rats. Then, in different conditions, they required either twenty, 30, or 40 seconds of drinking for ten, twenty, or 30 seconds of pressing (for twenty, 30, or 40 seconds of drinking, and so on). As the requirements for the two behaviors were varied, conservation obtained; the obtained and predicted time spent pressing and drinking corresponded closely. And, as the required pressing increased, total pressing increased but total drinking decreased. As the drinking required increased, total drinking increased but pressing decreased.

Other experiments corroborated the conservation theory, which is presented here only to give a flavor of the type of research that has evolved from the Premack Principle. You may see that there are some problems with the conservation model and with similar models, by examining the original article in *Journal of Experimental Psychology: General* (Allison, Miller, Wozny, 1979). That particular issue of the journal is devoted wholly to such models, and a survey of the various viewpoints presented shows that matters are not as simple as Allison, et al. considered them (see Staddon, 1979). Premack's Principle and derivatives certainly depart from older conceptions of the law of effect. They emphasize the relativity of reinforcement as a relation among activities, blur the distinction between behaviors and reinforcers, and surely advance our understanding beyond a simple hedonism.

The Matching Law:
The Molar Law of Effect

A great deal of interest has focused on what Herrnstein calls the matching law (e.g., 1958, 1970). Simply put, an organism obeying the *matching law* distributes responses so as to match the relative frequencies of reinforcement associated with a set of alternative behaviors.

For example, suppose that a child, a rat, or a pigeon spends an hour a day in front of a panel of buttons—one red, one white, and one green. Suppose also that button presses are reinforced occasionally by candy, food pellets, or grain. Finally, suppose that responding to green is reinforced ten times per hour, whereas responses to white and red are reinforced five times per hour, respectively. What will our subject spend its time doing? It could find that green pays off a lot more often than does white or red and simply spend all of the time pressing or pecking green. In the same way, an animal that forages in several areas could well spend all of its time in that area most abundant in food. Herrnstein (and many others since) have shown that the child, rat, or pigeon tends to match relative responding to the three alternatives to the relative frequencies of reinforcement associated with each. In our example, the relative frequencies of reinforcement for red, white, and green are 0.25, 0.25, and 0.50. Hence, our subject will make half of the responses to green and a quarter of the responses to each of the other buttons.

Why is this noteworthy? Assuming that the matching law is a basic fact of life (some dispute this), it tells us something interesting about the operation of the law of effect. That is, the effect of reinforcement of a given behavior is really predictable only if the effect of other behaviors and their reinforcement frequencies is taken into account. This is contrary to the common notion that we may select a target behavior and affect it independently of other behaviors; this view stresses the interdependence of behaviors.

This interdependence of behaviors and the relativity of the law of effect makes some sense of failures to find effects produced by many variations in reinforcement contingencies. For example, we believe that big reinforcers are more effective than little ones and that delay of reward is harmful. Yet, if you examine all of the studies seeking to show that this is the case, it becomes clear that the effects of magnitude and delay of reward are very unclear (e.g., Mackintosh, 1974). The rats running for the big pellet do not necessarily run faster than those getting the small one.

However, give the rodent exposure to both big and small pellets and it will quickly know about relative size. For example, the effect of a large reward is dependent upon the size of reward gotten a few minutes ago or concurrently available for other behaviors.

If the matching law still seems unrelated to things that happen in the real world, consider Baum's (1974) experiment. Oddly enough, Baum built the largest Skinner box in history by attaching a feeder and two response keys to the outside of his house. One perch was placed under the keys so that vagabond pigeons from the area could land, peck the keys, and be fed for it. Reinforcements were scheduled independently for pecking each key, and a group of twenty to thirty pigeons eventually came to congregate there. Now, the perch could accommodate only one bird at a time, so it is easy to imagine the flock of birds hovering around the feeder, successively displacing one another, with the key clicking, the feeder clanking, and the pigeons (clucking?) cooing. In spite of the bizarre conditions, the flock may as well have been one bird, because the relative frequency of pecking on the two keys closely matched the relative frequency of reinforcement for pecking them. This is therefore a powerful law, applying to the behavior of groups as well as individuals.

Self-Control

Rachlin and Green (1972) and Ainslee (1974) showed that one version of the matching law can account for self-control and went on to demonstrate this with pigeons. Their simplified version of the matching law was as follows:

$$\frac{V_1}{V_2} = \frac{A_1}{A_2} \cdot \frac{T_2}{T_1}$$

V_1 and V_2 refer to the relative values of two alternatives, A_1 and A_2 represent the amount of payoff for each, and T_1 and T_2 refer to the delays in obtaining the payoffs. Note that value is reasonably assumed to be proportional to the amount of payoff, and it is inversely proportional to the delay in receiving it. Thus, our valuation of an alternative depends on what we get and how long we have to wait. Since value depends on both of these factors, self-control occurs when we choose a larger payoff even though we have to wait for it. We save our money because of the large payoff we will receive in the future.

Time affects the relative values of alternatives so that it is easier to make a wise choice that commits us in advance of the actual choice between alternatives. If I find it difficult to save money, it is because the present value of spending it is greater than the value of accumulating it far in the future. Thus, given the choice, I spend it now. I may control myself by eliminating my choice: I sign up for the payroll savings plan and thus agree in advance to choose the wise alternative. This may be expressed as follows:

$$\frac{V_1 \text{ (spending)}}{V_2 \text{ (saving)}} = \frac{A_1 \text{ (use of small amount of money)}}{A_2 \text{ (use of large amount of money)}} \cdot \frac{T_2 \text{ (months)}}{T_1 \text{ (minimal)}}$$

With money in hand, V_1 is larger than V_2 and I spend my money. If I agree to begin the payroll savings plan in a month or two, I add that time period to both T_1 and T_2, thus neutralizing the discrepancy between them.

In illustrating this form of self-control in pigeons, Rachlin and Green (1972) presented the birds with two white keys. Pecking the right key changed the colors of the keys to red and green. Further pecks on the red key led to access to food for two seconds, and pecking the green key led to four seconds of food after a four-second delay. A wise bird would always peck the green key, wait for four seconds, and get twice as much access to food. But no bird did this. Given the choice between red (a small immediate reward) and green (a large delayed reward), every bird always pecked red, thus showing no self-control. Refer to Table 9.1.

If, when the keys were white, the left key was pecked, one key changed to green and the other went off. Only green could then be pecked, producing the large delayed reward. But no bird pecked the left key, though earlier forced trials ensured that the animals were familiar with the results of pecking either key. They always pecked the right key, giving them the choice of red or green, and then they always pecked red, giving them the small immediate payoff.

When the "payroll savings plan" was introduced, all of the birds eventually signed up. This amounted to interposing a delay before the keys turned to red and green. When the delay was greater than four seconds, the birds began showing self-control by pecking the left (commitment) key, which then (after four seconds) gave them only the option of pecking green (the wise choice). When the delay was as great as ten seconds, the left key was chosen much more often; the birds showed self-control.

TABLE 9.1 Two Examples of Self-Control and the Matching Law

1. $A_1 = 2s$ $T_1 = 1s$ $A_2 = 4s$ $T_2 = 5s$	No self-control $\frac{V_1}{V_2} = \frac{A_1}{A_2} \cdot \frac{T_2}{T_1}$ $= \frac{2s}{4s} \cdot \frac{5s}{1s} = \frac{10}{4}$
2. $A_1 = 2s$ $T_1 = 11s$ $A_2 = 4s$ $T_2 = 15s$	Self-control with 10s Delay $\frac{V_1}{V_2} = \frac{A_1}{A_2} \cdot \frac{T_2}{T_1}$ $= \frac{2s}{4s} \cdot \frac{15s}{11s} = \frac{30}{44}$

For people and pigeons this suggests that we always choose the higher-valued alternative and that value depends on both time and amount of payoff. I sign up for payroll savings in advance of receiving my pay, when the value of saving is greater than the value of spending, just as the pigeon "signs up" for the green key in advance. When I receive my pay, the value of spending is greater than the value of saving, just as the immediate small reward is of high value for the pigeon when it is faced with the red/green choice.

In both the human and the animal case, there is no real self-control; in both cases the choice is made for the higher-valued alternative *at the time.* Similarly, we place an alarm clock at a distance from the bed before going to sleep, when the value of getting up is high and the value of staying in bed is low. We do so because we know that the values may be reversed in the morning and that the value of staying in bed will likely be greater than the value of getting up. Are all instances of self-control reducible to such a principle?

Autoshaping: Relevant to the Operant/Respondent Distinction?

Many researchers have studied the behavior of pigeons pecking a lighted plastic response key. According to popular belief, one trains a pigeon to peck a key through shaping its behavior, using the method of successive approximations (see Skinner, 1953). This amounts to shaping behavior arbitrarily, "as a sculptor shapes a piece of clay." Not only key pecking may be shaped, of course, since we can easily shape all manner of behaviors. But most basic researchers in operant conditioning hand-shaped their birds to peck the key, often spending hours doing it with a difficult bird. Some birds just could not be hand-shaped, regardless of the skill of the shaper; these birds were kept in the uppermost cages in the bird room and were taken down only if one wanted to try (once again) to shape them.

Then Brown and Jenkins (1968) showed that the art of shaping was not necessary, at least in the case of key pecking by pigeons. They showed that all we need do is take any pigeon (even from the top row of cages) and induce it to eat from a grain hopper. In their experiment, they set up equipment to pair the lighting of the key with operation of the food magazine. The light came on for eight seconds, followed by the operation of the food magazine. The key light was then off for a variable period (such as a minute) and then was lighted again, followed by the delivery of food. Brown and Jenkins found that their 36 subjects all began pecking the key within 119 trials, with the average bird pecking on the 45th pairing of light and food.

If the birds pecked the lighted key, it immediately turned off and the food hopper operated. But pecking did not have to produce any effect. When the light stayed on for eight seconds even when the bird pecked it, eleven of twelve birds still began pecking, with the average bird beginning to peck on the 54th trial.

Brown and Jenkins noted that the procedure they used was essentially a classical conditioning procedure, with the key light acting as a CS and the food as the UCS. The reason that the birds pecked at the key was presumably the same reason that Pavlov's dogs licked a CS light if they could, though Brown and Jenkins suggested that it was due to a species-specific tendency for pigeons to peck what they look at. Finally, they suggested that this autoshaping could somehow depend on superstitious operant conditioning.

Williams and Williams (1969), however, showed that autoshaping was not due to superstitious reinforcement, and they did so in a simple and persuasive way. They used essentially the same procedure that Brown and Jenkins used, but with a single and important difference. If the pigeon pecked the lighted key, it prevented the operation of the food magazine on that trial. Thus, the key light was illuminated and food was delivered, unless the lighted key was pecked. In that case, no food came. Whether the birds involved had been previously hand-shaped, autoshaped, or were naive, they pecked! The key light went off and no food was delivered. Hence, the key peck in autoshaping is not maintained because of a superstitious connection with food; it persists even though it prevents food. Not surprisingly, it takes a number of light-food pairings, as is the case in autoshaping, before pecking begins; once it does begin and food is never received, it is apt to stop. This means that food will once more be delivered, which means that pecking will begin again, which will prevent the delivery of food, and so on.

Maybe pigeons just have to peck something when food is on the way, and if some alternative to the food-preventing key were available, the bird could peck harmlessly and still get food. The Williamses introduced a second key, which came on when the old key did but had no consequence when pecked. Reasonably enough, the birds quickly learned to peck this "irrelevant" key, thus having their pecking and their food too. But just any irrelevant key would not do; it had to signal food as did the other key. If an irrelevant key was left on all the time, except when the food was being delivered, it was seldom pecked; the birds went back to the old key and to their lost reinforcements. When they did peck this irrelevant key, it was always when the "relevant" key was on.

What do we make of this? The method of successive approximations still works, of course, even though it may be unnecessary if we wish to train a pigeon to peck a key for grain. But this case is clearly out of the reach of a standard operant analysis, meaning that there is something special about pigeons pecking for a food reinforcer. Is the key peck a CR, since autoshaping and automaintenance (shown by the Williamses) amount to classical conditioning and omission training, respectively? Recall from Chapter 3 that Sheffield (1965) used omission training to see whether CRs such as salivating were sensitive to their consequences. He found that dogs could not learn to withhold salivation when rewarded with food, showing that salivation was a true CR. Food is a UCS that elicits salivation, and salivation occurs even when it prevents the delivery of food. Since pecking occurs even when it prevents the delivery of food, is it a CR?

It seems that it clearly is, although debate raged on what seems to be such a simple matter. Pigeons eat by pecking at grain and other food items, followed by the ingesting and swallowing of the food. Just as salivation is a CR because it is a fractional component of the UCR, pecking is a CR because it is a fractional component of the pigeon's UCR to food.

Why is this finding damaging to traditional conceptions of learning? Well, if one takes Skinner's distinction between operant and respondent conditioning seriously (1938), and notes that key pecking has long been studied as a typical operant, the discovery of autoshaping is noteworthy. Thousands of studies have shown that key pecking is an operant, because it has been shown to be sensitive to its consequences more times than one can count. But autoshaping shows that key pecking acts as a CR, or respondent, as well. Many studies (e.g., Gamzu & Williams, 1971, 1973) have shown that autoshaped pecking occurs only under conditions in which Pavlovian CRs occur—that is, where there is a *contingency* between the CS and the UCS. Just pairing the key light and food is insufficient; the key light must predict the delivery of food, much as the CS must predict the UCS if classical conditioning is to occur.

Another reason for believing that the autoshaped key peck is a CR lies in the evidence that the form of the peck depends on the UCS, or reinforcer, involved. When the reinforcer is food, the autoshaped peck is made with beak open. When the reinforcer (or UCS) is water, the peck occurs with beak closed (e.g., Moore, 1973). It happens that pigeons peck at grain with an open beak and at water, if it could be called a peck, with beak closed. Thus, when autoshaped with food, the beak is open, as it is when the bird eats grain and when autoshaped with a water reinforcer, the (CR) peck is made with closed beak. The autoshaped response is thus a CR, since other CRs are fractional parts of the UCR to whatever UCS is coming.

Other evidence, cited in the scholarly review by Schwartz and Gamzu (1977), shows that when other organisms and reinforcers are used, the autoshaped response also resembles the normal response to the reinforcer. Thus, when access to a sexual partner is the UCS, the autoshaped response is similar to the sexual response, and when opportunity for aggression is the reinforcer, the autoshaped response is aggression. It has even been shown that pigeons may react to a stimulus preceding electric shock by striking at it with their wings (Rachlin, 1969).

Contrast and the Interdependence of Behaviors

Those favoring the operant model, both in behavior therapy and in basic research, have usually assumed that behavior is divisible into independent units. Hence, changes in the frequency of a target behavior may occur in the absence of changes in other behaviors, especially when the other behaviors seem quite different from the target behavior. In the same way, basic researchers have often confined their attention to whatever behavior operates the lever or key, assuming that other behaviors remain essentially unchanged (cf. Staddon & Simmelhag, 1971). If other behaviors do change when the target behavior does, the effect is usually attributed to generalization.

Recall that Skinner's early definition of the operant by no means restricted that unit to behaviors that appear similar in form or that conform to our notions of what a "behavior" should be (Skinner, 1931, 1935). An operant class may be composed of a number of behaviors that at first sight seem very dissimilar.

Wahler (1975) described data gathered from observation of children over several years and reported that changes in the frequency of a given behavior were often accompanied by changes in other behaviors that did not seem obviously related to the first. He referred to this as evidence for clusters of behaviors, which is another way of saying that the response class may be composed of quite different behaviors. They form a class because they tend to covary.

Another form of behavioral interdependence is quite different, appearing as a divergence or contrast among behaviors. Reynolds (1961a) showed that the frequency of one behavior may be greatly influenced by a change in the consequences produced by a second behavior. This effect, ***behavioral contrast***, has been found to be quite pervasive in animal learning and has clear parallels in human behavior, as Reynolds (1975) pointed out.

We will begin with an illustration with pigeons pecking a clear plastic response key. Initially, pecking is occasionally reinforced with food equally frequently when the key is lighted with either a red or green light. Some time later, reinforcement is discontinued when the key is green but continues at the same rate when the key is red. Responding in green decreases, of course, and it increases greatly in red, even though the reinforcement conditions in red are the same as always. Those conditions are better, however, relative to green, and the increase in responding in red is aptly called contrast.

Consider Figure 9.2. The left panel of the top figure shows responding to two stimuli, S+ and S−, when the frequency of reinforcement in each is equal. The right panel shows what typically occurs when the frequency of reinforcement in one stimulus is decreased, making S− a real S−. Response rate in S− decreases, of course, but the rate of response also increases in S+; why should this be? It would seem that only a very odd sort of person would find this occurrence strange. After all, is it not plausible to assume that the value of S+ increases because it is now associated with more frequent reinforcement relative to S−? Aren't our evaluations of things always relative? Is five dollars the same thing to me as to a Vanderbilt? Why should it be different for pigeons, rats, monkeys, children, and all of the other animals in which behavioral contrast has been observed?

This was essentially Reynolds's explanation, and it is essentially correct. He showed (1961b) that contrast occurs when a stimulus presented with VI reinforcement is alternated with time-out periods (no stimulus on the key and no reinforcement available for responding), or with extinction. But contrast does not occur when a VI schedule is alternated with a DRO schedule. A DRO (differential reinforcement of other behavior) schedule provides reinforcement for not responding. Such a schedule may provide as frequent reinforcement as a VI schedule, but rates of responding are near zero. This showed that a decrease in reinforcement frequency in S− was necessary for contrast; just a reduction in response rate is insufficient.

Local Contrast If you recall Pavlovian induction (Chapter 3), you will immediately recognize *local contrast* as the same phenomenon. Nevin and Shettleworth (1966) showed that during the formation of a discrimination responding in S+ and S− is influenced by the immediately preceding stimulus. The effects they found (and called transient contrast) are illustrated in the lower part of the figure.

The lower half of the figure illustrates local contrast occurring between presentations of two stimuli, S+ and S−. However, the data here

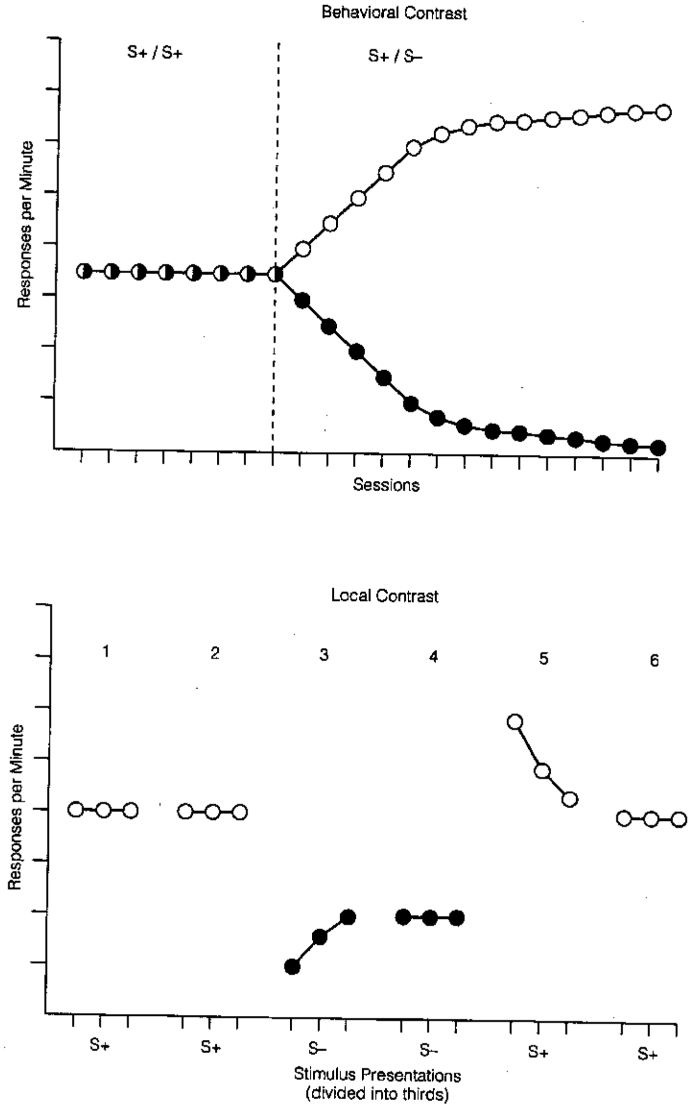

FIGURE 9.2 The upper panel shows behavioral contrast, a change in response rate appearing over sessions. An increase in responding in S+ occurs when reinforcement is discontinued in one stimulus, which becomes S−. This is positive behavioral contrast. Under other conditions, negative behavioral contrast occurs. The lower panel shows positive and negative local contrast. This effect appears among individual stimulus presentations. See text for details.

represent responding in stimuli as a function of the immediately preceding stimulus; typically, data show responding during successive segments of a stimulus presentation (that is first third, second third, third third). For example 1 to 2 in the figure show responding in S+ (2) when preceded by itself (1) averaged over one or more sessions. Responding in the two S+ presentations is similar.

On the other hand, 2 to 3 represents the presentation of S− (3) on those occasions when it was preceded by S+; note that responding is initially depressed and increases during the presentation. This is negative local contrast. Instances 4 to 5 show the effect of an S− (4) preceding an S+ (5); note that responding is initially elevated and decreases during the presentation. This is positive local contrast. These effects may last only a few sessions (e.g., Malone, 1976; Nevin & Shettleworth, 1966), as Pavlov (1927) often found to be true of Pavlovian induction. But if many stimiuli are used and/or they are very similar and therefore difficult to discriminate, local contrast may persist over months (Malone & Staddon, 1973). As far as is known, local contrast and Pavlovian induction share all important characteristics (Malone & Rowe, 1981).

Contrast and Behavior Therapy In the human case, contrast occurs, for example, when unruly behavior is maintained, if unintentionally, by consequences it produces on the playground and at home. In both cases, attention (at the very least) reinforces such behavior, just as the pigeon's pecks were reinforced in both red and green. Suppose that a teacher or behavior therapist manages to eliminate those consequences on the playground and that this leads to a decrease in unruly behavior there. Like the pigeon's pecks, which are still reinforced in red, unruly behavior will likely produce its customary consequences at home and one should therefore expect an increase in unruly behavior at home (Reynolds, 1975).

Reynolds's suggested parallel between pigeon and child is by no means farfetched, although at first sight it may appear to be. Wahler at the University of Tennessee Child Behavior Institute, has been interested in this parallel. He has noted (personal communication, 1989) that reports from parents of increased troublesome behavior in the home often accompany successful treatment of such behavior in the school. If we wish to decrease or eliminate the pigeon's pecking in red, we stop providing reinforcement in red. So too with the unruly child. Wahler, Winkel, Peterson, & Morrison (1965) found that the real method for dealing with unruly behavior is through training of parents as "therapists" so that the inadvertent reinforcement of such behavior in the home is ended.

This contrast among behaviors is not always the rule, of course, and behavior therapy is largely predicated on the assumption that treatment effects generalize to a greater or lesser degree across situations. But when might one expect such generalization and when might contrast be expected? There are cases in which treatment of troublesome behavior at school generalizes to behavior at home. Yet, there are cases when behavioral contrast seems to occur. The answer to this question is unknown at present. All we can turn to are the animal studies that have been done to date, and there have been many; but they have been unable to determine the conditions that lead to generalization or contrast. It is probably safe to say that if the behaviors in question (key pecks or troublesome behaviors) belong to the same response class, they are apt to covary in quite similar situations. (That is actually the definition of a response class.) Also, the more similar the situation (red/green, school/home), the more we may expect generalization of treatment effects from one situation to the other. The increase in the child's misbehavior at home depends in part on the dissimilarity of the school and home situations as well as the altered consequences for unruly behavior at school. This is not terribly helpful, but it is helpful to know that there is considerable evidence for contrast effects in animal learning and thus we should not be overly surprised when we find similar effects in analogous situations involving human behavior.

We might also know more about the causes of contrast if it were not for the great reluctance of basic researchers to accept such interactions among behaviors. It is more convenient to treat them as independent. One researcher (Bloomfield, 1969) went so far as to suggest that behavioral contrast shown by pigeons represented the avian analog of human psychopathology! That is, our pigeon pecking along in red had no right to peck faster in red when reinforcement was eliminated in green. Reinforcement conditions were unchanged in red and response rate should have not changed; hence, the bird was insane.

Neuringer and Chung: Conditioned Reinforcement

It has long been believed that any event that acts as a reinforcer of the behavior of a given species does so for one of two reasons. First, it may act as a reinforcer because of the nature of the organism, in which case its power would be evident at birth or shortly after. Warmth, food, water, and other *primary reinforcers* have obvious biological utility and their deprivation quite rightly gives them reinforcing power.

But few of us spend our lives as do most animals, constantly governed by needs for food, water, and other primary reinforcers. In fact, much of our behavior goes on in spite of such needs, especially when we are concerned with goals such as great achievement, riches, or fame. How do such "higher goals" gain their reinforcing power? We do not all wish fame or riches or great achievement. Often it is solitude and peace that seem the more powerful reinforcing agents. Yet, we do all like the attention of others; being ignored by all would be unpleasant indeed. How does attention from others gain reinforcing power?

The argument can be made that all events that act as reinforcers but do not act as such at birth (or shortly after) derive their power through association with primary reinforcers. This is a very popular view and, logically, is very hard to fault. We saw in Chapter 6 that this view was strongly criticized by both Allport and Harlow. Nonetheless, it has remained the leading interpretation of acquired reinforcement. For example, consider the infant whose motives are restricted to hunger, thirst, warmth, contact, and other primary reinforcers. At present our infant is hungry, thirsty, cold, and bored. An adult human appears and provides food, warmth, and entertainment, all of which act as reinforcers for the cry or whatever other behavior brought the adult. This scene happens over and over in the infant's early life and it is no surprise when someone suggests that the infant comes to associate the adult with satisfaction of basic needs. The adult acquires reinforcing power and thus the tale begins.

The adult indeed signals impending food, warmth, comfort, and entertainment, but only under certain conditions. An adult sitting quietly reading or knitting provides no such comforts; it is only when the adult's attention is directed toward the infant and the adult notes the infant that primary reinforcers are forthcoming. Thus, attention itself becomes associated with adults and primary reinforcers and gains reinforcing power that lasts throughout our lifetime. Later, it becomes clear that attention alone is often not enough to guarantee primary payoffs; approving, praiseful attention predicts food and the like more reliably. So we add approval and praise to the list of secondary reinforcers, the power of which still depends on the linkage with food, water, warmth, and so on.

Even curiosity and manipulation may be added to the list, thus dealing with Harlow's critique. When a cat or rat or monkey presses a lever for the opportunity to explore a maze or manipulate a puzzle, we are still dealing with acquired reinforcers ultimately dependent on associations with primary reinforcers. Is it not true that in the early life of any organism exploration of new surroundings is often rewarded with the discovery of food, water, and shelter? Although Harlow (1953 a, b) and others make a good case against such an account, the fact is that it makes a lot of sense. We are biological entities and life is basically the fulfilling of basic needs.

This general interpretation has appeared in so many texts, in so many forms, and has such a surface plausibility that it seems foolish to question it. Nevertheless, there is considerable evidence against it, derived both from daily life and from experience.

In brief, there is remarkably fragile evidence for the establishment of new reinforcers through association of some sort with already potent reinforcers. Gollub, who has spent many years concentrating on the nature of acquired reinforcers, comments on the irony of the situation at the conclusion of a very comprehensive review of basic research in the area (1977). The fact is that the aggregate of laboratory research shows that the pairing of new stimuli with current reinforcers lends only the feeblest reinforcing power to those stimuli. Yet, we see daily how powerfully tokens (including money) control behavior. How does that come to be?

What is bizarre here is not that tokens work but the belief that their power must somehow be based on real reinforcers and that acquired reinforcers are by their very nature of no value. Every time a child is reluctant to turn in tokens for candy, and every time we work on a term paper while tired and hungry, the plausibility of the traditional account is called to question. Isn't it odd that the acquired reinforcers, or *conditioned reinforcers*, in these cases seem more powerful than the primary reinforcers upon which their power is supposed to rest?

The truth is that new reinforcers are acquired and that they may have something to do with current reinforcers. But it is not the case that an acquired reinforcer acts as such simply because it has become a predictor of food, water, warmth, or other real and biologically reasonable events. The following is a very speculative outline of an alternative interpretation of acquired reinforcement. It is not without its problems, but see if it is not more reasonable than the classic explanation (or no explanation at all).

Neuringer and Chung (1967) found that new stimuli, such as brief darkness or tones, could act as potent reinforcers when presented in a specific way. In the simplest case, they reinforced pigeons' key pecks according to a variable-interval (VI) schedule of reinforcement. On the average of once a minute a key peck was followed by the delivery of food.

After baseline levels of responding were established, conditions were changed so that each eleventh response produced some consequence. In some conditions the consequence was a brief tone, in others the chamber was briefly darkened, and in others the key light was briefly turned off. The exact nature of the consequence did not matter; each eleventh key peck produced something. So the birds were getting food on the average of once a minute, and every eleventh peck produced some other brief consequence. What did the addition of such consequences do to the rate of pecking for food? Nothing, of course. Why should a food-deprived pigeon, pecking a key for occasional food, be at all mindful of occasional and meaningless added stimuli?

But a slight alteration in conditions did produce impressive results. Food delivery was scheduled by a timer that determined which pecks would be followed with food. Suppose that food was delivered only after a series of eleven pecks. That is, the bird would peck along, receiving a brief tone after every eleven pecks; there would be eleven pecks then a tone, eleven pecks then a tone, and then after two or three or six pecks, the timer would schedule a reinforcement. But the bird would have to complete its eleven-peck sequence before the food was delivered. In this way, something always happened eleven pecks from the last tone or food and never after more or less than eleven pecks. Eighty-five percent of the time eleven pecks produced only a tone, but fifteen percent of the time, they ended in food. Response rates almost immediately doubled. Why should the addition of a seemingly irrelevant consequence produce such an effect?

Here is a hint. In control procedures, birds receiving the same pattern of food and other consequences showed no change in response rates. For example, subjects receiving brief blackouts or tones after every eleventh response but receiving food for the first response after the timer scheduled reinforcement showed no change in response rates. Thus, an added irrelevant stimulus produced large changes in responding when it was produced in exactly the same way that food was produced—namely, by eleven responses. The same effect appeared when a fixed-interval-like requirement producing food or other consequences was added to the variable-interval schedule of food delivery.

Here is an intuitively plausible explanation of acquired reinforcement that does not depend upon the role of neutral events as predictors of events that already act as reinforcers. The brief consequences here really predict no food, at least for the short run, but they did confirm the effectiveness of responding in a certain way; something always happened after eleven responses and sometimes that something was food delivery.

Could it be that tokens, money, praise, and perhaps scolding and pain may come to act as acquired reinforcers if they are produced by the same behaviors that have in the past produced already effective reinforcers? When an infant is good, he or she is petted and stroked; later, the same behavior produces attention and praise, and later it produces a quarter. If hard work, formerly reinforced by money and praise, were to begin to be followed by hallucination of a green mist, wouldn't the significance of such experiences soon change for us? Needless to say, hard work must on other occasions be followed by more customary consequences, but doesn't the hallucination correspond closely to the feeling of a job well done? This explanation, speculative as it may be, attempts to account for the heretofore mysterious durability of conditioned reinforcers and their seemingly infinite variety without requiring that their power ultimately reduce to the prediction of food!

For practical purposes, the implications are obvious. Almost anything, be it a word, a token, or the opportunity to do something, may become an acquired reinforcer if it is made available according to the same contingencies that produce already effective reinforcers. Serious studying by a high school student may be reinforced both by good grades on an examination or by parental praise. Given such a history of payoff, other consequences may be substituted. "You've done such a good job with your homework that I'm going to let you rake the leaves for a while tomorrow." Imagine such a reward given seriously by a parent who normally gives only praise for such behavior and who has not used the opportunity to rake leaves as a form of punishment in the past. Would such an opportunity act as a reinforcer for doing homework?

Avoidance Learning

Suppose that a rat or pigeon or human subject were placed in a situation in which shocks occur every ten seconds. In the chamber is a lever (or button or key) that, when pressed, postpones the next shock—say, for a minute. Suppose also that a light comes on a few seconds before an impending shock. Most subjects, whether rodent, avian, or human, learn to press (or push or peck) when the light comes on and thus avoid the shock.

This is called signaled avoidance conditioning. A signal (CS or S^D) precedes shock, and a response turns off the signal and postpones the shock. If a subject responds at least once per minute (in our example), then it is never shocked. Rat and pigeon subjects do so, and the question arises: What reinforces the pressing or pecking response?

Early explanations (e.g., Mowrer, 1947; Miller, 1948) suggested that the signal preceding the shock became a conditioned aversive stimulus (CAS). Through classical conditioning, the signal became a CS, producing fear; the avoidance response, which turned off the signal and avoided the shock, was reinforced by the ending of the signal.

However, avoidance responses are learned even when there is no signal preceding shock. Sidman (e.g., 1953) showed that when shock is delivered at regular intervals, rat and pigeon subjects will press or peck to avoid them, even when there is no explicit signal preceding the shock. He suggested (after Schoenfeld, 1950) that this occurs because all other behaviors are eventually paired with shock and the only nonpunished behavior is pressing or pecking, which avoids shock! Surely, this explanation is one of the curiosities of psychology.

Still, why *is* the avoidance response made? Imagine yourself in a room with no features, but a large red button mounted on a wall. Every ten seconds you receive shocks that leave you trembling. Before long you would hit the red button and find that you weren't shocked for a while. Soon, you would be pressing at least once a minute and would never receive any shocks. Would you continue pressing for hours, days, and weeks after the last shock? Probably not. You would occasionally slack off for a minute, which would be long enough to remind you that shock comes when the button is not pressed for a minute.

But theorists wondered what reinforced the avoidance response. You press the button and nothing happens. (You or I might say that it is not *nothing*—a nonevent—per se; it is nothing in the context of *something*—a shock when no press is made.) This thinking, which requires that every discrete response is reinforced by a discrete event, is typical of much psychological thought. Freud's intrapsychic forces were invented to provide present causes for effects where the real causes happened years ago.

At any rate, Anger (1963) provided a present cause for avoidance responding. He suggested that, as time passed since the last shock or response, the subject "timed" the interval. As time passed, temporal stimuli became more aversive, since time elements that were more and more distant from a response or shock were more closely paired with shock. Thus, the passage of time includes temporal stimuli, which become increasingly aversive as the time for shock approaches. These conditioned aversive temporal stimuli (CATS) act as a drive to produce a response that "resets the clock" and reduces the drive. This theory, proposed by a follower of Skinner who thought in the discrete response/consequence terms of Hull, thus explains avoidance learning. The subject continues to make avoidance responses to escape the CATS occurring with the passage of time.

This two-factor theory, like the earlier view of Mowrer, explains unsignaled avoidance as it explains signaled avoidance. One just has to look harder for the signal. The subject responds, the signal is turned off (or reset, in the case of CATS), and the reduction in conditioned fear thus produced reinforces the avoidance response.

This theory, which was popular for many years and which required something to reinforce the avoidance response, underestimates the importance of contingencies acting over time, which is the emphasis of good Skinnerian theory. This was pointed out in an excellent experiment done by Herrnstein and Hineline (1966).

In their experiment, rats received shock according to a variable-time (VT) schedule with an average value of 6.7 seconds. That is, shock was delivered to the floor (and walls) of their chamber every 6.7 seconds, independent of their behavior. Or almost independent of their behavior. If they pressed a lever in the box, a different VT schedule was in effect, which produced shocks less frequently (every twenty seconds). When the shock came, and it did, the schedule was changed back to every 6.7 seconds.

Whatever the rats did, they were shocked. If they did not press, it came every 6.7 seconds (on the average), and if they did press it came on the average of every twenty seconds. Imagine yourself in their position, in a small room with a lever on one wall and intense shock delivered on the average of every few seconds. You press the lever and the schedule changes from shock every 6.7 seconds to shock every twenty seconds. But, depending on the tape programming the VT 20-second shocks, you could press the bar and immediately get a shock. How long would it take you, given no instructions, to learn to press the lever? You still get shocked, but *in the long run* you get shocked less if you frequently press the lever.

Rachlin (1976) compared the subject's dilemma to that of a New Hampshire resident who moves to Florida. If it's snowing on the day the person leaves and if the person arrives in Florida during a freak snowstorm, there is apparently nothing to distinguish the weather in the two states. Prior knowledge or time spent living in the two states would be required for the person to know that weather in the two states is different in the long run.

But seventeen of eighteen of Herrnstein and Hineline's rats showed stable lever pressing on the first day. Later, when the frequency of shock was made equal on the two tapes, responding eventually ceased. Thus, pressing was not merely shock-elicited flailing at the lever; it was maintained only when overall shock frequency was reduced by pressing. Lambert, Bersh, Hineline & Smith (1973) later showed that rats also learned to respond for reduction in overall shock frequency, even when each response produced an immediate shock.

What reinforced each press? Is that a reasonable question to ask? The fact is that the rats' behavior was best described as *lever pressing*, not as *lever presses*. That is, pressing was an activity that occurred over time and that was not best described as a set of discrete instances (that is, as presses). Pressing was reinforced by reduction in overall shock frequency; a behavior occurring over time depended on consequences occurring over time.

Azrin and the Effects of Punishment

It requires little reflection to see that punishment and the threat of punishment is the main instrument whereby society controls us. The law is a set of rules specifying which acts are punishable, not what will be rewarded. Can you imagine being stopped by the police and congratulated for driving safely? Parents spank their children and send them to bed, and teachers keep misbehaving students after school or force them to miss recess. The church tells us of heaven, but many denominations prefer to emphasize the horrors of hell. We are especially influenced by the punishments of the natural world around us, by the pain that punishes fire touching and the bruise that punishes slipping and falling.

We cannot deny that the schools, our families, the church, and nature itself do not reward us, but punishment also is used at least as often (especially by government) and has been for a very long time. With this in view, it may be surprising to learn that expert opinion has denied the effectiveness of punishment and opposed its use for many years.

You will recall that Thorndike decided, over the period from 1929 to 1932, that annoyers did not work to stamp out S-R connections, as proposed in his original rendition of the law of effect. Punishment, when it works at all, does so simply because it evokes a new behavior that can then be rewarded. Otherwise, "it tends to work best where it is least needed," leaving the "callous and base" unaffected, and it simply adds "fear and shame" to continued sinning. If punishment is to be used, Thorndike urged, we must at least avoid "doctrinaire, haphazard, fantastic, and perverted" punishments (1935).

Skinner (1938) made the same argument a few years later, pronouncing that punishment was generally ineffective, its influence on behavior was transient, and it produced unwanted by-products. This was substantiated in an experiment in which rats were first trained to press a bar, with occasional presses reinforced by the delivery of a food pellet. Such training was followed by periods of extinction, in which food delivery was discontinued until responding ceased. After a number of changes from reinforcement to extinction, Skinner assessed the effect of punishment on the rate of decline of responding during extinction.

After a period of food-reinforced pressing, extinction began and each press was punished in an interesting way: A device was connected to the mechanism controlling the pivoting of the bar in such a way that when the bar was pressed it pressed back, giving the rat a sharp slap to its paws. The slaps were discontinued after the first ten minutes of extinction. Skinner found that such punishment did indeed reduce the rate of bar pressing, but once it was discontinued pressing recovered, increasing to levels higher than those that occurred in the absence of punishment. It was as if the presses suppressed during punishment were "stored up" and appeared as soon as punishment ended. Other and less formal experiments used other punishers, such as throwing the rat into the air, pricking its tail with needles, and shining bright lights into its sensitive eyes. Again, the effect of punishment was only temporary, and suppressed responses seemed to be stored, augmenting response rate once punishment ceased. Additionally, the rats' behavior suggested that the punishers had emotional effects, and Skinner maintained that this could be sufficient to disrupt pressing enough to account for the decrease in response rate. (The rats were also less kindly disposed toward Skinner himself, who had become a conditioned aversive stimulus because of his association with aversive events!)

Sturdier support for these conclusions was provided by the data of Estes (1944), who followed Skinner's general procedure but used shock as a punisher. Estes found that mild shock produced the same effects reported by Skinner, a temporary suppression of responding followed by an increase after the termination of shock. If the shock were intense, responding was greatly suppressed over several postshock sessions. Even then the rat eventually recovered to the prepunishment level. Are the effects of punishment only temporary, as Skinner suggested?

Perhaps the most interesting aspect of Estes's data was the finding that the suppressing effects of shock were the same whether it was response dependent or not! *Punishment* means response dependent, of course. If shock presented independently of responding also suppresses behavior, we can hardly call it the effect of punishment. It amounts only to the general disruption in behavior, which could be produced by strong emotion, such as fear of shock. Thus, the decrease in responding was merely a by-product of emotional disruption produced by aversive stimulation (shock), exactly as Skinner had argued.

These data have had great influence for decades: I frequently have been asked for a good source that cites the best evidence showing that punishment doesn't work! Indeed, a special term was even invented for cases in which punishment did seem to work. Think about the possible meaning of *passive avoidance!* Active avoidance refers to behavior that is maintained because it prevents or postpones the occurrence of a noxious event. *Active* means that something is done; an act is performed. So what is passive avoidance? Is it the prevention or postponement of a noxious event by remaining passive? For example, when I placed my hand in a fire in the past (as a much younger person), it burned me. I now passively avoid such noxious events by keeping my flesh out of the fire. Put in a slightly different way, pain has punished the act of reaching into fires; I therefore no longer reach for such things. When punishment was banished, passive avoidance became its surrogate (Catania, 1984).

Happily, later data have come to the rescue of common sense; punishment is back. The credit for this belongs to Nathan Azrin and his coworkers, whose data led an authority on learning processes to conclude that the effectiveness of punishment now "admits of little doubt" (Mackintosh, 1974). Given the prevailing expert opinion at the time he was working (circa 1960), Azrin had to prove his point convincingly.

Basically, Azrin questioned the methods used by Skinner and Estes. When Skinner and Estes introduced punishment, it was done in extinction. Thus, the first shock told Estes's rats that further presses would be shocked and that there was no reason to press anyway, since food was no longer forthcoming. This surely is not the case in daily life. When a child is punished for bullying others, the punished behavior is maintained by the reinforcements it produces, such as the goods and services and fear and deference that come from those who are bullied. Occasional punishment is simply added to the consequences that already occur. When speeding is punished by a costly ticket, the reinforcers that maintain speeding are still available; speeding may be fun and it does get you where you are going in less time.

A more realistic test of the effects of punishment would apply it to behavior that is simultaneously maintained by reinforcement. Using pigeons as subjects, Azrin (1956) found that electric shock produced by key pecking suppressed pecking maintained by a schedule of food reinforcement. When shock was applied independently of responding, he found, unlike Estes, that responding was reduced relatively little; response-produced shock (that is, punishment) had a far greater effect.

In 1960 Azrin showed that the degree of response suppression depended upon the intensity of shock and that with fairly intense shock the suppression was long-lasting. In addition, the pattern of responding maintained by the schedule of food reinforcement was not disrupted. Thus, the decrease in response rate was not due simply to the generally disruptive emotional effects posited by Skinner and Estes. For example, Azrin (1959) showed that shock produced for every response maintained by a fixed-ratio schedule of food reinforcement left the fixed-ratio run (see Chapter 8) intact but produced an increase in the pause after reinforcement. This is exactly what happens when the ratio requirement is increased and the frequency of reinforcement therefore is decreased. Punishment thus seems to act in the same manner as a decrease in reinforcement frequency.

All of this shows that punishment does work directly; it is not effective only through emotional by-products. But Skinner and Estes also had charged that the effects of punishment were only temporary; in their view, when you stop the punishment the punished behavior returns stronger than ever. This is in fact the case, but it is hardly an argument against the effectiveness of punishment. Aren't the effects of reinforcement temporary? If my behaviors that are currently maintained by money, praise, feelings of health, and not falling down are no longer followed by such reinforcers, will I indefinitely continue to show such behaviors?

The one argument against the use of punishment that remains does warrant consideration. That is, punishers are often things that cause pain or physical injury, and thus their misuse is more serious than is the case with reinforcers. Also, as is the case with reinforcers, punishers work if applied immediately and appropriately (remember Thorndike's term *readiness*). The warning given a misbehaving child that he will "get it when his father comes home" is poor practice indeed. When his father comes and the misbehavior is long forgotten, a spanking may punish something but it will likely be the behavior of running up to and greeting his father! Further, as Guthrie made clear, punishment works best when it leads to a new behavior that can then be reinforced. We cannot quite follow Guthrie's advice when he points out that in training a dog to jump through a hoop using electric shock, "it is important at which end of the dog the shock is applied," but we can encourage acceptable, reinforceable behaviors when we punish the ones we want eliminated (e.g., Whiting & Mowrer, 1943; Herman & Azrin, 1964).

2. CLASSICAL CONDITIONING

Contiguity and Contingency: The Rescorla-Wagner Model

Since the 1920s, when classical conditioning was first introduced in America, American researchers and theorists have viewed it as a mechanical stamping-in process. A stimulus, such as a tone, becomes associated with another stimulus, such as food, because the two occur together. That is all there is to it. Pavlov often referred to the conditioned stimulus as a *signal* for the UCS, and the fact that it must act as such, rather than be merely paired with a UCS, was appreciated a half century after Pavlov's work became known. To be a signal, a stimulus must provide information, so to speak. There must be a contingent (if/then) relation between conditioned and unconditioned stimuli.

The truth of this was established by Rescorla (1967, 1988) who showed that the mere pairing of CS and UCS did not guarantee that classical conditioning would occur. For example, let us suppose that a brief tone is paired with the delivery of shock to your ankle. After a few pairings, the tone will take on a new meaning for you and, should the new meaning escape you, it will not escape the notice of your viscera! The tone will make you wince, move your ankle, increase your heart rate (and later, decrease your heart rate), and perhaps cry out. The pairing of tone and shock has produced a CR of fear, along with autonomic arousal and other anticipatory reactions. Suppose that you are left on a schedule of tone/shock pairings for some time and that CRs reliably occur, which would be the case. Further, suppose that extra tones that are not accompanied by shock are added to that schedule. You will still react to the tone, but less vigorously and reliably; the tone is no longer an absolutely reliable predictor of shock. Suppose next that extra shocks are added to your schedule, occurring in the absence of a warning tone. Your *Nos* and *pleases* may still occur but no longer in reaction to the tone, only to the shock. Your conditioned winces, heart palpitations, and so on will also fail to occur when the tone comes on, because the tone is no longer a CS. It tells you nothing about the occurrence of shock, and both you and your viscera see this. The original set of pairings of tone and shock remains as before, but the added tones and shocks have taken away the once predictive relationship. Classical conditioning requires a contingent relationship between CS and UCS; mere pairing is not enough.

A second, and related, effect has been studied by Kamin (1968) and a number of others. Kamin drew attention to what is known as *blocking*, which is related to the phenomenon of *overshadowing*, originally described by Pavlov. If in the situation above we maintain the tone/shock contingency so that the tone always precedes shock and also reliably produces unpleasant CRs, what will happen if we add a flashing light, so that light and tone always precede shock? You will continue to react as you did and, if we present the tone alone, we will find that it still has its power. But if we present the light alone there is no reaction! How can this be? The light has been paired with shock, as was the tone; should it not also produce fear and trembling? The answer is no and the reason is that the tone already tells us all there is to tell about the shock; the added light adds nothing. But why should this be? Why should the tone alone be an effective CS while the light, which also predicts shock, fails to become one? We shall see.

Overshadowing may provide a clue. If, strapped in your chair with the shock-giving electrode on your ankle, your initial experience involves the onset of a flashing light and the onset of a tone followed by shock, will both stimuli become capable of evoking equally strong CRs? Chances are that they will not. One of the stimuli will overshadow the other. One way of putting this is to say that one will be more salient. That stimulus, be it light or tone, may gain most or all of the CR-evoking power. This need not be, of course, since the tone and light could be more or less equally salient and thus share the CS power, but in many cases one or the other stimulus overshadows the other.

All of this caused a good deal of unrest during the late sixties; it appeared that classical conditioning was a complicated and less predictable process than had been believed. If it depends upon contingency rather than simple pairing by *contiguity*, and if a stimulus becomes a CS only if it provides information (is nonredundant), then classically conditioned organisms become contingency-evaluating, information-processing entities—a horrible prospect, since it means that what was simple and lawful is now vague and unpredictable!

Luckily, this problem was averted by Rescorla and Wagner (1972), who proposed a model for classical conditioning that has become extremely popular. We refer to it as the *Rescorla-Wagner model.* Taking their cue from a little-known paper of Hull's, published in 1947, Rescorla and Wagner showed how what seems to be contingency and informativeness can be reduced to simple pairing by contiguity. Their model is simple and reasonable, and it works.

They begin with the often-made observation that learning acquisition curves are negatively accelerated. That is, with practice we learn most on the early trials and progressively smaller increments are added as practice continues. In classical conditioning, the increase in magnitude and reliability of the CR increases rapidly at first and more slowly thereafter until asymptote is reached—until no further increase is possible unless conditions are changed (for example, by increasing the intensity of the UCS). Hull called this maximum *M*; Rescorla and Wagner call it λ (*lambda*).

On each reinforced trial improvement occurs in ever-decreasing increments. Hull suggested that the amount of increase (in habit strength, or ${}_{S}H_{R}$) is a constant fraction of the difference between current habit strength and M, or maximum habit strength. An early estimate of this fraction was $\frac{1}{25}$. Thus, on a given trial, the increment added to what is learned is $\frac{1}{25}(M - {}_{S}H_{R})$. As ${}_{S}H_{R}$ increases and M remains fixed, the quantity $(M - {}_{S}H_{R})$ decreases and the added increments decrease in magnitude. This may not seem fascinating, but it is a reasonable way to account for the progress of learning as long as one doesn't worry excessively about the problems involved in calculating initial values of ${}_{S}H_{R}$ and the value of M.

Rescorla and Wagner adopted this model with a change in terms, as follows:

$$\Delta V_{n+1} = K(\lambda - V_n)$$

Thus, the increment (ΔV) added to associative strength (V; habit strength is out) on the next trial ($n + 1$) is a fraction (K) of the difference between asymptotic associative strength (λ) and current associative strength (V_n). *K* represents the salience of the CS and falls between zero and one, and lambda depends upon the intensity of the UCS. To see how this accounts for the negative acceleration in the slope of the acquisition curve, consider the following series of three trials, letting *V* be set initially at 10, λ at 100, and *K* at 0.5:

1. $V_1 = 0.5(100 - 10) = 45$;
   V_2 is then $V_1 + V_1 = 55$
2. $V_2 = 0.5(100 - 55) = 22.5$;
   V_3 is then 77.5
3. $V_3 = 0.5(100 - 77.5) = 11.25$;
   V_4 is then 88.75

As is clear, the increments added to V become progressively smaller (45, 22.5, 11.25) as V increases (10, 55, 77.5).

There is nothing new here, but the next step is what counts and is what derives from Hull's 1947 paper. Rescorla and Wagner propose that all conditioning is actually compound conditioning; that is to say, associative strength develops not only to the CS but to the other stimuli in the background. This contributes to a total V and thus influences the rate of conditioning to a given CS. Given two explicit stimuli, the effects are more obvious.

First, take our light and tone paired with shock and suppose the light to be blindingly bright, near, and salient as can be, while the tone is faint, far, and less salient, though clearly noticeable. To examine the way in which conditioning affects them, it is necessary to consider them separately, using the equation above. Although they are considered separately, the fact is that they do appear together as a compound and it is their joint associative strength that is then important. This is the real point of the model and it is the joint associative strength (V_{ax}) that allows it to work its wonders. Remember that the light was more salient than the tone, so its K fraction will be greater. Let the initial associative strengths of the light and tone (L and T) be 10 each and let their saliences be 0.8 and 0.1. Lambda is still 100, and the

joint associative strength of the two stimuli is represented by V_{LT}. Here is what happens in a few trials:

$$\begin{aligned} \text{Light: } V_L &= L(\lambda - V_{LT}) \\ V_L &= 0.8(100 - 20) = 64 \\ \text{Tone: } V_T &= T(\lambda - V_{LT}) \\ V_T &= 0.1(100 - 20) = 8 \end{aligned}$$

After one trial, V_L is 74, V_T is 18, and V_{LT} is 92 (almost asymptote).

$$\begin{aligned} V_L &= 0.8(100 - 92) = 6.4 \\ V_T &= 0.1(100 - 92) = 0.\ 8 \end{aligned}$$

After two trials, V_L is 80.4, V_T is only 18.8, and V_{LT} is 99.2.

It can be seen here that the bulk of the associative strength goes to the light, and simple changes in parameters (such as setting V_L and V_T at initial values of 1 or 2, rather than 10) would result in V_L totally swamping V_T. This is the way in which the model accounts for overshadowing, and it is easy to see that it accounts for blocking as well. Pretraining with the tone alone, for example, would produce a high value of V_T, so that later introduction of the light (salient though it may be) would be to no avail; the tone could have already grabbed all of the V to be had.

The model also accounts for the effects of contingency noted above. It simply depends upon where the associative strength goes, to the CS or to the background stimuli. For example, if you introduce extra shocks, you introduce new conditioning to these background stimuli. Since there is a maximum of associative strength for any reinforcer (that is, any given frequency or intensity of shock in this case), the background stimuli compete with the tone. If the background stimuli acquire enough associative strength, the tone may have no more strength than the background stimuli have. In this way, it loses its function as an effective CS. This accounts for the effects of contingency, leaving the important factor as the simple contiguity of stimuli and UCS.

Solomon and Corbit: The Opponent-Process Model

The opponent-process model tells us something important about classical conditioning and about a variety of familiar phenomena, from risk taking to drug addiction to the basic categories of emotional experience. The major paper describing the model was published in 1974, but the real story began 30 years earlier, in the strange findings of Solomon, Kamin, and Wynne (1953), who were investigating methods for the curing of *phobias.*

These authors acted on the belief that unnatural fears of spiders, darkness, and so on, which greatly trouble many people, are simply CRs established earlier in life. For example, a spider present during the experience of great pain or fear produced by some other cause may become a CS producing fear. Their strategy was to establish a conditioned fear in dogs and then to investigate various methods of abolishing ("curing") it. Unlike a phobic reaction, the dog's fear was based on very reasonable grounds.

The dogs were placed (one at a time) in a shuttle box, a very popular device at the University of Pennsylvania. Such a box is made up of two compartments with a barrier in the middle. The barrier is about shoulder-high (on the dog) and may, with a bit of difficulty, be jumped over. The training began with the dog on one side of the box and with the route to the other compartment blocked by a gate, which the experimenters could raise. A buzzer sounded for ten seconds and the gate was raised. To the dog's surprise and dismay, this was followed by an extremely intense electric shock delivered through the floor to the footpads of the animal. As the authors described it, the shock level was "just subtetanizing," which means that it was just short of paralyzing the creature. Within a few trials, the dog learned to head for the barrier and leap it when the buzzer sounded and the gate was raised. If the dog did not "shuttle" to the other "safe" side, it received three seconds of the shock.

When the dog reached the other compartment, the gate was lowered and all was well—for a while. But soon (can it be?), the buzzer sounded, the gate was raised, and shock was forthcoming, so the dog jumped again. Then the buzzer sounded again, the gate was raised, and so on. Very quickly all dogs were shuttling back and forth, never getting shocked, and probably the exercise was doing them good. Here was a conditioned fear. The buzzer had become a CS producing fear, analogous to the phobic stimulus plaguing a human sufferer. How may such fear be eliminated, the dog's jumping be stopped, and the human's phobia be cured?

The simplest method is extinction, in which the CS (or phobic stimulus) is presented over and over until the reaction to it ceases. This was not at all successful in treating the dogs' fear of the buzzer, though. After the shock generator was turned off, two dogs continued jumping for 190 and 490 trials, respectively. The dogs showed no sign of stopping and their latencies actually showed a decrease with trials! They may have stopped after a few hundred (or a few thousand) further trials, but the authors wearied of going to work day after day watching the dogs endlessly jump. Clearly, extinction was not an efficient method for eliminating conditioned fears.

How about punishment? Although this would not increase the popularity of a therapist, it is possible that punishing phobic reactions with a strong shock or a slap in the face could work. Thus, conditions were changed in a novel way. As before, the buzzer sounded, the gate was raised, and the dog jumped—but not to safety. The gate was closed after the jump, and the dog found itself part of an electrical circuit; it had jumped into three seconds of intense and inescapable shock! Surely, this would give it pause. But it didn't; for 100 trials the animal jumped as quickly as it had before, and it knew what it was jumping into. As the experimenters reported, the dog's trip through the air was often accompanied by "an anticipatory 'yip' which changed to a 'yelp'" when the grid on the other side was reached.

Maybe the punishment wasn't strong enough, thought the experimenters, so the duration of shock was increased to ten seconds! Imagine, if you can, some time in your life when you received a strong and painful shock. Now imagine that shock (or a stronger one) lasting for ten seconds. In fact, try timing ten seconds to get a feel for the dog's plight. Yet, as you might have guessed, the dog continued jumping as rapidly as before!

This is indeed odd. One of the dogs initially had received only eleven pairings of buzzer and shock, and some of those shocks were experienced only briefly since the dog was jumping before the three-second shock period was over. Fear of the buzzer had lasted through 490 no-shock extinction trials, 100 trials punished by three seconds of inescapable shock, and 50 trials punished with ten seconds of inescapable shock. Was the dog's fear of the buzzer greater than its fear of the shock itself? (Is it true that we have nothing to fear but fear itself?)

Finally, the experimenters resorted to what they called "reality testing" to stop the jumping. As you probably have surmised, the dogs continued jumping because, although Solomon knew that the shock following the buzzer was discontinued, the dogs did not. How can we show the dog that it is safe to stay? (He knows that it is dangerous to jump.) The next day, the dog jumped for the 53rd, 54th, and 56th time into ten seconds of inescapable shock. But on the next four trials that day it didn't jump. This was because the experimenters installed a heavy glass plate over the barrier. The dog jumped, smashed into the glass, and landed in the same compartment from whence it had jumped. On subsequent trials it did not jump. What it did was stay on the same side barking, panting, and drooling. It was cured! After only 647 extinction trials, including over 150 with inescapable punishment, the fear produced by the pairing of eleven shocks with the buzzer was conquered. On the next day the buzzer sounded and the dog jumped, but only three times. It was indeed "cured." But why was the cure so difficult? Solomon must have wondered why, off and on for 30 years, until the answer dawned. The answer came during research in which dogs' heart rates were monitored during the administration of electric shock (Solomon & Corbit, 1974).

This time the dog being shocked was not in a shuttle box. It was simply resting in what is called a Pavlov sling, which is a canvas suspended from overheard with openings to accommodate the dog's four legs; in effect, the dog is hanging from the ceiling. As the ten-second shocks were applied to the dog's foot, its behavior was noted and its heart rate was monitored. The dog appeared terrified. It screeched, it thrashed, its pupils dilated, its eyes bulged, its hair stood on end, its ears were back, its tail was curled, it defecated and urinated, and its heart rate rose from 100 to 250 beats per minute. Fortunately, there is a happier ending coming. (In any event, this is not even a real dog. What is described is a composite of canine experiences in the laboratories of Solomon, his students, and collaborators.)

After a few shocks the dog was released from the sling and showed clear signs that it would now be a poor pet. It acted suspicious, unfriendly (wouldn't fetch), and stealthy—hardly surprising, given its recent treatment. The daily sessions of shock continued and, after a while, the dog's behavior changed. During the shock, it no longer appeared terrified but only anxious and annoyed. This was reflected in its heart rate, which scarcely increased at all and sometimes actually decreased during the shock periods. This reaction was not really surprising; we (and dogs) show *habituation* to prolonged stimulation, even when it is quite painful.

What was surprising was the behavior of the dog after these later shock sessions. Earlier, the dog seemed justifiably suspicious, but now it seemed perky, joyful, happy to the point of euphoria. In addition, its heart rate decreased as much as 50 to 60 beats per minute below baseline levels. This was a happy dog.

Why should painful electric shock have such an effect? Could it be that the dog was simply glad that it was over? Of course not. If that were the case, one would expect a euphoric dog after the earlier sessions, where the shock produced real terror, with its bodily accompaniments. Why should the dog be euphoric after shock sessions in which there is no sign of terror, no increase in heart rate, no bulging eyes? What is interesting here is that the sequence of states occurred as it did; terror was followed by stealthiness, and slight annoyance was followed by gaiety.

Noting this, as well as a number of similar cases, Solomon and Corbit proposed that a specific opponent process physiological/affective reaction was responsible. That is, any strong affect (pleasure, pain, terror, grief, and so on) may be accompanied by an affective reaction in the opposite direction. To keep things clear, we will stay with the shocked dog for a moment. The first few shocks produced strong negative affect (terror and sympathetic arousal), which was followed by a return to relative normalcy (stealth). The initial reaction (terror) is termed the *A state* and the opponent reaction the *B state*. Both occur at the same time and sum algebraically; hence, the shock produces sympathetic arousal and terror (A state) and a decrease in arousal (B state) simultaneously. The A state is much the stronger of the two, and evidence for the B state appears at first only after the offset of the shock and the end of the A state.

With repetition, the B state increases in strength and continues to outlast the A state. Since the reactions are literal opposites, the increase in magnitude of the B state subtracts progressively more from the A state until the A state virtually disappears. Thus, after a number of sessions the dog ceases to show terror during the shocks and its heart rate no longer increases (and may even decrease below baseline). After the shock, the now mighty B state shows itself in euphoric behavior and decreases in heart rate to a level far below baseline. Since the A and B states no longer resemble their original forms, the terms A′ and B′ are applied to these, the opponent reactions as they appear after long training.

The *opponent-process model* thus posits opponent reactions, A and B, where B is produced by (and simultaneously with) the A reaction. With repetition, the B reaction increases in strength and subtracts from the A reaction. After the stimulating event ends, the B reaction persists. In the case of the shocked dog, the persisting B (actually B′) reaction may last an hour or more; in other cases, the B′ state may endure for twenty years!

Human risk taking may provide an analog to the shocked dog. Humans who skydive for recreation must make that first jump, and for most people it is a terrifying experience. Behavior during the first few jumps closely approximates the signs of terror shown by the shocked dog, and the aftereffect is a feeling of quiet dullness. When we are momentarily filled with terror upon seeing someone backing down an exit ramp toward us we feel a sort of numbness thereafter, before the anger comes. Why would anyone continue skydiving if the effects were always terror and later numbness? The fact is that the skydiver, like the dog, later shows no sign of terror before and during the jump; what appears instead is a tenseness and eagerness. After the jump, the numbness is replaced by an exhilaration, which is accompanied by a lowered heart rate and which may last for hours. The parallel to the A and B states of the shocked dog giving way to the A′ and B′ states is obvious.

Where else may such a mechanism be operating? Solomon and Corbit suggest a number of other cases. For example, puppy love is characterized by an A state of excited happiness when the loved one is near and a B state of loneliness in his or her absence. After much experience, the loved one evokes an A′ state of normalcy, commonly reported by old married couples. But what is the B′ state, which is shown when the loved one is removed? It is the deep grief and depression that has come to be called the separation syndrome. The principle here is the same as that governing the shocked dog and the skydiver. Solomon applies the principle to a host of phenomena, including feeding in infants, imprinting in ducklings, and the behavior of humans struck by lightning (Solomon, 1980).

One significant aspect of this model is the evidence showing that the unconditioned response (UCR) is biphasic (Has A and B states). In addition, the fact that these reactions are unconditioned responses means that each can be evoked by conditioned stimuli (CSs). This has clear application to the problem of opiate addiction and was of central concern in Solomon and Corbit's argument for the model.

Users of opiates such as morphine and heroin typically report a pleasurable rush with each administration, followed by a later feeling of normalcy and perhaps a brief craving. This is true of new users. Users who have long experience have different reports, however. For them, the rush is minimal or completely gone and the drug is taken to eliminate a very strong craving which, according to many accounts, is agony itself. It thus appears that such addiction follows the opponent-process model and its sequence of stages. An initially pleasurable A state (the rush) is followed by a B state (normalcy or mild craving). With continued use of the opiate, the A state diminishes to an A′ state (normalcy) followed by a progressively stronger B state, the B′ state of craving (withdrawal). We simply have the mirror image of the shocked dog or skydiver and an exaggeration of the love/grief example.

If this is the case, how may drug addicts be treated? There is some evidence that the B state dissipates with time; the euphoria of the dog and the skydiver do not last forever, and even the separation syndrome eventually passes. But the B′ state appearing as withdrawal from heroin may last decades! Further, since the A and B states involved are really unconditioned responses to heroin, it is certain that a variety of conditioned stimuli are established that can make the addict's prognosis grim indeed.

Suppose that the cues present during the administration of heroin can become CSs for the A (and A′) reactions. The heroin itself, friends present, the hypodermic syringe, and even money (which becomes closely linked to the obtaining of a "fix") may become CSa's, which evoke a CRa, a fractional component of the A or A′ reactions. The evoking of the A reaction produces the B reaction, which thus strengthens the B state and increases the craving. Hence, it is best to stay away from such cues while trying to break the addiction.

But the B reaction is also a UCR and conditioned stimuli may as easily be associated with it! Thus, police stations, hospitals, non-drug-using friends, and other cues present during the B′ (craving) state may become CSb's and directly evoke the B state. When the B state is evoked, it gains in strength. Whichever way the addict turns there is no escape; the craving increases!

Solomon and Corbit provide convincing evidence that conditioned stimuli may become associated with A and B states in such a way and leave us wondering how any heroin addict is ever cured if such a process is going on. In fact, a lot of addicts don't get cured, and those who do overcome the addiction do so by biting the bullet, tolerating the agonies of withdrawal, and showing that the B process does eventually dissipate and that this dissipation slowly overcomes the augmenting that owes to the CSs that constantly add to the B state.

Other data, including some anomalies, may depend upon the same opponent-process mechanism. For example, Siegel (1978) has shown that the CR produced by a CS paired with morphine injection is opposite in direction to the effect of the UCS (morphine) itself. Morphine produces UCRs that include a decreased sensitivity to pain and increased body temperature. A CS paired with morphine produces a CR of increased sensitivity to pain and decreased body temperature. Similarly, glucose injections lead to hyperglycemia (elevated blood sugar), whereas a CS paired with glucose produces a CR of hypoglycemia. In addition, it has often been reported that the CR to a CS predicting shock is a decrease in heart rate, whereas the UCR for shock is, of course, increased heart rate. (Eikelboom and Stewart, 1982, dispute Siegel's interpretation. For example, they argue that glucose injection and hypoglycemia are not viewed properly as UCS and UCR. They are actually a single event, the UCS. The actual UCR is the release of insulin by the pancreas that causes a decrease in blood sugar level. Thus, the CR and UCR are not opposite in direction in this case.)

An opponent-process mechanism could well account for such instances of what seem to be homeostatic compensations. Very likely such opponent CRs represent the evoking of a B′ state, thus occurring after some minimal number of CS–UCS pairings. If this were the case, earlier presentations of the CS should produce CRs in the same direction as the UCRs. The data that could substantiate such a possibility are not available, but it is an interesting possibility.

In summary, the Solomon and Corbit model tells us some interesting things. The unconditioned response is not a unitary thing and the basic nature of its two states may change drastically with continued evocation. Terror/numbness may become annoyance/euphoria, and ecstasy/loneliness may turn to normalcy/grief. Also, the existence of opponent pairs of emotions may be more reasonable than has been supposed (cf. Wundt, 1896). But instead of happy/sad and fear/relaxation we may have happy/fear and love/sadness if Solomon and Corbit's opponents are what they seem to be.

Finally, and anticlimactically, we know why Solomon, Kamin, and Wynne's dogs continued jumping day after day and why such extraordinary means were required to stop them. Those first few shocks produced an A state of pain and fear, followed by a B state of relief as the dog jumped. Even after the shocks no longer were received, the buzzer and the raised gate acted as CSs for the A state, which produced a CRa and a CRb. Every jump by the dog was accompanied by CSs for the B (relief) state, and that state consequently was strengthened on every trial. Throughout their long training, the dogs were jumping into a huge B (actually B′) state, which was evidently more powerful than the effects of the punishments used to stop the jumping.

Learned Taste Aversions

John Garcia demonstrated a phenomenon that many viewed as very damaging to traditional learning theory and it was difficult for him to publish at first. What Garcia showed was the apparent uniqueness and specificity of a particular kind of learning, that of the association of taste and illness. He showed that taste, and not other cues, was specifically tied to subsequent illness and that the interval between the taste and the illness could be very long.

For example, in one of his best-known studies, Garcia and Koelling (1966) fed rats "bright-noisy-tasty" water, paired with electric shock or illness. The brightness came from flashing lights, the noise came from a clicking relay, and the taste came from saccharine or salt added to the water (see Figure 9.3). The electric shock came from a 0.5-second, 0.08–0.20-milliamp shock source and the illness was produced by a dose of metal poisoning (lithium chloride, or LiCl) mixed with drinking water, or by 54r (roentgens) of 250-kv (kilovolts) X rays delivered in twenty minutes. Both the LiCl and the X-irradiation has been shown to produce severe gastro-intestinal illness some time (many minutes or hours) after ingestion. What should happen in such a situation? Suppose I drink sweet water and become sick, or I drink salty water and become sick; suppose I drink sweet or salty water and receive shock. Would I avoid that water in the future?

Some of Garcia and Koelling's rats, which were poisoned, either with X-irradiation or lithium chloride, later avoided the sweet and the salty tastes, which had been paired with poisoning. But these rats showed no aversion to the flashing lights or the clicking relay, which also had been present when they were poisoned. On the other hand, the rats that were shocked while drinking sweet or salty water showed no later aversion to sweet or salty fluids; they avoided the flashing lights and the clicking relay.

What does this mean? What it surely suggests is that rats are more apt to associate tastes (sweet or salty) with subsequent illness than they are to associate lights and sounds with illness. By the same token, they associate lights and sounds with other unpleasant stimulation, such as the pain produced by shock. And this occurs for obvious adaptive reasons.

Assuming that we give some credence to the doctrine of evolution by natural selection (which has been ignored by some learning theories), we may expect that an omnivorous animal like the rat would be likely to pay attention to the tastes of the foods it eats and to the accompanying gastric effects. If a rat becomes violently ill, it consequently avoids any novel tasting foods it may have had during the recent past. This helps rats survive. And thus it is difficult to rid an area of rats by using poison.

But the rats did not avoid tastes that were accompanied by shock; they avoided the lights and clicks that were present with shock. And that also makes sense. They avoided those cues that accompany externally produced pain; those cues are usually externally sensed cues. When pain is internal, as in the case of nausea, an animal might "think" of what it ate and avoid eating it again. As evolution proceeded, rats that followed this rule survived to produce rats that behave similarly, and those rats that did not follow the rule did not survive to produce similar rats.

This was a striking finding and it has been repeated so many times that its truth is beyond question. Mammals associate the tastes of recent foods with subsequent illness. They do not easily associate the place where the food was gotten with illness, nor do they associate other sights or sounds with illness. The external sights and sounds are associated with pain produced by shock, and tastes are not so associated.

Equally startling was the finding that the CS-UCS interval could be measured in hours, rather than in minutes or seconds. Garcia, Ervin, and Koelling (1966) showed that a learned aversion to apomorphine poisoning (which causes regurgitation) occurred even when the saccharine taste that preceded the poisoning occurred more than 75 minutes before. In other cases, such as that reported by Smith and Roll (1967), an aversion was learned to saccharine-flavored water and irradiation after a CS-UCS interval of as long as twelve hours! In the Garcia, et al. study, rats experiencing electric shock showed no aversion to saccharine, as in the Garcia and Koelling study.

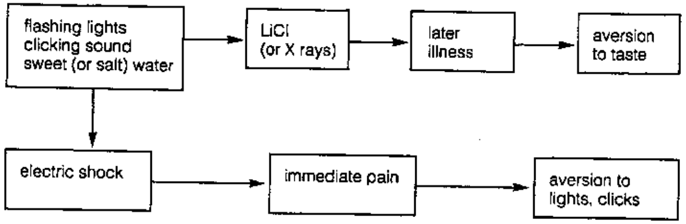

FIGURE 9.3 Learned taste aversions. *Garcia and Koelling, 1966.*

What does this mean? A rat, tasting saccharine or salt in the presence of flashing lights and clickers, is later (hours later) offered saccharine or salt or flashing lights or clickers. The rat will not drink fluids that taste like those it tasted before being poisoned an hour ago, or twelve hours ago. (The maximum duration is uncertain, but let us accept the twelve-hour figure for now.) But the poisoned rat will drink fluid that was presented hours ago with flashing lights and clicking sounds. What is this?

Garcia and others have reasoned that rats, being omnivorous "eaters of opportunity" are apt in their foraging to eat a number of toxic things. Those rats that have survived the process of natural selection and thus exist at present are those that have learned to react to nausea in the same way. When seriously ill, as is the case when Garcia exposed them to X-irradiation or lithium chloride poisoning, such rats looked back to what they last ate. They did not look back to where they last were, since they (and past rats) often ate different things in the same place. When sick, think of what you have eaten, not where you have been. On the other hand, when you have been injured from outside your body, by predator or otherwise, you are better off remembering the place where it occurred than what you had been eating.

Is Taste Aversion Learning Just Classical Conditioning? First, is it a mystery that taste, rather than lights and clicks, is aversive? Recall from Chapter 1 that one of the oldest of associationist principles is the law of association by similarity. Now consider the stimuli present when you drink sweet, bright, noisy water. As Garcia and Koelling suggested, it is possible that tastes become connected with nausea because of the similarity in the stimuli involved. That is, when a light flashes or a clicker clicks, it is a discrete, phasic stimulus; it has a clear beginning and an end. But when you become nauseated, it usually occurs more gradually. Unlike a light flash or a click, a taste also has a gradual onset, much like the stimulus pattern in nausea. Thus, we and the rat associate taste and nausea because of their similar patterns of gradual onset and offset, but we associate physical pain, such as that produced by shock, with abrupt signals from our surroundings, since such signals more often predict pain of that sort than they do illness. Testa (1975) illustrated that this might be the case by showing that rats are particularly apt to show fear when a light pulses at the frequency at which shock has been delivered in the past. Thus, it could be that the specificity of the taste-illness connection is no more than an instance of the old GPLT principle of association by contiguity.

But what of the long CS–UCS delay, perhaps extending over twelve hours? Imagine yourself made sick. You drink some sweet-tasting solution and thereafter are subjected to a dose of radiation or of lithium chloride poisoning. Hours later you are sick—very sick! When we are very sick in this way, we exhibit emesis—which is to say that the reflex mechanisms in the gastrointestinal tract begin reverse peristalysis to bring up the offending substance. We regurgitate. While we are sick, we are apt to partially regurgitate a number of times, and the food that we recently ate is thus apt to be retasted. Thus, there is really no hours-long delay between taste and illness; when we retaste the food while sick, we pair the taste and illness. This could be the explanation for the long delay between CS (taste) and UCS (feelings of illness) in taste aversion learning.

Tempting and reasonable as that may sound, it is not the answer. The reason we know this is simple: rats do not regurgitate, as humans do. When rats become ill, the contents of their stomachs are prevented from returning up the esophagus by the cardiac sphincter. And even if some aftertaste from the stomach does find its way up, the learned aversion still is demonstrably not due to the association of present illness with a present aftertaste.

Revusky and Bedarf (1967) showed that novel tastes are more apt to produce learned aversions than are familiar tastes, even while the latter were more recently experienced (and thus more likely to produce aftertastes during subsequent illness). In two experiments, they familiarized different groups of rats with either milk or sucrose or with sucrose and grape juice. Then each group was given a taste of both the familiar and the unfamiliar flavor (for example, first sucrose, then grape juice) in both orders. A rat familiar with sucrose might get a taste of sucrose followed by a taste of grape juice, while another rat also familiar with sucrose might taste the grape juice and then the sucrose. The rats were then exposed to X-irradiation, which made them ill later (see Figure 9.4).

During subsequent choice tests, the rats avoided the unfamiliar flavor that had preceded poisoning, for obvious reasons. It was as if they retrospected and behaved as we would, blaming their illness on something recently tasted and also unfamiliar. Taste aversions to novel tastes are learned far more readily than to familiar ones. Since these results applied to novel tastes presented both after and before the familiar ones, it seems obvious that an aftertaste produced by the illness could not be the cause of the aversion. (Rozin, 1969, showed that aversions may quickly be learned to specific concentrations of taste solutions. This occurs even when poisoning occurs a half hour later, when differences in aftertastes should be negligible.)

Specific Hungers Rozin (1967) suggested that learned taste aversions provide a simple explanation for an intriguing set of phenomena investigated since the 1920s and 1930s by Curt Richter and others. This is the effect appearing as specific hungers, and it is one of the most impressive instances of an "isn't nature wonderful?" phenomenon (cf. Chapter 6).

Suppose that we deprive a rat of thiamine (vitamin B_1) by feeding it commercially available rat food that lacks thiamine. Then, when we give the rat a choice of foods, it selects thiamine-rich food, which will remedy its deficiency. By 1930, researchers demonstrated that a variety of animals deprived of sodium, vitamins, proteins, sugars, fats, and minerals would select those foods that supplied whatever was missing. This is really an amazing finding.

The reason that it is amazing lies in the mechanisms that must be necessary to allow self-selection of needed nutrients. Since there are at least a dozen specific hungers that have been demonstrated, it follows that there must be at least a dozen specific sensors in the body that monitor the levels of those nutrients in the blood. When the level of one of them is down, the sensor (presumably in the hypothalamus) must detect it and somehow induce the organism to eat foods that contain that nutrient. The "somehow" is the real trick! If I am low on fats or vitamins, for example, does my body tell me to eat peanut butter? If I am low on calcium, how does my body tell me to drink milk?

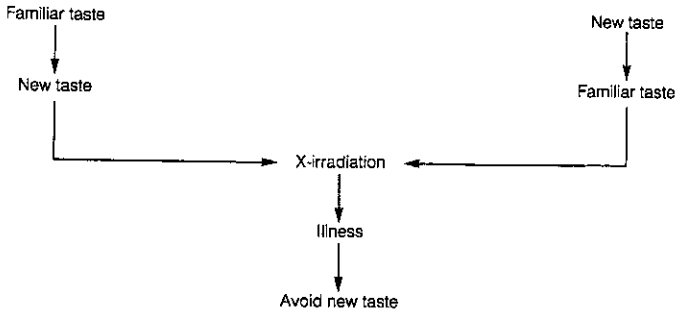

FIGURE 9.4 Familiar versus novel tastes in learned taste aversions. *Revusky and Bedarf, 1967.*

What must happen (somehow) is a change in the attractiveness of some tastes. If I am low on fats or vitamins, I must say to myself, "I really have a taste for peanut butter." A calcium deficiency should make me think, "I would love a glass of milk—not Coke or orange juice, but milk!" The whole thing is simple if we are talking about us and our long history of experience with a large variety of foods. If I feel sluggish, a doughnut seems appealing, because in the past when I ate a doughnut I was temporarily pepped up. Maybe the other tastes I have from time to time—for pizza, Coke, or coffee—come from the fact that in the past my feelings of disquiet have been quieted when I ate specific foods. I have learned what to eat to feel better.

But what of young rats that have eaten only laboratory food pellets during their entire lives? If they know to eat that which contains the nutrients they need, then the whole matter is basic biology: Nature *is* wonderful! The sensors in the body detect a deficiency in some vital nutrient and then they somehow make foods that contain such nutrients taste better! Can we really assume that there are twelve or thirteen special sensors to influence preferences for a dozen different tastes? Probably not.

Rozin (1967) suggested an alternative that can explain specific hungers far more simply (see Figure 9.5). Instead of an organism sensing a deficiency in one of a dozen nutrients and seeking out foods that contain the needed nutrient, as suggested by the "isn't nature wonderful?" account, Rozin suggested that all the creature needs to know is that it is now sick and to remember what it has eaten recently. Specific hungers could be merely an instance of taste aversion learning.

That is, when an organism has been eating a diet deficient in thiamine, it associates the taste of that diet with feelings of illness, just as does the rat poisoned with lithium chloride. And, just as is the case with the poisoned rat, the thiamine-deficient rat avoids food recently eaten (and thus develops paleophobia); if given a choice, the rat eats unfamiliar foods or old foods that were not associated with illness. So the occurrence of specific hungers is not really evidence that creatures seek out foods containing specific nutrients. In fact, sick animals avoid familiar foods because they may have caused the sickness and settle for novel foods, whether those foods contain the missing nutrient or not (that is, show neophilia). Of course, if they eat the new food and remain sick, the new food becomes old and is apt to be avoided as is the original old food (cf. Rozin & Kalat, 1972).

Rozin suggested that this analysis may account for many cases of diet selection but not all cases.

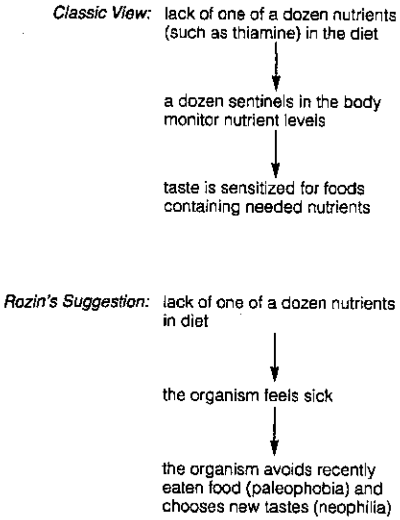

FIGURE 9.5 Classic view versus Rozin's view of specific hungers.

For example, there apparently is a real specific hunger for sodium, and there are cases in which rats and humans select a balanced diet from a set of foods presented them. There are many animals running around in the wilds right now that manage to eat a more or less balanced diet. Such wise food selection cannot be explained entirely in terms of learned taste aversions, can it?

The Significance of Taste Aversion Learning Since the first reports of learned taste aversions by Garcia and his colleagues, there have been countless studies replicating and extending the effect. Why has such attention been paid to such a simple effect? How could such a finding play a large part in changing the direction of experimental psychology?

The answer is simple, though hard to believe. It lies in the fact that taste aversion learning seems obviously to fit the simple paradigm of classical conditioning but has characteristics not shared by other forms of classical conditioning. In classical conditioning, common belief had it that all stimuli could be as effective as CSs, regardless of the UCS used. We may use a tone, bubbling water, a light, a word, or anything as a CS, and it should work about as well as any other CS, as long as the intensities are not wildly different. And classical conditioning occurs only if the interval between CS and UCS is relatively brief; it may be a matter of seconds or of minutes, but not much longer.

Then we discover taste aversion learning, half a century after Pavlov introduced classical conditioning to us. We find here that just any stimulus is not effective; rats are affected by taste and birds by color. Evidently, whatever cues are used in normal feeding are effective in the learning of taste aversions. We also find that the interval between the CS (taste) and the UCS (illness) may be very long, perhaps several hours.

All of this tells us that nature has made taste aversion learning special. Those organisms that learned to avoid tastes that made them sick lived to bear offspring with the same talent. Perhaps taste aversion learning is not an odd case of classical conditioning; maybe it is a species-specific form of learning. And that means there must be other species-specific forms of learning. Don't most humans learn language remarkably quickly? Don't birds fly and don't fish swim? How about the migration of the salmon up the streams of their youth at spawning time? How about the migration of the arctic tern? Surely we knew that there were many species-specific behavior tendencies.

The discovery of taste aversion learning and the realization that it involved hereditary dispositions came as a shock to many who had placed too much faith in the simplified versions of conditioning and learning theories that appeared in the textbooks of the 1950s and 1960s. These books emphasized primarily the views of Hull and his followers; when other theorists, such as Tolman or Skinner, were covered, their views were presented as variations of Hull's theory.

Thus, by the 1960s psychologists had favored a simple S-R associationism, and that view did not have a place for instinct.

Some denied that learned taste aversions were special or that they really violated the laws of learning. For example, Revusky (1971) suggested that such learning was an instance of classical conditioning, differing only in the fact that it is taste (in the case of mammals) that is associated with illness. This "belongingness" of taste and illness is what makes the phenomenon unique, not the fact that the CS–UCS interval may span hours. Revusky suggested that in ordinary classical conditioning, such delays prevent learning because of interference produced by stimuli during the delay.

For example, if we present a brief tone CS and follow it reliably by an electric shock an hour later, we are not likely to see a CR produced by the tone. This is because of the interfering effect of all of the sights and sounds that are sure to occur during a one-hour CS–UCS interval. But in the case of taste aversions, the only stimuli capable of interfering are other tastes, and it is likely that none (or few) will occur during the hour interval between taste and illness.

Revusky considers taste aversion learning semispecial, therefore. The possible long CS–UCS interval is not special; it is simply made possible by the "belongingness" principle, which does make the learning special. Revusky and others have pointed out that taste aversions resemble normal classical conditioning in most respects. That is, learning is better when the CS and UCS are moderately intense, when the CS–UCS interval is brief, and when other stimuli do not block, overshadow, or appear during the CS–UCS interval.

It could be that Revusky did not go far enough. Maybe taste aversion learning is really just classical conditioning, and maybe "belongingness" can be explained in terms other than those of hereditary dispositions. In fact, Mazur (1986) suggested that taste aversion learning has become accepted as a case of classical conditioning and pointed out that the taste-poison paradigm has become one of the favorites for studying classical conditioning!

SUMMARY

Instrumental Conditioning

Although the law of effect has little to do with hedonism, reinforcers and punishers often have been treated as equivalent to pleasures and pains. This is unfortunate, since hedonism is a very vulnerable point of view. Much of the recent research in instrumental conditioning shows, among other things, that it is long past time to bury hedonism once and for all.

It was the implicit belief in hedonism that led to a misinterpretation of Olds's discovery; what were formerly called "pleasure centers in the brain" are now better understood. Something that acts as a powerful reinforcer need by no means be "pleasant." Similarly, so-called punishment centers were mislabeled. Like the effects of stimulation elsewhere in the limbic system (for example, evoking feeding, fasting, or aggression), the effects are far more general than had been believed. Olds's discovery did not advance our understanding of the law of effect.

Premack showed that reinforcers and punishers are not fixed classes; they are better treated as activities, and, under the right circumstances, any activity can act as a reinforcer or punisher of any other activity. The flaw in Premack's theory was the necessity to scale activities on a continuum of value. This was not practical and turned out to be unnecessary in any event. The theory of response deprivation requires only that one activity be available less than is preferred by a subject; access to it may then be used as a reinforcer for another activity. Allison suggested that a sort of conservation applies here, such that the amount of reinforced and reinforcing behavior remains constant as access to each is altered.

All of the research stemming from Premack's theory treats the law of effect as a molar entity, defined as changes in relative amounts of behaviors over time. This is true also of the matching law, first proposed by Herrnstein. Masses of data show that animal and human subjects distribute their responses to a number of alternatives so as to match the relative payoff provided by each alternative. Like the Premack Principle, the matching law shows that reinforcement is a relative thing; the effect of a given frequency or magnitude of payoff depends upon the payoffs available for other available alternatives. The matching law may even explain cases of self-control, if we consider time as well as magnitude of payoff.

Autoshaping was originally believed to be damaging to extant theories of learning, since it showed that much of what was assumed to be operant behavior was actually classically conditioned. This is really only of great interest if one believes that the operant/respondent distinction is very important and that responses have to be emitted *or* elicited and kept separate! Although autoshaped behaviors may be elicited, we find that this has little effect on our understanding of behavior in general.

Researchers over the years assumed that behaviors could be treated as independent and that manipulation of one could be carried out without affecting others. Behavioral contrast and local contrast have now been documented often enough to show that they are reliable and that (once again) effects are always relative. Responding to one stimulus depends in part on the rate of reinforcement in other, successively presented components.

Understanding conditioned reinforcement has been a problem for many years. Although it seems that tokens and other acquired reinforcers gain their power from pairing with already effective reinforcers, research has not clearly supported the pairing theory. An alternative interpretation is that things become acquired reinforcers when they are produced in the same way as are already effective reinforcers. This explains why simply pairing *may* be effective and it explains why many consequences of our daily behavior come to act as reinforcers.

Avoidance learning is obviously important to many species. Yet, it has proven difficult to explain how such learning takes place. The main difficulty has been in understanding how an avoidance response is maintained when it is never reinforced, except by the nonoccurrence of the harmful event. Two-factor theory assumed that some signal always accompanied harmful events and that responses that removed such warning stimuli were reinforced by the reduction of fear. More recent research has shown that organisms respond to decrease the overall level of the stimulus (such as shock), even when each response is followed by the stimulus. Such findings invalidate two-factor theory and show that a molar viewpoint removes the mystery that comes with molecular viewpoints, as was the case with the matching law and conditioned reinforcement.

After 1930, Thorndike argued that annoyers work indirectly, by encouraging new behavior. This position was similar to that of Skinner, who believed punishment to be a by-product of emotion produced by the aversive event. This view was supported by evidence, such as Estes's, suggesting that punishment had temporary effects no stronger than those produced by randomly administered aversive events. However, Azrin and his colleagues demonstrated in the 1960s that punishment does work as directly as does reinforcement, though with opposite effect.

Classical Conditioning

Rescorla showed in the 1960s that classical conditioning occurs only when the CS predicts the UCS; it is not enough that it is paired frequently with it if it also frequently appears independently of it. Shortly thereafter, Rescorla and Wagner proposed a variant of a very old model to explain how Rescorla's evidence for the importance of contingency could be explained as contiguity (pairing) of CS and UCS. The model also explains overshadowing and blocking and has been applied to stimulus generalization phenomena.

Solomon and Corbit proposed the opponent-process model to account for a number of motivational and emotional phenomena, as well as for drug addiction. Their evidence shows that many UCRs are biphasic; a strong response during earlier conditioning trials diminishes over sessions, due to the development of an opponent response. Thus, the effect of a UCS depends upon frequency and spacing of application. Their model may account for seemingly paradoxical effects, such as those reported by Siegel.

John Garcia showed that many organisms are predisposed to associate taste with subsequent gastrointestinal illness. Aversions following illness are specific to tastes accompanying poisoning, even when taste and illness are separated by many hours. Taste aversion learning initially was believed to be damaging to current learning theories, but subsequent research has not supported this view. In fact, the taste aversion paradigm is now a popular procedure for those studying classical conditioning. Learned taste aversions may well explain many instances of what have been called specific hungers.

GLOSSARY

A state Solomon and Corbit's term for the initial reaction to an affect-producing UCS, such as electric shock. With repetition, the A state diminishes in strength. This occurs because the opponent B state increases in strength over trials. (See *opponent-process model.*)

B state Solomon and Corbit's term for the opponent reaction produced when the A state is produced. If the A reaction appears as an increase in heart rate when shock is applied, the B reaction is a decrease in heart rate. The B reaction is assumed to increase in strength with repetition and to thus eventually reduce the A state to the diminished A′ state. (See *opponent-process model.*)

Behavioral contrast An increase or decrease in response rate to a stimulus when a change in conditions occurs during a second stimulus. When reinforcement rate is decreased for responses to a red light, response rate may increase greatly during green, even though there is no increase in reinforcement frequency in green.

Blocking Term used by Kamin to refer to the effect of pretraining with one CS on subsequent failure to establish a CR to a second stimulus presented along with the original CS. In terms of the Rescorla-Wagner model, the first CS gains the majority of the associative strength available, preventing additional conditioning to the added CS.

Commensurability of units The major problem with Premack's Principle, as originally proposed, in which widely different activities are scaled on a common continuum of value. It becomes difficult to assume that one unit of an activity is commensurate, or translatable, into one unit of another activity. For example, is one minute spent eating commensurate with one minute reading?

Conditioned reinforcer An acquired reinforcer; one that gains its reinforcing power during an individual's lifetime.

Conservation Principle applied to allocation of activities under schedule constraints. Allison and his colleagues have shown that total behavior per session remains constant under many conditions. Thus, more of one activity means less of some others.

Contiguity Closeness in space and/or time. Pavlovian conditioning was formerly thought to depend only on the simple contiguity of CS and UCS in time.

Contingency Contingent events have an if/then relationship. As an example, a CS which unequivocally predicts a UCS has a clear contingent relationship with the UCS. A CS which is always followed by a UCS does not have such a contingent relationship if the UCS also appears at other times.

Drive The chief way of treating motivation in many psychological theories, especially Hull's. Biological imbalances, or needs, lead to behavior aimed at reducing the need. A need acting to produce behavior is termed a drive.

Habituation A decrease in response with repeated stimulation. For example, the startle response to a gunshot diminishes with repetition.

Hedonism The doctrine that our conduct is regulated largely by pleasure and pain.

Local contrast A sequential effect that occurs during discrimination learning. (See Pavlovian induction, Chapter 3.)

Matching law Also called the molar law of effect. Herrnstein demonstrated that relative response rate is equal to (or matches) relative reinforcement frequency.

Opponent-process model Solomon and Corbit's theory of acquired motivation. According to this model, strong stimulation produces either positive or negative reactions (the A state) as well as a compensating reaction in the opposite direction. With repeated stimulation, the opposing reaction (the B state) increases in strength and subtracts from the A reaction. The model provides a plausible account of the mechanisms involved in risk taking, drug addiction, and other phenomena.

Overshadowing Effect first reported by Pavlov, who found that when more than one CS is presented at the same time, the more salient one may "overshadow" the less salient one, leaving the latter ineffective.

Passive avoidance Avoidance of noxious stimulation achieved by not responding. We passively avoid burns by not placing our hands in fires. The term was often used in place of the term punishment during the several decades when it was believed that punishment was ineffective.

Phobia An unnatural and usually unreasonable fear of common stimuli. Agoraphobia, for example, refers to a fear of open spaces; siderophobia is a fear of railroad trains.

Premack Principle Revision of the law of effect first proposed by David Premack in 1959. According to this view, activities in which we engage may be ranked along a continuum of value. Reinforcement occurs when lower-valued activities provide access to higher-valued activities. When a higher-valued activity produces access to lower-valued activities, the former decrease in frequency (are punished).

Primary reinforcer A reinforcer that acts as such in the absence of prior experience with it; its power does not depend upon learning. Food and water are usually viewed as primary reinforcers.

Punishment A decrease in the probability of a behavior owing to the consequences it produces. Electric shock produced by a response is a commonly used punisher.

Rescorla-Wagner model Theory of classical conditioning published by Rescorla and Wagner in 1972. According to this theory, conditioning always involves a compound CS and a limited amount of available associative strength, determined by the nature of the UCS. Associative strength gained by one element of the compound CS detracts from that available to other elements. Increments in associative strength on successive trials depend upon the maximum (asymptotic) level possible and the amount of associative strength already gained. The model accounts for overshadowing and blocking, for which it was formulated, as well as for other phenomena.

Response deprivation Name given by Timberlake and Allison to their revision of the Premack Principle. According to this theory, reinforcement occurs when a behavior is restricted so that it can occur only at less-than-baseline level and when a second behavior is required for access to the restricted behavior. The second behavior will increase in frequency (be reinforced) as long as the restricted behavior remains available at less than its baseline (unrestricted) level.

Response language Name given by Premack to his way of viewing the law of effect. Instead of speaking of behaviors producing reinforcing stimuli, such as food, Premack spoke of behaviors producing access to other behaviors, such as eating.

Transsituational Term used by Meehl in 1950 to refer to the power of a reinforcer of one behavior (such as bar pressing) to act the same in different situations and with different behaviors (such as key pecking).

Value Intervening variable used as a basic term in Premack's theory. Value refers to the attractiveness of activities and may be assessed in terms of time spent in one or another activity.

RECOMMENDED READINGS

Brain Stimulation (or So-Called Pleasure Centers in the Brain)

Valenstein, E. S. (1973). *Brain control.* New York: Wiley.

Valenstein is a well-known researcher in this field and presents an entertaining and authoritative summary of research on brain stimulation and psychosurgery.

Pribram, K. (1971). *Languages of the brain: Experimental paradoxes and principles in neuropsychology.* Englewood Cliffs, NJ: Prentice-Hall.

This book provides the advanced student with a scholarly review of findings in neuropsychology as well as a speculative model for memory.

Instrumental Conditioning

Honig, W. K., & Staddon, J.E.R. (Eds.), (1977). *Handbook of operant behavior.* Englewood Cliffs, NJ: Prentice-Hall.

This handbook contains chapters written by authorities in fields ranging from the control of feeding to language. More recent research may be found in the *Journal of the Experimental Analysis of Behavior.*

Mackintosh, N. J. (1974). *The psychology of animal learning.* London: Academic Press.

This is a difficult, scholarly compendium of research and theory in the associationist tradition. A shorter but equally difficult treatment of more recent work appears in Mackintosh (1983). See also the *Journal of Experimental Psychology: Animal Behavior Processes.*

Classical Conditioning

See recommended readings following Chapter 3.