Chapter **8**

# Reinforcement Theory

*I have but one lamp by which my feet are guided, and that is the lamp of experience. I know no way of judging of the future but by the past.*

—PATRICK HENRY

Most voluntary behavior appears to be **purposive**. It is forward-looking and directed toward a goal. Attainment of that goal—finding food or a mate, escaping or avoiding fear or stress—is the **consequence** that directs behavior this way or that.

In this chapter and the next, we will look at contrasting theories about what makes behavior purposive. This chapter will focus on **reinforcement theory**, a cornerstone of modern behaviorism. Briefly, the theory says: *The behavior we call purposive is behavior that is maintained by its reinforcing consequences.*

## REINFORCEMENT AND PURPOSE

From very early on in this book, we have distinguished between a stimulus-bound reflex or action pattern on the one hand, and a purposive, goal-directed motivational state on the other. A motivational state, we have seen, is one that can be expressed by an arbitrary response. If we see that the subject *does whatever is available* to obtain a goal, then we say that a motivational state exists. Now in practice, when we show that an animal will make an arbitrary response to obtain a goal, we almost always use a reinforcement procedure.

### The Use of Reinforcement to Diagnose Purpose

Think again about a rat pressing a lever to get food. We may train the rat to press a lever, but we could just as well train it to turn left at a choice point or dance around the cage. We select an arbitrary response—pressing or turning left or dancing—and then *reinforce* it with a bit of food. If the experiment is successful, then we have shown a motivational state.

In such a case the reinforcer—food—is an *environmental* object or stimulus. But it is easy to see how internal factors, as discussed in Chapters 3 and 4, can affect behavior by way of these environmental events. To a food-deprived rat or human, food is a reinforcer; to a sexually motivated one, access to a mate is a reinforcer; and so on. Internal factors can affect behavior by *setting the reinforcing value of environmental commodities*.

This is similar to the idea we developed on pages 93–95: that drive states can determine the *pleasantness* of an object such as food. But the notion of reinforcement is more general than that, as we will see in a minute.

### Reinforcement Theory: Purpose *Means* Reinforcement

Thus we use reinforcement *procedures* as a way of deciding whether or not an action reflects a motivational state. But we also speak of reinforcement *theory.* This is the modern behaviorist's theory about the nature of purposive behavior, and it offers a solution to the great paradox with which such behavior confronts us.

The paradox is this. We speak of behavior as influenced by its consequences. But when the behavior occurs, those consequences haven't happened yet. They lie in the future. And if there is one thing we insist on, it is that causes precede their effects; they don't follow them. Hence the problem: How can future events cause present actions? How do we get the future into the present?

Reinforcement theory says: We don't. The action is affected not by its *future* consequences, but by what the consequences of such action have been *in the past.* If in the past an action has been reinforced, then the organism is changed in such a way that that action is more likely to occur thereafter.

Reinforcement theory asserts that we can understand all purposive behavior this way. The causes of action are not in the future; they are in the past where they belong. And the *consequences* that control action are not future ones, but past ones.

That is the theory. Before we examine its accomplishments, let us review where it came from.

## DEVELOPMENT OF THE REINFORCEMENT PRINCIPLE

In Chapter 2, we saw some forerunners of the reinforcement principle. There was Hobbes's *hedonistic* theory of action: We perform acts that are

*associated with pleasure,* and refrain from ones associated with displeasure. Why are they associated with one or the other? Because they have led to one or the other in the past.

The modern statement of the idea began with Thorndike's *law of effect.* Thorndike saw the attainment of satisfying consequences as connecting the response more firmly to the stimulus situation. For him, the *association* was between the stimulus situation and the response. Reward strengthened that association; punishment weakened it.

## Skinner's System

During the 1930s, B. F. Skinner presented a different way of looking at the law of effect.* His formulation is the one that is most influential at the present time.

### Operants and Respondents

Skinner begins by distinguishing two kinds of actions. Responses that are *elicited by a stimulus* are called **respondents.** This class of responses includes built-in reflexes; for example, a pinprick *elicits* leg withdrawal. It includes classically-conditioned responses; for example, after an appropriate conditioning history, Pavlov's bell *elicited* salivation in dogs. It also includes action patterns evoked by releasing stimuli. In short, respondents are any acts that are *stimulus-bound* (pp. 104–105).

**Operants** are responses that are *not* elicited by stimuli. These are the actions that we speak of as voluntary rather than reflexive. Bar-pressing in the rat does not bear any fixed relation to an eliciting stimulus. Thorndike saw such a response as becoming attached to the stimuli provided by the situation; but Skinner sees no need to talk that way. The responses occur, but no specific environmental events *make* them occur; they surely have causes, but those causes are not environmental stimuli. Hence, Skinner speaks of such responses as *emitted* by the organism, rather than as elicited by stimuli. And it is these responses that are related to their consequences.

### Reinforcement and Extinction

The heart of Skinner's system is his restatement of the law of effect. He presents his **law of conditioning** for operants: "If the occurrence of an operant is followed by presentation of a reinforcing stimulus, the strength [of the operant] is increased." And there is the converse **law of extinction:** "If the occurrence of an operant already strengthened

*There is a great deal of confusion about this, because careless writers have lumped Skinner with the Thorndike–Watson "stimulus-response" or S-R tradition within behaviorism. Skinner himself has made it very clear, however, that he is not an S-R psychologist (he does speak of stimulus control, but he means something quite different by it); and his views on behavior are fundamentally different from Thorndike's.

through conditioning is not followed by the reinforcing stimulus, the strength is decreased."[1]

And what is a reinforcing stimulus? It is one that does strengthen the response that precedes it. If Response R is followed by Stimulus S, and if Response R increases in strength, then Stimulus S is a reinforcer by definition. Strictly speaking, therefore, the law of conditioning is not a law at all in the usual sense. Rather it is the definition of a reinforcing stimulus.

Notice that a reinforcer, so defined, is not assigned any specific properties. In particular, nothing in the definition says that the reinforcer must be pleasant or satisfying, or even that it must be perceived by the subject. It is a reinforcer if, and only if, it reinforces. Skinner's system specifies nothing else about it.

### Discriminative Stimuli and Stimulus Control

Although operants are not *elicited* by external stimuli as reflexes are, external stimuli can come to control them in another way. Think again about a rat in a box, pressing a lever for food. Let there be a light in the box that can be turned on or off. When it is on, lever-pressing will be reinforced with food; when it is off, lever pressing will not be reinforced. The rat will eventually come to press the lever only when the light is on. When it is off, lever-pressing will weaken—undergo extinction—and cease to occur. When that has happened—when the rat presses the lever if the light is on, and not when it is off—we say that the light has acquired **stimulus control** over the emission of the response.

A case of stimulus control in humans is our behavior at traffic lights. If the light is green, we drive on; if red, we hit the brake. Other examples abound—whether to dress formally or casually, whether to use the formal or the familiar form of address in French. In all such cases, what we learn is not the behavior itself, but what behavior is appropriate in what situation. When we have learned it, then we say that the situation exerts *stimulus control* over the behavior.*

This role of the stimulus is quite different from the eliciting or releasing function of the stimulus in the case of respondents. There is nothing about the light that causes a bar-press from the rat; there is nothing about the red light that causes brake-pressing in humans. Look at it this way: In both cases, the conditions of reinforcement could just as well be reversed. The light could signal the *unavailability* of food, not its availability. Then the rat would learn to press when the light was off, not when it was on. We could have set up the rule that green means stop, and red means go; then drivers would learn that.

[1]Skinner, 1938, p. 21.

*The human cases are more complicated because, in most cases, the stimulus control is not developed by reinforcement and extinction, but by *instruction*—we are *told* what to do when. Still we can say that the situation exerts stimulus control over our actions if, in fact, we emit different operants in different situations.

In short, discriminative stimuli are not eliciting stimuli. Rather, they provide *information.* The stimulus tells the organism that the situation is one in which that operant will be reinforced.

### The Terminology of Reinforcement and Punishment

If a rat is allowed food after pressing a bar, and bar-pressing increases as a result, then we speak of the food as a **positive reinforcer**. It is something the animal will work to obtain. On the other hand, if bar-pressing serves to turn off a painful shock, then shock is a **negative reinforcer**. Its *offset,* rather than its *onset,* reinforces the bar-press; it is something the animal will work to get rid of. Finally, if we follow a response with the *onset of a negative reinforcer* such as shock, then we are employing **punishment**. Punishment is the *onset of a negative reinforcer, dependent on the occurrence of a response.* Thus punishment and negative reinforcement are not the same; don't confuse them.

Notice also that what is reinforced or punished is the response, not the actor. We don't reinforce rats; we reinforce lever-pressing. We don't punish a child, we punish its misbehavior. This sounds picky perhaps; but on the other hand, if we are careful about what we say, it may help us be careful about what we do.[2] Suppose we have a child who snitches cookies between meals, and we wish to suppress this. We see it happen, call the child to us, and administer a punishment. Now the child has been punished, all right. But the *response* that has been most immediately punished is not snitching cookies but coming to us. The child may come to inhibit approaches to us—the punished response—rather than snitching cookies.

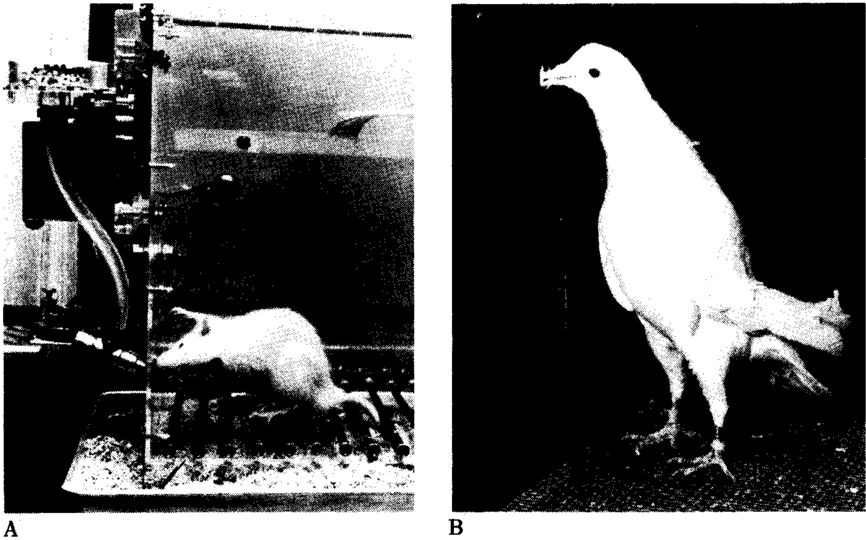
We have introduced the most important concepts used by modern reinforcement theorists: The operant-respondent distinction, the concept of reinforcement, and the concept of stimulus control. Now let us look more closely at some of the ways in which these concepts are investigated. Then we will see some ways in which they are applied.

## METHODS IN THE STUDY OF OPERANT BEHAVIOR

The phrase **operant conditioning** covers any experiment in which a reinforcer is delivered if, and only if, a specified response occurs. The famous maze, in which the rat must make a correct turn or series of turns to collect its reward, is technically a case of operant conditioning.

There is a problem with mazes and related methods, however. The rat, let us say, reaches the goal and eats its food. The experimenter must corral the animal and, at some later time, return it to the apparatus for another trial. The animal's stream of behavior is broken up into a series of distinct events called **trials**, whereas real-world behavior is a continuous flow.

[2]See Catania, 1979.



A B

**Figure 8-1.**
Experimental chambers for the study of reinforcement. (A) A rat pressing a lever for small pellets of food. (B) A pigeon pecking a lit pecking key for food reinforcement.

## The Free Operant

Because it preserves that continuity of behavior, the **free operant** method has become the most popular way of studying operants. Here we encounter the famous Skinner box (Figure 8-1).

The subject is enclosed in a small chamber, with some well-defined response available at all times: a lever for a rat to press; a disk or *pecking key* mounted on the wall for a pigeon to peck; or a button to be pressed by a human subject. When the response is made, reinforcement is delivered. A pellet of food may drop into a food cup for the rat. A hopper full of grain may be brought up so that a pigeon can feed from it for a few seconds. Or a counter may record points earned by a human subject.

After reinforcement occurs, the situation does not change. The subject need not be removed and replaced, and there need be no trials. The rat can simply emit the operant again, and earn another pellet. The lever is always available, and the rat is free to press it at will, so to speak, over and over again. That is why this procedure is called *free operant.*

In such an experiment, the most obvious measure of the strength of an operant is its frequency, or **rate of occurrence.** A bar-press response that occurs more rapidly, over time, is presumably stronger than one that occurs only at a low rate, with long delays in between. More generally, one can think of the rate of a response as an estimate of its *probability* —that is, its frequency of occurrence relative to other responses that the situation allows. Obviously, the greater the rate at which a rat presses the

lever, the less time it is spending doing other things; and, therefore, the greater is the probability that it will be bar-pressing at any given instant.

With this method, one can watch the rate of a response change *continuously* over time, without interruption. Reinforce key-pecking with food, and a hungry pigeon will very shortly come to peck at a high rate. Stop reinforcing the response and its rate goes down again; this is *extinction* of the response. If we reinforce key-pecking only when a tone is sounding, or only when the key is illuminated with green light, then the response will undergo extinction at other times, and the tone or the color of the key will acquire *stimulus control* over the rate of emission of the operant.

### Schedules of Reinforcement

Of course, there is no law that requires *every* emission of the operant to be reinforced, even when reinforcement conditions are in effect. Instead, one may use **partial reinforcement**, in which only some occurrences of the response are reinforced. Very often, a **schedule of reinforcement** is imposed; that is, some rule relates the occurrence of reinforcement to the passage of time or to number of responses.

In an **interval schedule**, reinforcement is made available by a timing device. For example, a 5-minute fixed-interval schedule would be one in which, after each reinforcement is delivered, 5 minutes goes by before the next reinforcement is available. After 5 minutes is up, then the next response is reinforced. Responses made during the 5-minute period are recorded, but not reinforced.

Or one may impose a **ratio schedule**, where the ratio is number of responses to number of reinforcements. On fixed-ratio 10, for example, every tenth response is reinforced (that is, responses/reinforcements = 10).

Partial-reinforcement schedules such as these may lead to remarkably persistent responding. It is possible to generate an enormous amount of behavior for very few reinforcements—a point, reinforcement theorists would say, that has not been lost on owners of slot-machines and similar devices. The large literature on reinforcement schedules is discussed in books on learning; we simply lack space to consider it further here.

## APPLICATIONS OF REINFORCEMENT

The use of reinforcement extends far beyond the laboratory. In recent years, psychologists dealing with personal, academic, and industrial problems have paid close attention to the possible role of reinforcement in complex human behavior. Many of the behavior therapist's techniques depend upon the identification, and modification, of the conditions of reinforcement that operate in natural settings. Stated briefly, these inter-

ventions revolve around the attempt to see that appropriate behaviors get reinforced—and that inappropriate behaviors do not.*

## Environmental Maintenance of Disordered Behavior

One of the best examples of reinforcement therapy is also one of the earliest.[3] In mental hospitals, severely disturbed patients all too often receive little more than custodial care. Their needs are met, things are done for them that they cannot or will not do for themselves; but they may receive little attention or social interaction unless their behavior becomes bizarre, assaultive, or self-destructive. In that case, nurses and attendants may flock to calm them down.

This caring, helping behavior is of course understandable and its intentions are admirable. But what if attention and soothing actions serve to reinforce disruptive behavior? What if doing things for the patient reinforces his not doing them for himself? Quite possibly we are reinforcing, and thus maintaining, the very kind of behavior we would like to get rid of—for the patient's benefit even more than our own.

Two therapists recognized this situation and attempted to turn it around—to *stop reinforcing inappropriate actions.* They began by training, not the patients, but the psychiatric nurses at the hospital where they worked. The concept of reinforcement was explained to the nurses in ways like this:

> Reinforcement is something you do for or with a patient, for example, offering candy or a cigarette. Any way you convey attention to the patient is reinforcing. . . . When we say "do not reinforce a behavior," we are actually saying "ignore the behavior and act deaf and blind whenever it occurs."[4]

Training the nurses was not easy. Yet the results showed some dramatic changes in the patients' behavior. One female patient persistently entered the nurses' office and distracted them from their work. Previously, this behavior had been reinforced! The nurses would take the patient by the hand and lead her away. When they ignored her instead, her intrusions underwent extinction and occurred much less often (Figure 8-2). This shows that the nurses' previous reaction was in fact a reinforcer for the disruptive actions. In other words, what was maintaining those actions was not the patient's illness alone, but also the *environmental consequences* of the actions—that is, the nurses' reactions to them.

Two other psychotic patients refused to feed themselves, and had to be spoon-fed by the staff. In seeking an effective reinforcer for self-feeding,

*For an excellent survey of both the rationale and the findings of these methods, see Rimm and Masters, 1974.

[3]Ayllon and Michael, 1959.

[4]Ibid., p. 325.

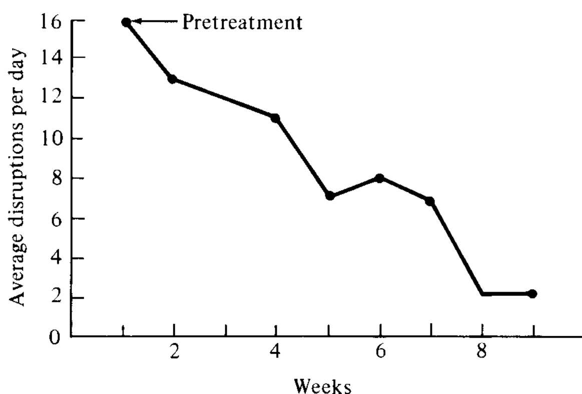



**Figure 8-2.**
Declining frequency with which a patient entered the nurses' office and disrupted it when the nurses stopped reinforcing these intrusions. (From Ayllon and Michael, 1959.)

Ayllon and Michael looked carefully at these patients' likes and dislikes. It turned out that both patients were concerned about the neatness of their clothes, and didn't like to get stains on them. So the nurses were instructed as follows: The patients would be spoon-fed if they insisted, but it was to be done without conversation, and carelessly, so that drops of food would fall on a patient's dress. If the patient fed herself, then the nurse was instructed to remain with her and talk with her. Here again we see social attention used as a reinforcer, this time combined with an *avoidance* procedure (Chapter 7): To avoid spots on her dress, the patient must feed herself.

The result was a dramatic increase in self-feeding by both patients. Indeed, one of the patients, who initially weighed only 99 pounds, gained more than 20 pounds over the next 8 weeks.

Does this sound like a cold and heartless procedure—deliberately staining a patient's dress, and refusing to converse with her unless she behaved "properly"? One could argue that it is. On the other hand, is it kinder to keep a patient indefinitely in a state of helpless, spoon-fed passivity? The nurses' previous actions were doing just that, as shown by what happened when those actions were changed.

The use of such procedures is not limited to institutions. Family members at home also may be powerful sources of reinforcement—for better or for worse. For example, there is the case of a woman who went into a deep depression following the death of her mother.[5] The actions typical of depression—lassitude, weeping, expressions of helplessness—met with sympathy and consolation from the woman's husband and children. This

[5]Liberman and Raskin, 1971.

is understandable, but shortsighted too, if these attentions were acting as reinforcers for the unproductive and miserable depressive actions.

On the advice of the therapists, the woman's family members agreed to change their reactions. As in the hospital study, the family members forced themselves to ignore depressive behaviors, but to respond with attention and support when positive coping behaviors took place.

Figure 8-3 shows the results. Depressive behaviors were high, coping behaviors low, over a 1-week baseline period. When the new reinforcement conditions were put into effect, these two classes of behavior promptly changed places. When the initial conditions of reinforcement were reinstated, depressive behaviors again began to climb; but then, when reinforcement for coping behaviors was applied again, coping behaviors rose again, and depressive behaviors fell to low levels and remained there. We see again the reinforcement of appropriate behaviors combined with the extinction of inappropriate behaviors.

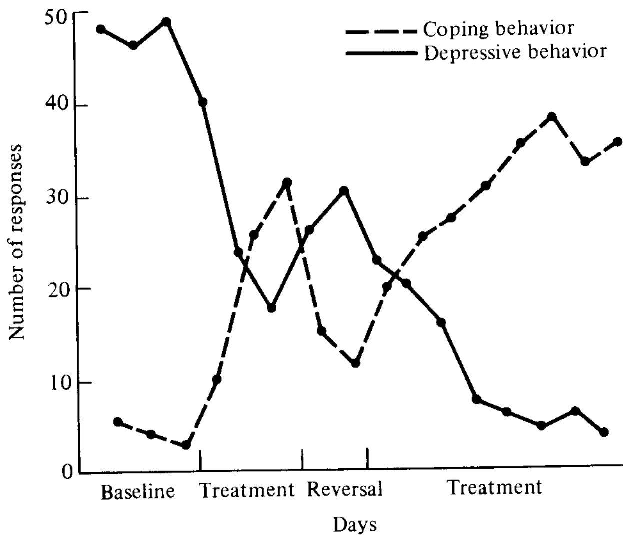



**Figure 8-3.**
Frequency of depressive and of coping responses. Before treatment (Baseline), when the family reinforced depressive responses, coping responses were few and there were many depressive responses. The relative frequencies reversed when treatment (the new reinforcement conditions) was imposed (Treatment). They reversed again when the family returned to its previous reactions (Reversal), but depressive responses fell once again and remained low when the new program was restored and left in effect (Treatment). (From Liberman and Raskin, 1971.)

## Behavior Modification in the Classroom

Experiments much like the hospital one have been conducted in schools, to reduce disruptive behavior in the classroom.[6] But my favorite example is the delightful study in which a class of junior high school students was given a cram course in reinforcement theory, and it was pointed out to them that they could increase the number of pleasant contacts between the teacher and themselves by reinforcing pleasant behavior on the *teacher's* part. They could reinforce it by smiles, sincere efforts to learn, and discussions before and after class.[7] But these reinforcers were to be delivered only when the teacher was pleasant. They were not "free" reinforcements; the teacher had to earn them.

As Figure 8-4 shows, pleasant student-teacher contacts were greatly outnumbered by snarly ones, at the beginning—but not after 6 weeks of teacher modification by students.

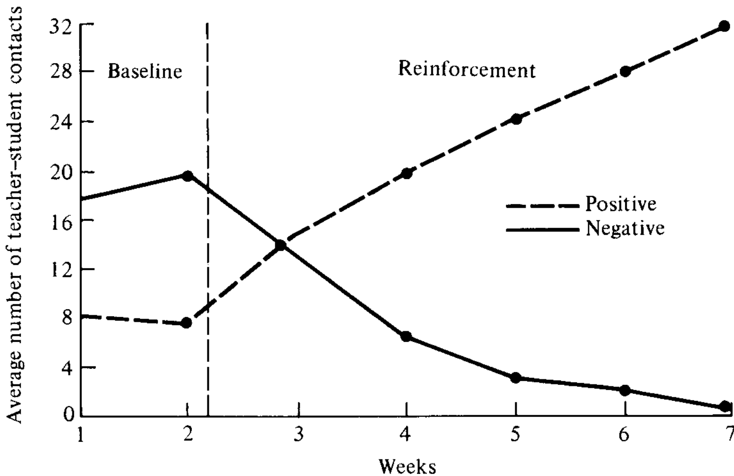



**Figure 8-4.**
Teacher-student contacts were more often negative than positive during the pre-experimental baseline period. Positive reactions were much more frequent when students reinforced the teacher's friendly behavior. (From Gray et al., 1974.)

## Token Reinforcement in the Workplace

For the next example, recall our discussion of *conditioned reinforcers.* A poker chip, or similar token, can be made exchangeable for some more

[6]See for example Hall, Lund, and Jackson, 1968.
[7]Gray, Graubard, and Rosenberg, 1974.

natural reinforcer, such as grapes for a chimp or privileges for a hospital patient. The token then becomes a reinforcer—a conditioned reinforcer.

Because they are cheap and portable, such tokens can be used when it would be difficult to follow up desired behavior with a natural reinforcer right away. In some hospitals, whole **token economies** have been set up, based solely upon tokens.[8] Even short of that, they can be highly effective.

In a factory in Mexico City, there was a severe problem of worker tardiness. Seven hundred fifty instances were accumulated in 1970 alone. So a team of researchers tried a program of token reinforcement: Arriving for work on time was reinforced with tokens, each exchangeable for a small amount of money. What happened? The incidence of tardiness fell from about 10% to about 2%, and stayed there.

These findings are all the more striking when we consider that the workers could earn the equivalent of about $40 a year with the extra bonus system—and that the company had already been paying an *annual* bonus of $40 for consistent punctuality, with no effect at all! It seems that *frequent* and *immediate* reinforcement has an impact of its own, apart from long-term gain.[9]

## Reinforcers in the Natural Environment

Reinforcement theory holds that all purposive behavior is maintained by its reinforcing consequences. Therefore, it becomes important to identify the reinforcers that operate in natural settings—especially the reinforcers by which we, as social creatures, affect each other's behavior.

In families of aggressive children, Gerald Patterson and his colleagues have tabulated the occurrence of various kinds of interactions among family members, and the characteristic sequences that they follow.[10] A common one goes like this: The mother is engaged in some activity, the child makes a demand, the mother ignores the child, the child whines, the mother begins to attend to the child, and the child stops whining.

Each event in this chain affects the next. The child's whining is unpleasant to the mother. By attending to him, she removes that unpleasant stimulus; this is *negative reinforcement* for her attending, and makes it likely that she will do it again next time. Unfortunately, her attending acts as a positive reinforcer for the child's whining, making it more likely that *he* will do it again next time. So the mother reinforces the child's whining, and the child reinforces the mother's giving in—angrily and resentfully —to the whining.

Patterson calls such interactions **reinforcement traps**. Each participant's behavior is reinforced when it occurs. But that means that, in the long run, each reinforces the other's tendency to behave in an unpleasant way. So the level of unpleasant interactions remains high. Patterson's

[8]See Rimm and Masters, 1979.
[9]Herman et al., 1973.
[10]See for example Patterson, 1975.

group has also shown how such patterns can change, if one of the pair is taught to react to the other in a different way.

## A Look Backward: The Operant Philosophy

In running through these examples, we see not only the power of reinforcement, but also a definite point of view about the nature of behavior, its causation, and its problems. This operant philosophy, a product of the behavioristic approach, has certain emphases and assumptions that characterize it.

### The Focus on the Environment

In all our examples, we see behavior therapists addressing a problem *behavior,* and attempting to modify *the behavior itself.* And this is done by modifying the *environmental conditions that affect that behavior.* One does not ask what the behavior means to the actor, or refer it to an underlying anxiety or personality trait. The operant philosophy assumes that conditions *outside the person*—stimulus control and the reinforcements the environment provides—provide the answer to the question, "Why does she do that?"

Allow me a personal example, which brings the point home to me each semester. I routinely set up class demonstrations of operant conditioning in rats. Of course, few things run smoothly for very long, and often I find my rats behaving improperly—not displaying the behavior I wish to show my students. Something is wrong, and I must find it and fix it. And, invariably, what I find and fix is not in the rats but in the apparatus. The fault in behavior is a fault in the environmental conditions, set up by the apparatus, that produce that behavior. "The organism is always right," Skinner has said; it does what it ought to do under the conditions imposed upon it. That is the operant philosophy in a nutshell.

A behavior therapist would address problematic human behavior in the same way. He does *not* assume that "people are like rats"; no sane person does. He does assume that behavior, in people or in rats, occurs in the environment and is responsive to that environment, and will change if the environment does.

Of course, especially in the human case, other people are a major part of one's environment. That is why a behavior therapist so often has to begin his work with the nurses, teachers, or family members who provide reinforcements for people. This can present great difficulties of its own, for apart from the natural reaction "Why are you talking to *me? He's* the one with the problem!", one may have to combat very strongly entrenched attitudes, habits, and vested interests in those who interact with the "problem person."

### Identification of Behavioral Goals

To produce the desired behavior, we must set up the environmental conditions that will in turn produce it. This means that we must think about

the behavior we wish to bring about, at least as carefully as we think about the problematic behavior we wish to change.

If this seems obvious enough, think about how persistently we refuse to consider it. Psychiatric nurses attend to disruptive behavior, not to effective social behavior, in their patients. In the workplace, Dale Carnegie saw the paradox over half a century ago. How does a supervisor, as part of the worker's "apparatus," typically behave? "If he doesn't like a thing, he raises the Old Harry; if he does like it, he says nothing."[11]

An insightful paper written for business managers[12] points out "the folly of rewarding A, while hoping for B," and describes case after case of it. In politics, we *hope* that candidates will have in mind clear and specific solutions to problems, but we *reward* statements that "may be relied on to offend absolutely no one,"[13] and wonder why we have few statesmen. Businesses *hope* that administrators will consider costs and opportunities over the long term, and initiate programs that will pay off some years down the road. But if they *reward* only short-term sales or profits, they virtually ensure that only the short term will affect the actions actually taken. Many firms *hope* for performance but *reward* attendance, and attendance is what they get. And schools—at all levels? What is hoped for? What is rewarded? In your experience, how does the one match up with the other?

## THEORETICAL ISSUES IN REINFORCEMENT

We have seen examples of how reinforcers affect behavior in a variety of settings. Now we need to take a closer look at just what we are doing when we reinforce a response. What is a reinforcing event, anyway? What does it do and how does it do it?

### What Is a Reinforcer?

Skinner defines a reinforcer by its effects. A reinforcer is something that reinforces. It may indeed turn out that that is the best we will be able to do. But it is natural to wonder whether different reinforcing events have something in common. If they do, and if we knew what it is, we would be able to specify in advance whether a given event will be a reinforcer or not. There have been several suggestions.

#### THE HEDONIC HYPOTHESIS

The notion that *we seek what brings us pleasure* is a venerable one, dating back through Hobbes to the Greeks and probably much farther than that (Chapter 2). One might suggest that the reinforcement principle is just a special case of that common-sense axiom. Food is pleasant to a hungry

[11]Carnegie, 1936.
[12]Kerr, 1975.
[13]Ibid., p. 769.

rat, attention is pleasant to a disruptive patient (is it?), and so these organisms act in ways that produce these pleasures.

There are some obvious difficulties here. First, we cannot ask an animal what is pleasant.* And if we infer pleasure from behavior, we end up arguing in a circle. Why does a hungry rat work for food? Because food is pleasant. How do we know it is pleasant? Because the rat works for it. We must do better than that.

### The Drive-Reduction Hypothesis

Clark Hull's theory of behavior (see pp. 192–194) included the reinforcement principle explicitly, and he linked that principle to the concept of *homeostasis.* The idea is this: If a biological drive exists, and if a response results in a reduction of that drive—thus reducing the homeostatic disturbance as well—then that response will be reinforced. In other words, *reinforcement occurs when a drive is reduced.* That way, the animal learns to do what meets its body's needs.

The **drive-reduction hypothesis** was a plausible and attractive theory. Unfortunately, there are many powerful reinforcers that cannot be shown to have anything to do with meeting physiological needs.

Rats, for instance, will repeatedly run down an alley for an opportunity to copulate, even if ejaculation—the presumed drive reduction—is not allowed.[14] Rats will drink a saccharin solution avidly—that is, the solution will support the operants of going to where the bottle is, and lapping from it—even though the rats are not deprived and the saccharin has no food value.[15] Monkeys will take apart a mechanical puzzle, time after time, for no reward but the opportunity to do it again.[16] And teachers become more friendly if their students are pleasant and attentive; what drive does that reduce? The conclusion is clear: Drive reduction characterizes some reinforcers, but not all.*

### Response Factors and the Relativity of Reinforcement

A quite different approach to the nature of reinforcement shifts the emphasis from the reinforcing stimulus or event, to the *response* that such an event permits. We present food to a hungry rat, and the food permits the rat to eat. What if the actual reinforcing event is not the food itself, but the *opportunity to eat?*

*We can ask a human, it is true; but even then we must wonder what the answer means. Do you mean by pleasure what I do? Is my red the same as your red?

[14]Sheffield, Wolf, and Backer, 1951.

[15]Ernits and Corbit, 1973; Sheffield and Roby, 1950.

[16]Harlow, Harlow, and Meyer, 1950.

*Of course we could hypothesize drives to be reduced as the case requires—for example, an exploratory drive for the monkeys, reduced by exploring; or an affiliative drive for the teachers, reduced by the students' sociability. And it must be said that some writers have played that game of pseudo-explanation with enthusiasm. Some suggested, for instance, that Harlow's monkeys were anxious in the presence of a mechanical puzzle, until it had been thoroughly explored and found safe. Harlow's answer was eloquent: "It is not the monkeys that are anxious, but the drive-reduction theorists."

Strong evidence for such an idea was provided by an ingenious experiment by David Premack.[17] Consider a rat that lives in a cage, with food, water, and a running wheel available. The rat spends a certain amount of time drinking, a certain amount running, and a certain amount eating.

Now we impose some rules. The running wheel is locked, so that the rat cannot run in it—unless he drinks some water first. Now drinking is the operant, and opportunity to run is the reinforcer. Sure enough, under these conditions the amount of drinking rises. Opportunity to run will reinforce drinking behavior.

Now we reverse the situation. Water is withheld from the rat, unless it runs in the wheel. Running in the wheel makes water available for a short time, and then the rat must run some more to obtain more water. Sure enough again: The amount of running increases. Opportunity to run will reinforce drinking; but also, opportunity to drink will reinforce running.

This demonstration makes the important point that we cannot think of reinforcers just as events of a particular kind. It is not that some *stimuli* are reinforcers and others are not. Instead, the reinforcing event may involve the relation between one response and another. One response is reinforced because it permits another response to occur.

This takes us to the **response-deprivation hypothesis,**[18] which says: If a response is prevented from occurring as frequently as it normally would, then the opportunity to engage in that response is a reinforcer. Thus, opportunity to run will reinforce drinking, but only if the rat has first been prevented from running. Opportunity to drink will reinforce running, but only if the rat has first been prevented from drinking.

Thus we see that the operant-reinforcer relation is not fixed but relative. Under some conditions, A will reinforce B; but under others, B will reinforce A. Reinforcers cannot be defined independently of the response they reinforce, the constraints on the emission of that response, and the frequency with which it would occur if those constraints were not there.

This idea explains how drive states produced by deprivation set the reinforcing power of various commodities. Food deprivation, we say, makes food a reinforcer. But a better way of saying it might be: Deprivation of opportunity-to-eat makes opportunity-to-eat a reinforcer. The deprivation need not be tied to homeostatic needs. Deprivation of opportunity-to-run makes opportunity-to-run a reinforcer, though there is no homeostatic *need* to run.

The idea may also make sense of some well-known phenomena that at first glance look problematic for reinforcement theory. For instance, food is often a good reinforcer, and a child could be rewarded with food for performing some chore. However, eating food can also be an operant that gets reinforced, rather than a reinforcing event. Consider how children may be persuaded to finish their dinners by allowing them to play only

[17]Premack, 1962.
[18]Dunham, 1977.

after they have done so. In the first case, opportunity to eat is denied until the chore is done; so opportunity to eat is a reinforcer. In the second case, opportunity to play is denied until the meal is eaten; so opportunity to play becomes a reinforcer, and eating becomes the operant.

Or consider a child doing poorly in school, who plays with friends a lot but seldom reads a book. A counselor, parent, or behavior therapist might prevent the child from playing until after a period of reading. On the other hand, what about a bookish child who seldom plays with others, and so is not picking up the social skills he needs? For such a child, it might be best to turn the situation around: Access to a book could be denied until the child has played with other children for a while.

The response-deprivation hypothesis has much to recommend it. We still have to wonder, though, whether it can really encompass *all* instances of reinforcement. What about the teacher's actions reinforced by students' behavior? Was there any response the teacher could then make, that she was otherwise prevented from making? We could speculate about that, but certainly it has not been demonstrated. So, although this idea fits many cases of reinforcement, we cannot say that it fits all.

Moreover, there can be real difficulties in establishing just when one is deprived of this or that response. We speak of response deprivation if the response is restricted below its spontaneous rate of occurrence. Well, suppose we go to a movie once a week, on the average. Suppose an experimenter wants to deprive us of movies, and doesn't let us go. At what point do we become "movie-deprived"? After the *average* inter-movie interval? After the *longest* time without a movie that we have let pass spontaneously? Or what? The concept of response deprivation is less precise than it sounds.

### The Empirical Law of Effect

We have seen three answers to the question: What is a reinforcer? One is hard to evaluate (the hedonic hypothesis); one clearly won't do for all cases (the drive-reduction hypothesis); and the third may or may not apply to all cases (the response-deprivation hypothesis). In practice, most reinforcement theorists have fallen back on the **empirical law of effect**. An event is reinforcing if it reinforces.

This simply means that the reinforcing properties of an event are demonstrated by the effect of that event. For example, if we find that a child shakes a rattle that makes noise, and goes on shaking it; stops shaking it if it stops making noise; and goes back to shaking it if it begins making noise again—then we have discovered that the noise is a reinforcer for her rattle-shaking operant.

That conclusion leaves open the possibility that different events may reinforce for different reasons. And, after all, there is no particular reason why all reinforcers should work in the same way. Consider three people who emit the operants of walking to the coffee shop and paying for coffee. For Anne, the reinforcer might be the taste of the coffee. For Brad, it

might be the lift that the caffeine provides. For Carol, it might be the opportunity to engage in conversation with the other two.

Here are three different reinforcers for the same operant. And these *are* reinforcers, in the strict technical sense of the term. They maintain coffee-shop-going behavior, and if they were no longer forthcoming the behavior would extinguish. If the shop began serving bad-tasting coffee, Anne would stop going. If it began serving coffee with no caffeine in it, Brad would stop going. And if Anne and Brad stopped going, Carol would stop going. The reinforcers are very different, but they all reinforce.

In this sense, the reinforcement principle is somewhat like the Darwinian conception of **fitness**. One form of a species is fitter than another, by definition, if its reproductive success is greater. But one form may reproduce more than its cousins because it is stronger, or because it is smarter, or because it is more cowardly and thus escapes predation better, and so on. Fitness can take many forms. Perhaps *reinforcement* has many possible mechanisms.

## What Do Reinforcers Do?

As we have seen, a reinforcer is defined as a stimulus that increases the strength—that is, the probability—of a response that it follows. However, there are several ways that a stimulus could have that effect.

The most obvious starting point is to take the strengthening idea literally. A reinforcing stimulus might directly increase the tendency to emit the most recent response. This idea may be correct, but it probably is not the whole story. We will return to it, but first let's look at another possibility.

### Reinforcement as Selection

If there are only so many responses that might occur, and if one comes to occur with high probability, that could be because it has grown stronger (see below). Or it could be because the others have grown weaker! The effect of reinforcement, then, might be to *protect* the reinforced response from the weakening that unreinforced responses undergo. John Staddon and Virginia Simmelhag[19] performed a classic investigation that not only supported this idea, but also pointed out some of its important ramifications.

A hungry pigeon was placed in an experimental chamber, with a disk that it could peck. Pecking, however, was *not required* to produce food. Rather, brief access to food was provided automatically, irrespective of what the pigeon did, at regular 12-second intervals. Observers watched the birds, tallying various behaviors as to how often, and when, they occurred.

During the intervals between food deliveries, the birds would engage

[19]Staddon and Simmelhag, 1971.

in a variety of activities: wing-flapping, turning in circles, preening, etc. These behaviors were highly variable from one bird to another. But as the interval neared its end and the next food delivery approached, all the pigeons began to peck at something. Some might peck at the pecking key, some at the feeder, some at spots on the wall—but they all pecked at *something.* These pecks *had no effect on the delivery of food,* but they appeared anyway as the inter-feeding interval approached its termination.

Then Staddon and Simmelhag changed the rules. Now, access to food was made available at the same intervals as before, but it was no longer delivered automatically. It took a peck on the pecking key to produce the food. In other words, every 12 seconds the apparatus set up a reinforcement, so that the next peck at the key would deliver it. But a key-peck was required now, whereas it had not been before.

What did that change do to the frequency of pecking? Nothing at all. Pecking did not become any more frequent than it had been before; but pecking *at the key,* rather than at something else, did. In other words, when key-pecking was specifically reinforced with food, the *frequency* of pecking was not affected. Only the *location* of the pecks was changed.

In interpreting these data, Staddon and Simmelhag suggest the following. In situations that are associated with (or predict) the arrival of food, certain actions are likely to occur. These are not produced by reinforcement. They are behaviors that occur when a motive is present, but cannot be satisfied.

What responses occur under such conditions? That varies with the species; in pigeons, pecking is definitely included. What responses occur is determined by **principles of variation**—genetic factors that produce instinctive action patterns (see pp. 101–104), and prior learning experiences. These principles of variation determine which members of an organism's repertoire *might occur,* or are *likely to occur,* in a given situation.

When a response is required to produce food, then the reinforcement principle takes over as a **principle of selection.** Its effect is not to *strengthen* the reinforced response, but to *select* that response from the various responses, already strong, that the principles of variation make available to the organism (Figure 8-5). Reinforcement determines which of the responses that *might occur* is the one that *will occur.* And since the reinforced response is not strengthened by reinforcement, for it is already strong, its occurrence with increased frequency must mean that other responses that might occur become *less* likely. It is not that reinforcement strengthens; it is that non-reinforcement weakens. When key-pecking was required to produce food, pecking did not become more frequent. Rather, pecking at things other than the key became less so.

This theory of **reinforcement as selection** bears an interesting parallel to the notion of *selection pressure* in evolutionary theory. Species evolve, not because their environment produces new structures that are adapted to it, but because it *selects* life forms with structures that already are

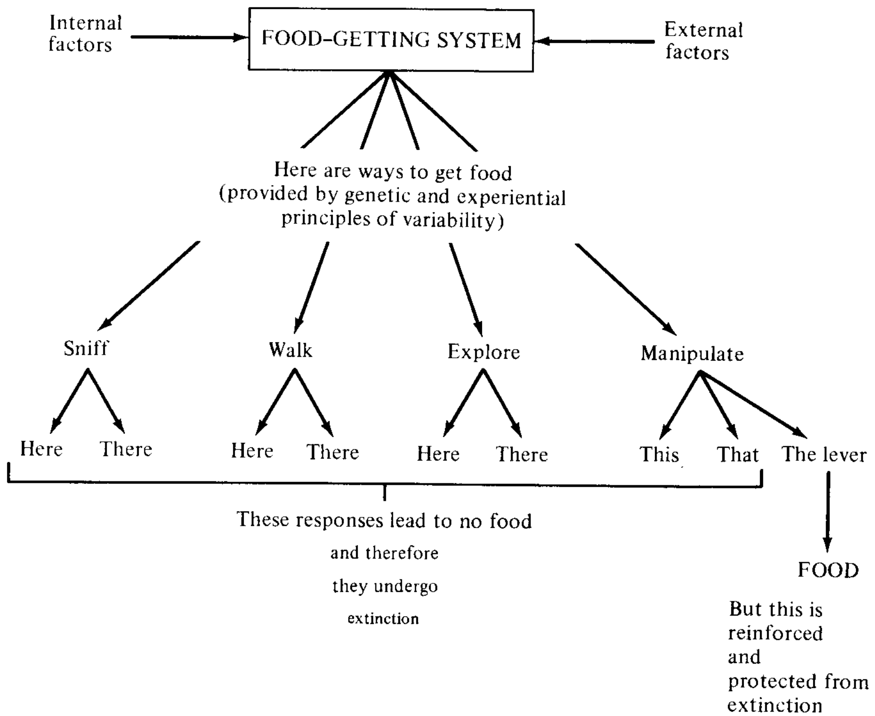



**Figure 8-5.**
Reinforcement as selection. Nonreinforced responses undergo extinction; the effect of reinforcement is to protect the response that precedes it from that extinction.

adapted to it. And it selects these from the life forms that already exist. Forms of life that are poorly adapted to their environment have few descendants and so undergo extinction, leaving only the well-adapted forms. It is the same for behavior. Forms of behavior—responses—undergo extinction if *they* are poorly adapted to the environment. If a response is poorly adapted—that is, if it doesn't produce reinforcement—then it drops out. That leaves the well-adapted, successful responses. In other words, an animal adapts over its lifetime, and a species adapts over many generations, by the same process—trying various things (variation), and discarding what doesn't work (selection).

The theory also makes contact with another idea that we are familiar with. It is virtually identical with Bolles's analysis of avoidance conditioning (pp. 236–238). Training in avoidance, Bolles argued, does not teach the rat new tricks. It teaches it which of the tricks it already knows is the one that works.

Staddon and Simmelhag's analysis can be seen as a generalization of Bolles's idea. Perhaps operant conditioning in general does not teach an organism new tricks. It teaches it which trick, already "known" or available through the principles of variation, is the one that works.

### REINFORCEMENT AS STRENGTHENING

Although the idea that reinforcers protect responses from weakening has much to recommend it, there are other data suggesting that reinforcers may strengthen responses directly. Here is one example.[20]

In this experiment, pigeons were confronted with a 4 × 4 matrix of lights, one of which was illuminated, and two pecking keys (Figure 8-6). The rules of the game were as follows: The upper left key was illuminated. A certain number of pecks at one key drove the light one row to the right. A certain number of pecks at the other key drove the light one row down. The bird's task was to get the light to the lower right position. Then the pigeon was paid with food and the light was reset to upper left for the next trial.

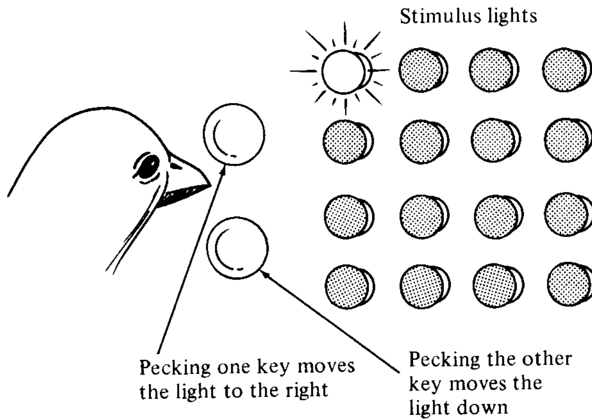



**Figure 8-6.**
The Vogel and Annau experiment. In the starting position, the upper left stimulus light is illuminated. By pecking the keys that move the illuminated light from one location to another, the pigeon must get the light to the lower right position.

In this case, we consider a *response* not just a key peck, but a whole sequence of pecks that controls a series of movements of the light. It is easy to see that there were many possible responses in that sense, any of which would be reinforced—that is, there were many ways of getting the light from top left to bottom right. The bird could have moved the light four places to the right, then four down; or four down, then four to the right; or one to the right, then one down, then one more to the right, and so on.

In fact, early in the experiment the birds did use a variety of sequences. But as sessions went on, each bird tended to settle on some one sequence and to repeat it trial after trial. The behavior, in other words, became

[20]Vogel and Annau, 1973.

more fixed and stereotyped than the reinforcement conditions required it to be.

Why did the pigeons do that? What happened to all the other sequences that were tried early, and that worked, but were abandoned as training went on? We cannot say that they were weakened by non-reinforcement. They were reinforced when they occurred, but they stopped occurring anyway.

The data make sense if we suppose that when a given sequence worked, and was reinforced, it was made *stronger.* Then, suppose that for whatever reason, one or another sequence happened to occur more frequently than any other. Each time it occurred it would be reinforced and strengthened; then it was even more likely to be emitted again, and made stronger still, and so be still *more* likely to be emitted again, and so on, until it became by far the most probable sequence. If that is what happened, then it means that a reinforcement can in fact make a response stronger, and not just protect it from weakening.

A Look Backward: Strengthening or Selecting or Both?

Where are we? There is evidence that reinforcement selects responses by protecting them from weakening. There is other evidence that it can strengthen a response directly. Perhaps it does both. Just as different reinforcers may be reinforcing for different reasons, so their reinforcing effect may be exerted in different ways—protecting the reinforced response, or strengthening it, or both.

## BIOLOGICAL CONSTRAINTS ON REINFORCEMENT

Throughout this discussion, I have made no mention of the hierarchical structures of motivational systems. Until recently, reinforcement theory took little account of them. Why? Because it was assumed that the *environment* determined these structures.

> In any given environment an organism confronts a broad . . . set of options in dividing its behavioral stream into response classes. The contingencies of reinforcement carve the operants out of the behavioral stream.[21]

What that means is this. A rat's lever-pressing response, like a stickleback's reproductive behavior, can be described at any level we choose —the level of muscle twitches, or of paw movements, or as anything that gets the lever down. At what level should we describe it? What is a response, anyway—an outcome of action (the lever gets depressed), or a muscle movement? Reinforcement theorists answer: A response is whatever the conditions of reinforcement specify.

So, if anything that gets the lever down is reinforced, and nothing else

[21]Herrnstein and Vaughan, 1980, p. 173. See also Schwartz, 1981.

is, then lever-pressing is the operant. But suppose we require the rat to press the lever with its left paw, so presses with the right paw are not reinforced. Then pressing with the left paw is the operant. And we could, if we chose, somehow record and reinforce only a left-paw lever press of specified force. Then that kind of lever press is the operant. In short, any action at any level can be an operant, if the *conditions of reinforcement* —in the *environment*—make it one.

Now if it were only the conditions of reinforcement that "carve the operants out of the behavioral stream," then we could ignore the hierarchical structure of action that is built into the animal by evolution. We ought to be able to define any operant, provide reinforcement for it, and increase the frequency of just that operant and nothing else. In recent years, however, it has become increasingly clear that we cannot ignore the wired-in connections between one movement and another. Certain actions are coupled with certain others, at higher and at lower levels—where the conditions of reinforcement do not provide these couplings. This can best be made clear by turning at once to some examples.

## Top-Down Constraints: The Misbehavior of Organisms

The first set of examples was also among the first to be discovered. Briefly, what they show is this: If we activate a higher level of a lattice hierarchy, we may end up activating many, or all, of the lower-down components that are linked with it. And this can happen whether the reinforcement conditions specify it or not, and whether we like it or not.

### The Case of the Hammering Chicken

The research that led to this discovery began as an exercise in applied learning theory. Keller and Marian Breland were training animals for commercial purposes: advertising displays in store windows, television commercials, and the like. Trained animals can make effective, eye-catching displays (Figure 8-7).

In some cases, the training, based on operant-conditioning principles, proceeded smoothly. In other cases, it ran into unexpected difficulties. The Case of the Hammering Chicken is illustrative of the snags that can arise:

> The observer sees a hopper full of oval plastic capsules which contain small toys, charms, and the like. When [a light] is presented to the chicken, she pulls a rubber loop which releases one of the capsules onto a slide. . . . The capsule rolls down the slide and comes to rest near its end. Here one or two sharp, straight pecks by the chicken will knock it forward off the slide and out to the observer, and the chicken is then reinforced by an automatic feeder. . . .
>
> However, a good 20% of all chickens . . . fail to make the grade. After they have pecked a few capsules off the slide, they begin to grab at the

**Figure 8-7.**
The piggy that went to market. The pig was trained with operant techniques to push a market cart. In this case, the experiment was successful; "misbehavior" did not intrude.

capsules and drag them backwards into the cage. Here they pound them up and down on the floor of the cage. Of course, this results in no reinforcement for the chicken, and yet some chickens will pull in over half of all the capsules presented to them.[22]

Obviously, the dragging-and-hammering behavior was not what the hungry chicken was reinforced for doing. It was not part of, and it interfered with, the food-getting operants. It is an example of what the Brelands happily call the **misbehavior of organisms.*** The animal simply does not do what it is reinforced for doing. It does something else instead.

We can understand this intrusive hammering action as a response that already is a component of chicken feeding behavior. The chicken that hammers capsules is exhibiting instinctive behavior having to do with breaking open seed pods and killing insects, grubs, and the like.

The Brelands provide several other examples of the misbehavior of

[22]Breland and Breland, 1961.
*The phrase is a takeoff on Skinner's classic book *The Behavior of Organisms.*

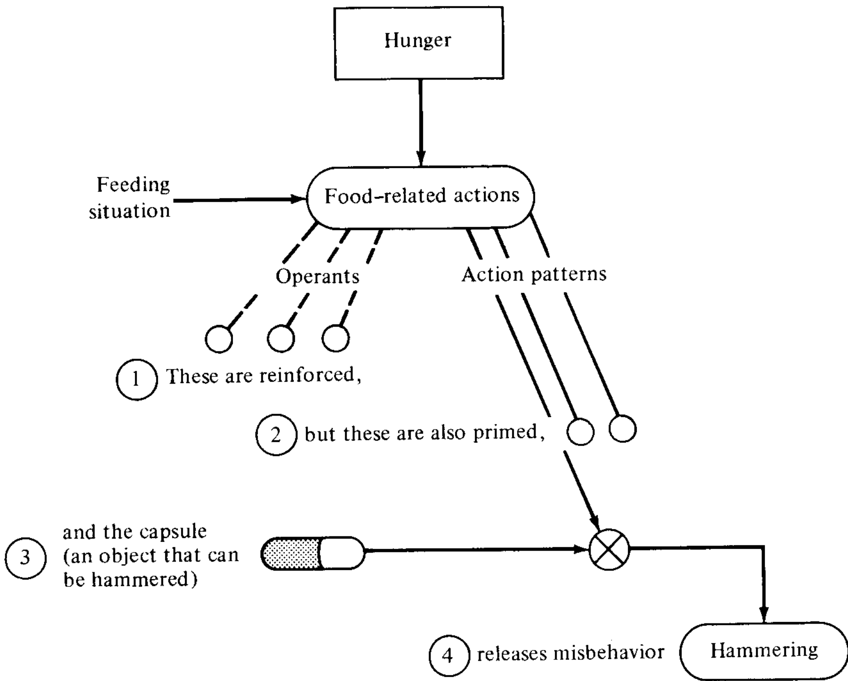



**Figure 8-8.**
How instinctive action patterns can intrude into operant-conditioning experiments, producing such misbehavior as the hammering pattern even though quite different responses are reinforced.

organisms. The general principle seems to be that when a motivational system is aroused, the sub-units of that system will occur—not because they are reinforced but because they *are* components of that system (Figure 8-8). Here, hammering is a component of the food-getting hierarchy that was activated in the conditioning situation. We might think of all those components as being primed, or facilitated, by higher levels. And the little capsules, which the bird was supposed to knock toward the observer, were enough like seeds or grubs to release the hammering action pattern instead. And so that action drifted into the bird's behavior, to the detriment of the response that the reinforcement conditions required.

### AUTOSHAPING

Another example is a phenomenon known as **autoshaping.**[23] Pigeons are placed in a Skinner box where a response key is available. From time to time the key will be illuminated for a few seconds, and then food is delivered. The food is free. The pigeons don't have to peck in order to get

[23]Brown and Jenkins, 1968.

it. But they begin to peck at the key anyway, and they will persist indefinitely in this useless behavior. That is autoshaping.*

Autoshaping does require that food be offered from time to time, and that its approach be signaled to the bird. Pigeons do not peck keys if they are not fed in the apparatus, so they are not pecking just for the fun of it. Further experiments showed that pigeons would persist in pecking, under such conditions, even if pecking *cost* them food.[24] In these experiments, each peck at the key would actually postpone the next delivery of food. To get fed, the pigeons had to *refrain* from pecking when the key illumination came on. But the birds would not or could not refrain. They persisted in pecking.

To make any sense of this, we must suppose that pecking is a strong component of a pigeon's feeding system. Pigeons peck in situations where food is forthcoming, whether or not they are reinforced for doing so, and even if they are punished for doing so by postponement of food. The pecking component of the hierarchy of pigeon feeding behavior is extremely difficult to suppress, even if it costs the bird food.

#### AUTOSHAPING, AVAILABILITY, AND VARIABILITY

Notice a close resemblance between the autoshaping experiments and the Staddon and Simmelhag experiment described earlier (pp. 280–282). Indeed, the two are virtually identical. In the autoshaping experiment, illumination of the key predicts food; in the other case, the passage of time does. And in both cases, pigeons peck when the situation predicts the arrival of food. In both situations, they peck when they don't have to; in the autoshaping case, they peck even when pecking costs them food.

Perhaps the *principles of variation* will help us understand both cases. We might express it this way: When food is coming, pecking is such a strong instinctive behavior in pigeons that it virtually *must* occur. The principles of variation do not leave the pigeon any options. Pecking, when food is coming, *will* occur because it is about the only thing that *might* occur. As a result, it is extremely difficult to teach a pigeon to refrain from pecking in order to get food. *Not pecking* simply is not a response that *might* occur; and so it *does not* occur.

### Bottom-Up Constraints: Recruitment

There is another kind of interaction among levels that has received remarkably little study, though it is potentially of great importance. It appears that under some circumstances, the activation of components can bring into operation, or **recruit,** the higher-order system of which the components form a part.

*The name comes from the practice of carefully training, or "shaping," the key-peck response for food reward in pigeons. It seems that the careful training is unnecessary after all. The pigeons will "shape" the key-peck themselves—hence, autoshaping, or "self-shaping"—even if reward does not depend on it.

[24]Williams and Williams, 1969.

A dramatic example of this sort of thing has been observed in pigeons.[25] The investigators were studying the operant conditioning of aggressive behavior. To obtain food reinforcement, their experimental pigeon was required to peck at another pigeon. The experimenters had no trouble training a bird to peck another bird, but the interesting thing was this. Once it had begun pecking the target bird, the experimental pigeon began launching full-blown attacks on that bird, with jumping and vigorous beating with the wings. These latter responses, *which were not required to produce reinforcement,* are naturally-occurring components of a pigeon's attack pattern. It appears that the act of pecking another bird was able to recruit the attack pattern as a whole, so that all its components were expressed.

In people as in pigeons, it seems that the activation of low-level components can bring whole systems into play. Many people have found, and it can be shown by direct experiment, that pretending happiness can actually improve one's mood. Pretending anger, as in a stage play, can increase one's self-reported feelings of anger (see Chapter 12).

A more involved and grimmer phenomenon *may* be another example of this kind of recruitment. This is the simulated-prison study conducted by Philip Zimbardo and his colleagues. They built an artificial prison in the basement of the psychology department building at Stanford University. Of the experimental subjects—normal, healthy, mature young men —half were assigned at random to play the role of prisoners, half the role of prison guards.

The subjects began playing the roles assigned to them. They began, of course, by just going through the motions—literally, by deliberately activating the specific responses that characterize guard-like and prisoner-like acts. But in a very short time, their emotions and their spontaneous behaviors were recruited into the roles they played.

> It was no longer apparent to us or most of the subjects where they ended and their roles began. . . . We were horrified because we saw some boys (guards) treat other boys as if they were despicable animals, taking pleasure in cruelty, while other boys (prisoners) became servile, dehumanized robots who thought only of escape, of their own individual survival, and of their mounting hatred of the guards.[26]

The role-playing marathon had been planned to run 2 weeks. In fact, Zimbardo stopped the experiment early, for fear of his subjects' physical and mental safety. How long did this take? Just 6 days.

[25] Azrin and Hutchinson, 1967.
[26] Zimbardo, 1981, p. 131.

### A Look Backward: Learning and Instinct Revisited

It is clear from all this that if we set out to reinforce some arbitrary response, we may get more, or less, than we bargained for. Other, related responses may be strengthened besides the one we reinforce. And the one we select may not respond to reinforcement if a competing member of the system is strong. Pigeons have trouble *not pecking* to get food; pecking as a food-getting response is too strong.

In short, we cannot pick any response we like and expect to strengthen just that response and nothing else. To know what effect a reinforcer will have, we must know what responses are already connected to the one we reinforce. Such connections may be formed by genetic factors or by previous learning—the *principles of variability.*

And, to the extent that the connections are genetically determined—that is, are untaught, or instinctive—this means we must know what the structure of a species' instinctive repertoire is like. *It is not enough to know what its environment is like.* The impact of the environment will depend on the structure that is already there, and the study of reinforcement includes the study of the *interaction between learning and instinct.*

## THE COST OF REWARD

There is another and quite different way in which the use of reinforcers may give us more than we bargained for—or less. Earlier, we saw how reinforcement can increase the frequency of desired behavior in hospital, school, home, and industry. But now we must consider some other effects of such procedures. They may strengthen desired behavior—but what else do they do?

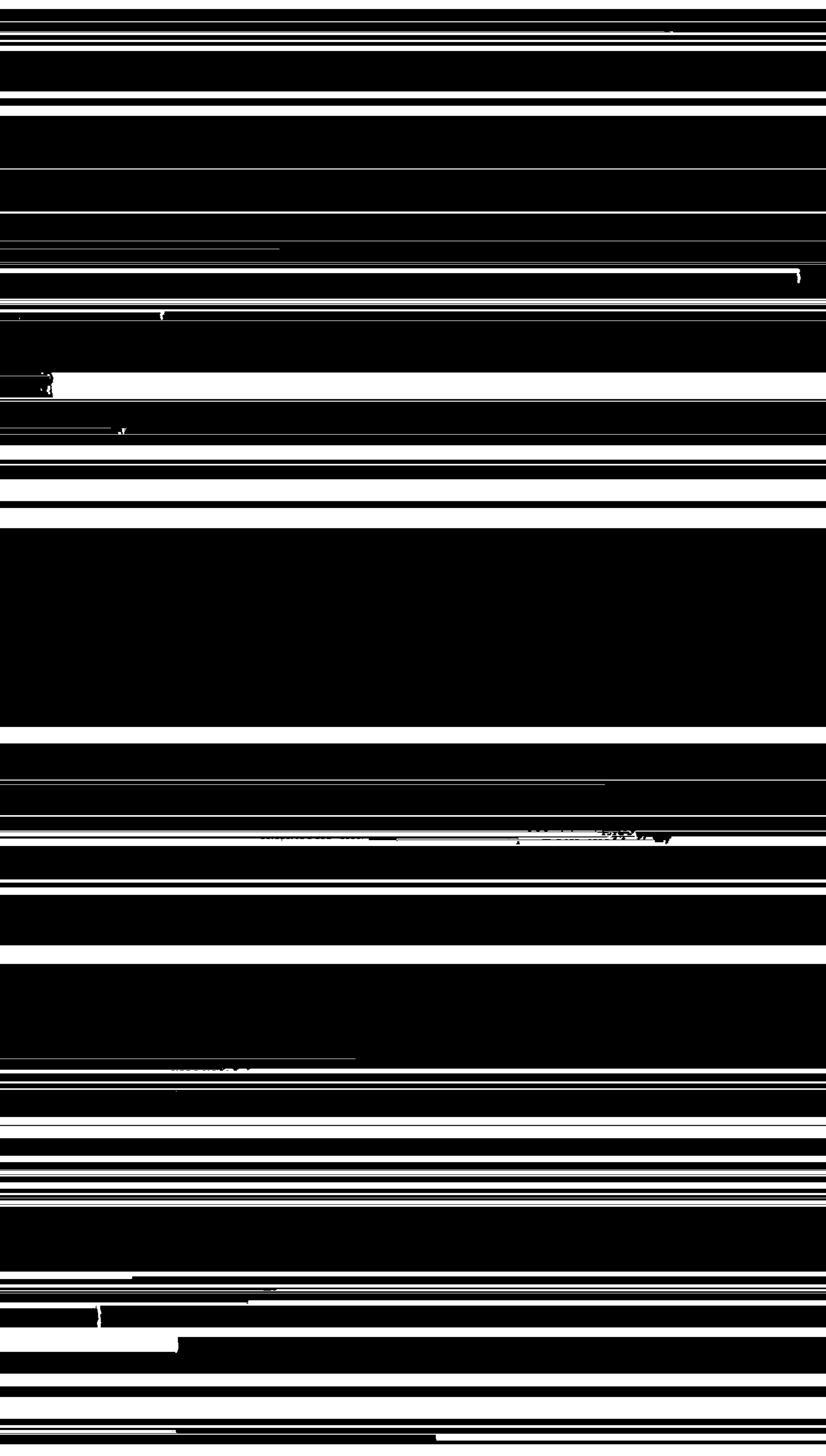
### Undermining Intrinsic Reinforcement

Consider this experiment.[27] Nursery school children were reinforced with tokens for drawing. The reinforcement was effective: The children worked vigorously at drawing while the reinforcement conditions were in effect. But when the tokens were no longer given, the children stopped drawing, and actually did *less* drawing, later, than children who had never been reinforced for doing so. The *intrinsic* rewards of drawing—that is, the satisfactions that the activity itself provides—no longer maintained the children's drawing activity.

Note that this finding is not a failure of reinforcement theory. The reinforcers did just what they ought to do; they increased the amount of drawing that occurred, *as long as they were delivered.* The problem was the weakening of **intrinsic reinforcement**—the rewards of drawing for its own sake.

Why should this happen? We are not sure. It might have to do with how

[27]Lepper, Greene, and Nisbett, 1973.



they could learn nothing new by doing so. Or they would return to hypotheses about what the rule was that they had already ruled out earlier. It appeared that working for reinforcement alone had actually switched off part of the subjects' rule-finding, hypothesis-testing apparatus—and that they did not find it easy to switch it on again!

Here again, the findings do not imply a failure of reinforcement theory. The reinforcers worked. If a sequence was reinforced, that sequence was likely to be repeated. The problem is that sequences were strengthened *at the expense* of cognitive operations leading to rule discovery. That may not be the outcome we want:

> [S]uppose that our educational goals include developing the ability to discover general rules, as scientists do, and fostering an interest in learning so that the rewards [of] acquiring knowledge are an inherent part of the learning process itself. The [rule-discovery experiment] suggests that operant contingencies are not an effective tool for developing rule discovery. Indeed, they might even interfere with the development of this ability. And the [intrinsic motivation] experiment . . . suggests that operant contingencies may take control away from the rewards inherent in an activity, rather than promoting them. Thus, for people who view these educational goals as significant, the lesson of these two experiments may be that operant contingencies can interfere with effective education rather than facilitate it.[30]

## The Reinforcement Theorist Replies

Now a reinforcement theorist would not accept that criticism. He might reply:

> Look, it is well known that weak reinforcers—intrinsic reinforcers, for instance—become weaker still when they are contrasted with more powerful ones.* Behavior therapists are well aware of this problem. Some have stated it as a rule: Never reinforce a response unless the natural environment will continue to reinforce it.
>
> As for rule-finding and other forms of comprehension, we know that it's folly to reinforce A while hoping for B [see above, pp. 275–276]. If we want subjects to find rules and principles, and not just settle on a response sequence that works, then we must identify and reinforce the behavior that leads to rule-finding.

Such a reply may make good theoretical sense, but the critic of reinforcement procedures is not likely to be satisfied with it. The problem is whether the theory can be implemented in practice. It is easy to say, "Don't reinforce a response unless the natural environment will reinforce

[30]Schwartz and Lacey, 1982, p. 250.
*See the discussion of the matching law, pp. 344–348.

it." That, however, may be more a wistful hope than a practical principle. In the classroom, reading and studying can be promoted by reinforcement.[31] But what happens when the child leaves school, if she returns to a home and a peer group in which reading and studying are not reinforced? Just how do we get that natural environment to reinforce reading, studying, and the like? And if we do not yet know just what the behaviors are that lead to comprehension rather than rote performance, maybe we should wait until we do know before we rush in with our reinforcement-based interventions.

### A Look Backward: The Risks of Reinforcement

These findings about the costs of reward do not pose a challenge to reinforcement *theory.* But they certainly do pose a challenge to its simple-minded *application.* And precisely because operant techniques have been used so successfully in schools, factories, and hospitals, some writers fear that we may adopt them with an enthusiasm that does not consider their dangers.

## CRITIQUE OF REINFORCEMENT THEORY

In the last section we looked at some dangers in the *application* of reinforcement. What about the reinforcement principle as a *theory* of purposive behavior?

As we saw at the beginning of this chapter, many behaviorist writers argue that purposive behavior *is* behavior controlled by reinforcement. The concept of reinforcement may replace the vague notion of *purpose* with a specific and observable process—that is, the effect of the organism's past experience on its present behavior.

The advantages of this view are considerable—if it will work. If we speak of seeking some future goal state, we seem to be implying that the future can in some mysterious way act back to affect the present, and we do not really believe that. If we reply that it is not the future but our present *hopes* or *wishes* for the future that affect our actions, we are speaking of internal agents that need to be explained in turn (Chapter 1).

The reinforcement theorist sees no need to talk about unobservable internal events. Rather than doing the things that we *hope will work* in the future, we do the things that *have worked* in the past. That is how the light of experience guides our feet. What more needs to be said?

Despite its appeal, however, the attempt to make reinforcement "all there is" to purposive behavior faces a difficulty. The problem is so simple, yet so basic, that it must stand as the Achilles' heel of the enterprise, despite the volumes that have been written in attempts to argue it away.

It is this: To be reinforced, a response must occur and be followed by the reinforcing event. To show that it has been reinforced, we must show

[31]See for example Hall, Lund, and Jackson, 1968.

that it is more likely to occur on a later occasion. In other words, if we are to speak of reinforcement at all, then *the response in question must occur at least twice, and the reinforcing event must occur at least once.*

But what do we say about an action the *first* time it occurs? And what of outcomes that do not occur at all? Let us look at an example that raises both these problems.

## First-Time Actions and Non-occurring Outcomes

To dramatize the situation, suppose that we live in a never-never-land called Titipu. Suppose further that we have just been informed of a decree by the omnipotent ruler of the land: Henceforth, flirting is a capital offense!

> *He decreed in words succinct*
> *That all who flirted, leered, or winked*
> *(unless connubially linked)*
> *Should forthwith be beheaded.*
>
> —W. S. GILBERT, *The Mikado*

Having heard that, would we flirt, leer, or wink? Probably not—certainly not without a careful look round for the secret police. The lethal *consequences* of our flirting will very likely suppress that behavior. The question is: Is this because of what the consequences have been in the past, or of what we expect them to be in the future?

Granted, expectations of the future—if they exist—must be based on building blocks that past experience provides. Experience has taught us what beheading is, and what flirting is, and, perhaps, whether the secret police in Titipu are competent enough to worry about. But the punishment at issue is not part of history. We have never been beheaded before. Our flirting behavior has never been punished before. And, quite possibly, the act of *refraining from flirting* is one we have never emitted before.

Rather than being a product of *past* reinforcement history, it seems clear that our new cautious demeanor is under the control of its anticipated *future* consequences. A new situation confronts us: We *will be* beheaded if we flirt. And so we don't. The concept of reinforcement, as defined, simply does not apply. So a cognitivist theorist would argue.

## Rule-following Behavior and Its Difficulties

The debate does not end there, however. A reinforcement theorist would object that what we do, in refraining from flirting, is to emit an operant that has in fact been reinforced in the past—and often. It is the operant of *following a rule.* Thus a behaviorist writes:

> [W]ith some choices . . . we cannot directly experience the consequences of the alternatives before making a decision. Taking a particular job, getting married, buying a house are all important choices, but

*purposive* the *first* time they occur, and/or that are affected by *conse-quences* that have not occurred. The consequences may never occur—a law-abiding, non–flirting population may never experience a single beheading in Titipu. Rather than stretch the meaning of the reinforcement concept to cover these troublesome cases, it seems better to accept the very real insights that reinforcement theory offers—but also its limitations, and the need for something more.

What that "something more" might look like, we will explore in the next chapter.

## SUMMARY

Reinforcement *procedures* are used to train animals to perform arbitrary acts. We use them to diagnose motivational states, as opposed to stimulus-bound reflexive or instinctive actions. Reinforcement *theory* holds that all purposive behavior is maintained by its reinforcing consequences. To call an act *purposive* is to say, not that future consequences are anticipated, but that such an act was established and maintained by reinforcement in the past. Internal states such as food deprivation affect purposive behavior because they set the reinforcing value of environmental events; for example, food deprivation makes food a reinforcer.

Modern reinforcement theory uses concepts developed by B. F. Skinner. It distinguishes between *respondents,* which are elicited by stimuli, and *operants,* which are emitted by the organism without stimuli that cause them. If an event increases the strength of a response that it follows, then that event is defined as a reinforcer. Positive reinforcers are ones whose onset is reinforcing. With negative reinforcers, the offset of the event is reinforcing; termination of painful shock is an example. If a response is followed by the onset of a negative reinforcer, we speak of *punishment.* These events are defined by their effect on a response; it is responses, not organisms, that are reinforced. If a response is reinforced in the presence of a stimulus, and not reinforced in its absence, then it will come to occur only when that stimulus is present. We then call the stimulus a *discriminative stimulus,* and say that it has acquired *stimulus control* over the emission of the operant.

Operant behavior is frequently studied in a chamber (operant chamber or Skinner box) with some defined response available, such as pressing a lever for a rat or pecking at a key for a pigeon. Reinforcements can be delivered, and rate of response measured, without change in the situation and without interrupting the continuity of behavior.

Reinforcement procedures have been applied in a variety of settings—hospitals, homes, schools, workplaces—to increase the frequency of effective behavior. Often *conditioned reinforcers,* such as money or tokens, can be used where natural reinforcers would be difficult or inconvenient to use. These methods reflect a definite point of view—an operant philosophy—about the nature and causes of behavior. They center on identifying reinforcing conditions, identifying the behavior that is desired and/or the behavior that is problematic, and arranging that the one be reinforced and the other not. Thus they focus on the *environmental* determinants of behavior. Since other people are an important part of the environment, that can mean that one must try to change the behavior of the target person's family or caretakers, not just his own. The attempt may encounter great resistance.

What makes a reinforcer reinforcing? The hedonic hypothesis—rein-

forcers are pleasant—is hard to test. The drive-reduction hypothesis may cover some cases of reinforcement, but not all. The response-deprivation hypothesis says that opportunity to make a response, when such opportunity is otherwise restricted, is reinforcing. That idea explains the fact that Response A may reinforce Response B under some conditions, but B may reinforce A under others. But not all reinforcers depend on response restriction in any obvious way, nor is it always clear what responses are restricted and how much. We are left with the *empirical law of effect:* An event is a reinforcer if it reinforces. Perhaps different reinforcers have that effect for different reasons.

What exactly do reinforcers do? In some settings, certain *principles of variation*—genetic factors and prior learning—make only certain responses likely to occur. Reinforcement then acts as a *principle of selection.* It selects the responses that are reinforced, not because it makes them stronger but because non-reinforced responses get weaker and cease to occur. In other situations, however, reinforcement may strengthen responses directly. When any of a number of responses could be reinforced, animals and humans are likely to settle on only one, and keep repeating it. This suggests that that response is made stronger each time it is reinforced. If so, then reinforcement can strengthen responses directly as well as protect them from weakening.

For a long time it was assumed that reinforcement conditions can take hold of any response, at any level, and cause just that response to become more frequent. More recently we have realized that reinforcement taps into the hierarchical structure of behavior systems. Food reinforcement, for instance, brings an animal's repertoire of food-getting behaviors into play. Unwanted components of that system may occur, even at the expense of the reinforced response. Autoshaping is an especially vivid example: In a situation where food is predicted, pigeons peck, even if it costs them reinforcements (food). Besides these top-down effects, there may be bottom-up effects: Strengthening one component of a hierarchy can recruit the entire system, with its other components. In pigeons and perhaps in people, if one goes through the motions of behaving aggressively, aggressive motivational states may follow.

There are other ways in which reinforcement procedures may have unintended consequences. Powerful reinforcers may undermine the intrinsic motivation for an activity—the reward of performing the activity for its own sake. They may encourage the rote performance of successful actions as opposed to the discovery of principles, rules, and concepts. Both these effects are compatible with the theory of reinforcement; still, they show that the use of reinforcement has its dangers.

Does reinforcement account for all purposive behavior? Many writers think so, but others, myself included, do not. A response can be affected by its consequences the *first* time it occurs, when it has had no consequences before. And it may be affected by potential consequences that never occur at all. We must have a way of anticipating future events, as well as being affected by past ones.