

# Conditioned Reinforcement

Matthew C. Bell and Margaret A. McDevitt

## Introduction

The idea that a neutral stimulus can function as a learned (i.e., conditioned) reinforcer has a long and venerated place in learning theory. These conditioned reinforcers are sometimes said to have acquired conditioned *value*, a term we use to indicate that the stimulus has acquired the ability to function as a reinforcer for behavior. One of the first textbook accounts of conditioned reinforcement appeared in Keller and Schoenfeld (1950):

In reaching our goal [of understanding human behavior], the principle of secondary [conditioned] reinforcement will be of great analytical assistance. When added to the other functions of stimuli, it gives us a powerful and indispensable tool for the solution of many vexing and absorbing problems of human actions. (Keller & Schoenfeld, 1950, p. 260)

Skinner also recognized the importance in his seminal textbook,

Although it is characteristic of human behavior that primary reinforcers may be effective after a long delay, this is presumably only because intervening events become conditioned reinforcers. (Skinner, 1953, p. 76)

In general terms, learning processes have evolved in humans and other animals because of the survival advantage they offer. Although there is some cost to an organism (and a species) to have the capacity to learn, learning provides virtually unlimited behavioral flexibility. Imagine what would happen if most of our behaviors were genetically determined. Specific behavior patterns might be very useful in one environment, but what works in one environment might easily prove to be a hindrance

in another. If you consider how drastically human life has changed in the last few hundred years, you quickly realize the importance of learning. Without the capacity to learn, we would not be able to adapt quickly to a rapidly changing environment. The ability to learn is one of the most important contributions of evolution.

Conditioned reinforcers help navigate and guide our behavior through a complex world. Food and other primary reinforcers are crucial for the survival of humans and other animals, and learning about how they function is quite important. In order to respond effectively, though, it is equally important to learn about the myriad of other events and stimuli that provide information about those primary reinforcers. They have survival value as well. In fact, as Skinner (1953) pointed out above, most human behavior is guided by conditioned reinforcement. Most of what motivates our behavior is not direct access to primary reinforcers like food and water, but rather things that lead to those primary reinforcers.

### **Defining Characteristics of Conditioned Reinforcement**

Conditioned reinforcers are previously neutral stimuli that gain reinforcing effectiveness either by their association with primary reinforcers or by association with already-established conditioned reinforcers. For example, the sound of a food hopper (a grain delivery mechanism) being raised may become a conditioned reinforcer for a pigeon that has been working for food in an experimental chamber. The sound of the hopper being raised immediately precedes access to and consumption of grain, and therefore over time that sound itself becomes a reinforcing stimulus. Similarly, the sound of our friend's voice may become a conditioned reinforcer over time, as the presence of their voice often signals the reinforcing aspects of the friendship.

Conditioned reinforcers are traditionally assumed to have one essential characteristic: When presented as a consequence immediately following a response, the conditioned reinforcer functions to increase the probability of future responding. This function is thought to be due to its history of association with already-established reinforcers.

Zimmerman, Hanford, and Brown (1967) provide one example of an early experiment designed to demonstrate conditioned reinforcement effects. Pigeons were presented with a two-component multiple schedule (where two or more simple schedules are arranged in sequence; see, for example, Ferster & Skinner, 1957). In both components food delivery was not contingent upon responding and any food presentation that was scheduled required 6s without a key peck. Response keys were illuminated blue and yellow, respectively, for the two components. Responding to the illuminated keys resulted in brief presentations of the food magazine sound (the conditioned reinforcer). The first component delivered reinforcement according to variable interval (VI) schedules ranging from a VI 1-min to a VI 12-min across conditions in addition to two extinction conditions (key pecks never produced the conditioned reinforcer). A VI schedule provides reinforcement for a single response following a varying period of time across opportunities. The second component delivered food according to a VI 1-min schedule as well. Responding was well maintained over a substantial period of time (the study lasted for over 100 sessions) even though pecking at the keys was never reinforced with food.

Zimmerman et al.'s (1967) findings seem to necessitate an explanation in terms of conditioned reinforcement; the concept's use as an explanatory concept has been more controversial for other behavioral procedures, some of which we will describe below. It is important to clarify when conditioned reinforcement is a necessary and/or sufficient explanation for a behavioral finding and when other explanatory frameworks are required. For example, many of the findings attributed to conditioned reinforcement have alternatively been interpreted in terms of discriminative functions. We do not advocate conditioned reinforcement as a one-stop catch-all explanatory mechanism. Multiple behavioral procedures are involved in most behavioral phenomena, and it is a mistake to view them as mutually exclusive.

Zimmerman et al.'s (1967) procedure is but one example of several procedures used to investigate conditioned reinforcement. In the next section, we describe a number of different procedures that have been used to study conditioned reinforcement. As we will see, there are a number of challenges in studying pure conditioned reinforcement effects. Although there is no single definitive procedure for studying conditioned reinforcement effects, the data from a rich variety of procedures provide considerable evidence of the nature of the effect and the importance of the concept.

In addition to reinforcing properties, stimuli may also have discriminative properties. In other words, a stimulus may signal a particular operant contingency and, functioning as a discriminative stimulus ( $S^D$ ), it may lead to differential responding independent of any reinforcing properties. Distinguishing between the discriminative and reinforcing functions of stimuli has proven difficult. Early theorists suggested that in order for a neutral stimulus to become a conditioned reinforcer that it would have to become a discriminative stimulus for an operant contingency (see, for example, Schoenfeld, Antonitis, & Bersh, 1950; Skinner, 1938; Wyckoff, 1952), although later work suggests this may not be a necessary condition (see, Jenkins & Boakes, 1973; Kendall, 1973a, 1973b). The interplay between conditioned reinforcing and discriminative properties of stimuli is an important topic, but it is largely beyond the scope of this chapter.

## Studying Conditioned Reinforcement

Several procedures have been used to study conditioned reinforcement. In this section, we describe a variety of procedures to illustrate both some historical approaches and to show the challenges presented in the study of conditioned reinforcement. While this list is not exhaustive, it does cover the major approaches.

### Train a New Response

One of the earliest approaches to studying conditioned reinforcement effects focused on demonstrating the response-strengthening aspect of conditioned reinforcement. In an initial condition, a neutral stimulus is repeatedly paired with a primary reinforcer and then, in extinction, that previously neutral stimulus is tested to see if it functions as a reinforcer for a new response. Skinner (1938) provided one of the earliest demonstrations of this approach: he trained rats to approach a food cup when the

mechanism operated. (Its operation produced a sound which was used as the conditioned reinforcer.) Following 60 sound-food presentations, he stopped presenting food in the situation and introduced a lever. Responding on the lever produced the sound. Skinner successfully shaped bar pressing and reported increased response rates over a brief period. Responding then decreased, presumably as the conditioned reinforcing value of the sound extinguished.

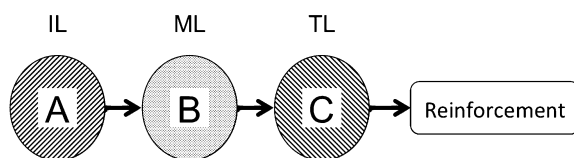
Another early example with a more applied emphasis was reported by Wolfe (1936) who compared the effectiveness of primary and conditioned reinforcement in establishing and maintaining the behavior of chimpanzees. While there was some decrement associated with the use of conditioned reinforcement (tokens), the conditioned reinforcement was generally similar to food in its effectiveness.<sup>1</sup>

Although it provided some interesting and useful data, this general procedure proved to be problematic for two reasons. First, it is an extinction procedure. Any existing conditioned value diminishes during testing because the stimulus functioning as a conditioned reinforcer is no longer paired with primary reinforcement, and therefore it is no longer paired with the source of conditioned value. Second, the change from the baseline conditions to the conditioned reinforcement tests produced stimulus generalization decrement effects. This is because removal of the primary reinforcer made the testing condition noticeably different from the training condition. In other words, responding might decrease not merely because of any change in conditioned value per se, but rather because the change of conditions can be discriminated. Regardless of which of these factors might be implicated, effects found using this paradigm were typically fleeting and thus this approach proved to be quite limited.

### Extinction Rates for an Established Response

A similar approach is to maintain a response with both a conditioned reinforcer and primary reinforcer. The primary reinforcer is then removed (extinction) and responding is compared across conditions with or without the putative conditioned reinforcer present. Higher, more persistent responding in conditions where the stimulus followed a response would suggest that the stimulus functioned as a conditioned reinforcer.

An early example of this approach was reported by Zimmerman (1957) using water-deprived rats. In the training phase, a 2-s buzzer was followed immediately by presentations of water using a dipper. Trials occurred approximately once a minute. Zimmerman began with a 1:1 buzzer-water pairing and thinned out the schedule until buzzer presentations were followed by water on approximately 10% of the trials. Then, in the test phase, Zimmerman tested whether or not the buzzer would function as a conditioned reinforcer by using it as a consequence for lever pressing (according to an intermittent reinforcement schedule). Importantly, water was never presented during these sessions. Zimmerman found that responding was maintained by the conditioned reinforcing properties of the buzzer. Zimmerman also reported that lever pressing extinguished when the buzzer was removed, providing additional evidence that the buzzer functioned as a conditioned reinforcer (but see Wyckoff, Sidowski, & Chambliss, 1958, for an alternative interpretation).



**Figure 10.1** Procedure schematic of a typical three-link chain schedule used to study conditioned reinforcement effects. Each link of the chain is correlated with a specific schedule of reinforcement. When the schedule requirement is satisfied, responding to the initial link (IL) stimulus (A) leads to the middle link (ML) stimulus (B) which then leads to the terminal link (TL) stimulus (C) and finally primary reinforcement.

### Chain Schedules

A chain schedule is a fixed sequence of stimuli correlated with different reinforcement schedules where the final link of the chain ends in primary reinforcement (see Figure 10.1). Each link in the chain is presented successively and operates according to its own independent schedule of reinforcement. Imagine, for example, that a pigeon is presented with a three-link chain schedule, in which green, yellow, and red colored keylights are presented in sequence. Food is presented following completion of the schedule requirement correlated with the red stimulus, but to get to the red stimulus, our pigeon must respond to first the green, and then the yellow stimulus, completing the schedule requirements in each of those links. What reinforces responding to the green and yellow stimuli?

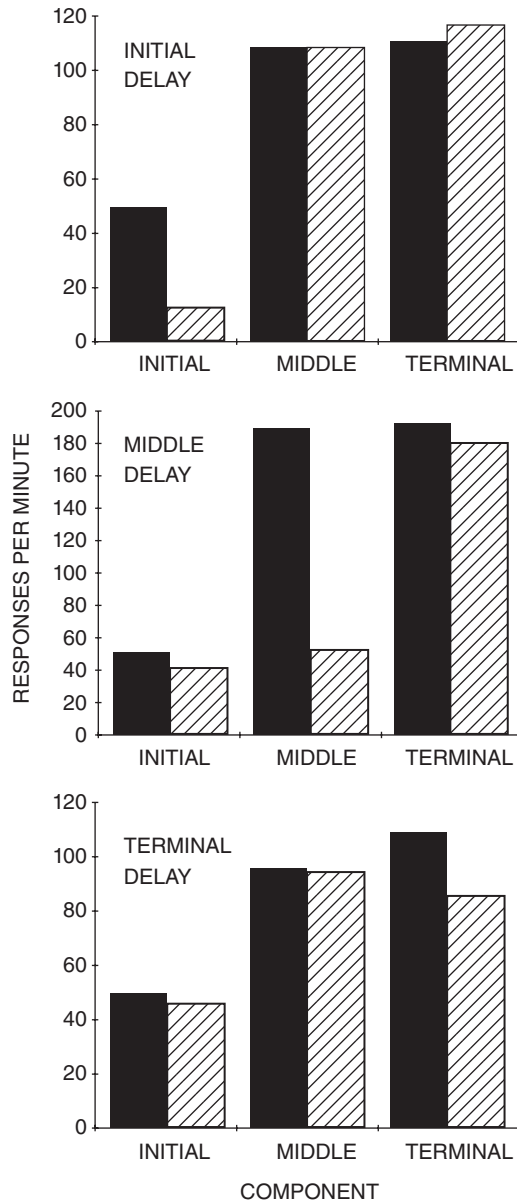
One possibility could be the delayed primary reinforcer at the end of the chain. However, depending on the schedule values used in the chain schedule, the overall delay to food might be so long that responding may not be adequately maintained by the food at the end of the chain. Another possibility is that responding to green is reinforced by the presentation of yellow, and responding to yellow is reinforced by the presentation of red. In this way, conditioned reinforcement value is traditionally believed to extend backwards through the chain. Because red is directly associated with food, it quickly gains value as a conditioned reinforcer. As red gains in reinforcing strength, yellow begins to gain value as well, through its association with red (and so on, back through the chain). The stimulus at the end of the chain is always the most valuable conditioned reinforcer (due to its direct association with the primary reinforcer), and reinforcing effectiveness decreases with each prior link of the chain, as you get further away from the primary reinforcer at the end (see, for example, Williams, Ploog, & Bell, 1995; Williams, 1997). The stimuli in the chain appear to have acquired conditioned value and function as conditioned reinforcers for responding in earlier links of the chain.

This approach avoids the problems associated with testing during extinction by examining steady-state behavior maintained by primary reinforcement at the end of the chain. Conditioned reinforcement effects are seen by examining performance in links of the chain not directly paired with primary reinforcement.

Chain schedule performance has been compared to a variety of control conditions, often tandem schedules. Tandem schedules differ from chain schedules in that there are no stimulus changes between links; they are otherwise functionally identical to chain schedules. Control of responding in early links of a chain schedule was typically thought to be under direct control of the immediate consequences. In other words, the conditioned reinforcement value of the following stimulus functions to reinforce responding in the preceding link. Eliminating the stimulus transition should eliminate conditioned reinforcement and, therefore, lower the response rate.

An early failure to show conditioned reinforcement effects in the chain schedule demonstrates the difficulties encountered in studying these effects. Gollub (1958, as described in Gollub, 1977 and Fantino & Logan, 1979) compared responding across three groups. He used fixed-interval (FI; for more, see Ferster & Skinner, 1957) schedules of reinforcement, in which only the first response that occurred after a fixed amount of time resulted in delivery of the reinforcer. In the first group, pigeons were presented with a simple FI 5-min schedule. The second group of pigeons was presented with a chain schedule with five FI 1-min links. The third group of pigeons was presented a tandem schedule with five FI 1-min links. Responding on the simple and tandem schedules were comparable, but responding on the chain schedule was weakly maintained, if at all. The problem with this particular procedure, and a challenge associated with chain schedules more generally, is that stimuli in early links of a chain function as signals for nonreinforcement. Thus, responding in earlier links of a chain will typically be slower than responding in later links. Responding is difficult to maintain in the early links of chains with three or more links.

Despite the failure of Gollub's (1958) study to demonstrate support for conditioned reinforcement within a chain schedule, later research which employed shorter chains (i.e., fewer than five links) and VI schedules (instead of FI) have successfully investigated the conditioned reinforcement relationship between links in a chain schedule. For example, Royalty, Williams, and Fantino (1987) systematically included an unsignaled delay of reinforcement in different links of a three-link chain schedule. In the typical chain schedule, the response that satisfies the schedule requirement is immediately followed by a stimulus change. When an unsignaled delay is included, that same response is followed by a brief delay prior to the stimulus change (note that additional responses have no effect). Royalty et al. found evidence that unsignaled delays can have substantial effects on responding in the link associated with the unsignaled delay. In their study, pigeons responded to a three-link chain where, in baseline conditions, each link was a VI 33-s schedule. They systematically inserted a 3-s unsignaled delay between the response and the consequence—in other words, the VI 33-s schedule was changed to a tandem VI 30-s fixed-time (FT, an FT schedule delivers reinforcers independently of responding after a fixed amount of time) 3-s schedule—for each of the links. One important feature of this design is that the overall time to primary reinforcement was held constant. That is, the relationship between each stimulus and primary reinforcement remained unchanged. Their results showed lower responding in only the link with the unsignaled delay (Figure 10.2) caused by the degraded contingency between the response and the conditioned reinforcer in the following link of the chain (and not by any changes in timing of stimulus changes or in time to reinforcement). The results of this study suggest that responding in chain schedules is, in part, maintained by conditioned reinforcement.



**Figure 10.2** Mean response rate (response/min) for each chain component. The dark bars represent baseline responding. Hashed bars represent responding during the various delay conditions, with the initial-link delay condition shown in the top graph, middle-link delay condition shown in the middle graph, and terminal-link delay condition shown in the bottom graph. Ordinate scaling varies among panels. (Royalty, Williams, and Fantino, 1987, Figure 7. Reproduced with permission from Wiley).

### Observing Responses

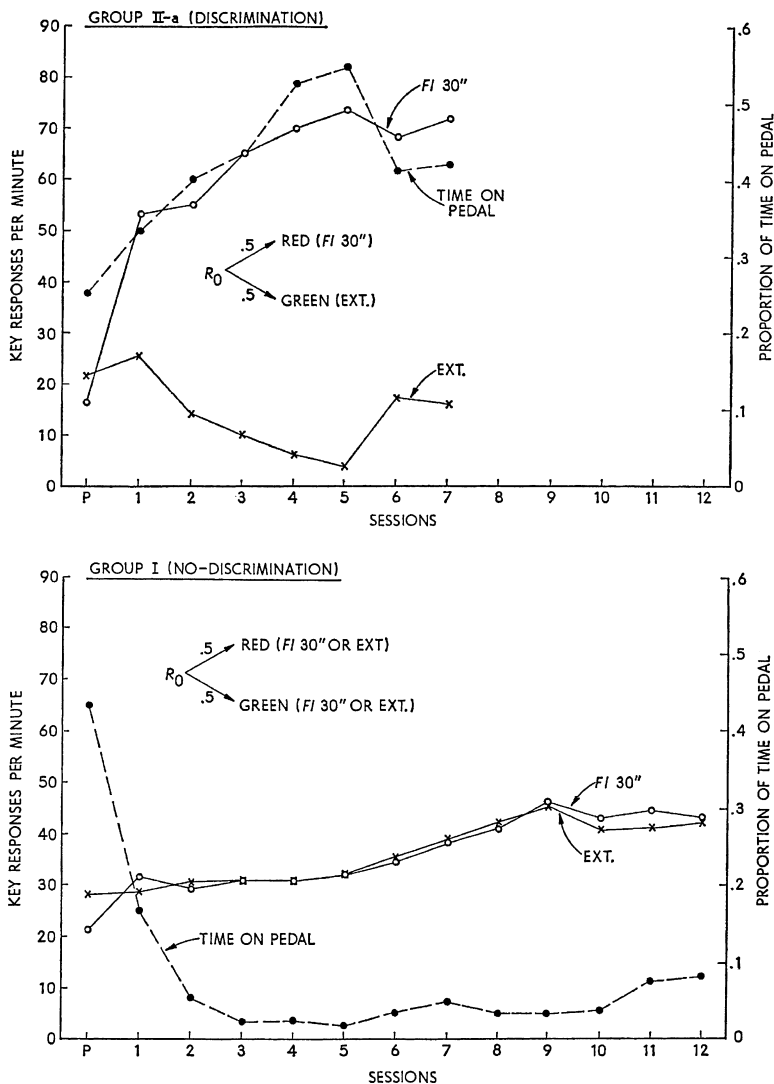
The observing response procedure takes a different approach. This procedure assesses the degree to which a conditioned reinforcer has value (recall that this term is used to indicate the ability of a reinforcer to maintain behavior). Wyckoff (1952, 1969) presented pigeons with two schedules of reinforcement, an FI 30-s schedule of food reinforcement and extinction. These schedules alternated within experimental sessions. Initially, the schedules were correlated with the same white key light stimulus (a mixed schedule). A separate observing response (a pedal press) changed the key light stimulus from white to red if food was available or from white to green if extinction was in effect. Thus, observing responses converted the mixed schedule to a multiple schedule, but had no other consequence. In control conditions, the stimuli were not correlated with reinforcement (i.e., red and green were equally associated with the FI 30-s schedule and the extinction schedule). When the stimuli were correlated with reinforcement, subjects reliably responded to the pedal and to the red stimulus and responded at a much lower rate to the green stimulus. When the stimuli were uncorrelated, the observing rate was low and subjects responded equally to the two schedules. This study shows that subjects will respond for the signal indicating which reinforcement schedule is in effect when the signals are correlated with the reinforcement and extinction schedules (top graph in Figure 10.3) and will respond differentially to the two schedules. When the signals are uncorrelated, subjects make very few observing responses and responding to the two schedules is not different (bottom graph in Figure 10.3). Thus, when the signals are correlated, it appears that the signal for the reinforcement schedule functions as a conditioned reinforcer. One alternative interpretation of these findings is that subjects are responding because of the information value of the signal, independent from any association with reinforcement. However, the information hypothesis was not supported by subsequent research (e.g., Fantino 1977).

### Concurrent Chains

The concurrent chains procedure (Figure 10.4) was developed to address the challenges of discriminating between schedule-related effects (like the scallop pattern of responding often seen in FI schedules) and conditioned reinforcement effects (Fantino, 1977). That is, in simple choice situations there is a confound between any putative conditioned reinforcement effects and schedule demands. For example, if subjects were presented with a choice between a differential reinforcement of high rate schedule (DRH; this is a schedule where the response rate must be higher than a pre-defined response rate threshold for reinforcement to be delivered), and a differential reinforcement of low rate schedule (DRL; a schedule where the response rate must be lower than a pre-defined response rate threshold for reinforcement to be delivered), it would appear as if the DRH were highly preferred. Clearly, the high rate of responding in the DRH is not necessarily an indication of preference for that schedule.

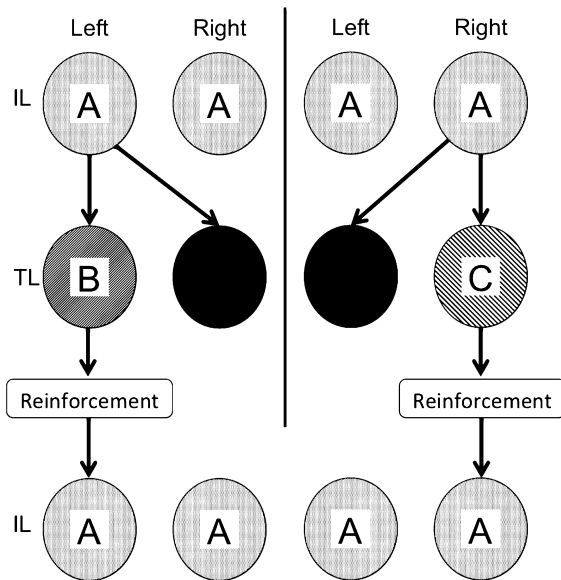
The concurrent chains procedure addressed this problem by presenting subjects with a choice between alternatives with two equal initial links, typically equal VI schedules available concurrently. Any differences seen in initial-link responding are then attributable to differences in what follows those choice alternatives.<sup>2</sup>





**Figure 10.3** The top graph shows mean responding when pedal presses produced signals (red and green stimuli) correlated with the fixed-interval (FI) 30-s and extinction (EXT) schedules. The bottom graph shows mean responding when the pedal presses produced signals uncorrelated with the two schedules. When the signals were correlated, subjects made observing responses and responded differentially to the two schedules. When the signals were uncorrelated, subjects made very few observing responses and did not respond differentially to the two schedules (reprinted from Wyckoff, 1969, Figures 9.1 [bottom panel] and 9.2 [top panel]).

A counterintuitive example of how conditioned reinforcement effects are tied to the schedule of primary reinforcement is illustrated by studies that evaluated preference between a reliable alternative (one that leads to food 100% of the time) and an unreliable alternative (one that leads to food 50% of the time and blackout 50% of the time; Belke & Spetch, 1994; Dunn & Spetch, 1990; Kendall, 1974; McDavitt,



**Figure 10.4** Procedure schematic of a typical concurrent chains schedule used to study conditioned reinforcement effects. When subjects complete the schedule requirement on the left initial link (IL) (shown in the left panel), the IL stimuli extinguish and the left terminal-link (TL) schedule begins. Completing the terminal-link schedule results in reinforcement. The right side functions in a similar fashion. Typically IL schedules consist of equal VI schedules. TL schedules vary by experiment. Since IL schedules are equal, any differences in performance to those is presumed to be a function of what follows those links.

Spetch, & Dunn, 1997; Spetch, Belke, Barnet, Dunn, & Pierce, 1990). In a replication of the basic finding, McDevitt et al. showed that pigeons were largely indifferent between a 100% reinforcement alternative and a 50% alternative when the food and blackout outcomes were signaled by the terminal-link stimulus on the 50% alternative (i.e., when one stimulus was correlated with trials ending in food and a different stimulus was correlated with trials ending in blackout). It was hypothesized that the signal for food on the 50% alternative served as a stronger conditioned reinforcer than the signal for food on the 100% alternative, since subjects should have exclusively preferred the 100% alternative because it provided twice as much food. In a series of conditions, McDevitt et al. assessed the strength of the terminal-link stimuli as conditioned reinforcers by interposing a short gap (a 5-s period in which the keylights were darkened) between the choice response and some or all of the terminal-link stimuli. If a terminal-link stimulus functioned as a conditioned reinforcer, the imposition of the gap should reduce its effectiveness, and choice behavior should shift away from the alternative with the gap. The results supported the notion that conditioned reinforcement underlies the finding of suboptimal preference in probabilistic (100% vs. 50%) reinforcement procedures (i.e., when the gap was placed at the beginning of the signal for food on the 50% alternative, choice shifted more than when the gap was placed at the beginning of the signal for food on the 100% alternative). Thus,

the evidence indicated that the signal for food on the 50% reinforcement alternative functioned as a stronger conditioned reinforcer than the signal for food on the 100% reinforcement alternative.

McDevitt and Williams (2010) used a concurrent chains procedure to provide evidence that both the frequency and the value of conditioned reinforcers influence choice in a concurrent chains procedure. In their study, pigeons responded on a concurrent chain where one choice option presented a VI 30-s VI 60-s chain and the other option presented a VI 60-s VI 30-s chain. Again, only the initial links were concurrent. This manipulation equated total time to reinforcement for the two alternatives. Importantly, McDevitt and Williams manipulated whether the terminal links of the two chains had different or identical signals. In conditions with different signals, subjects preferred the chain with the shorter terminal link, a finding consistent with other research (e.g., Duncan & Fantino, 1970). When identical stimuli were used for the terminal links, preference reversed in favor of the alternative with the shorter initial link (i.e., the VI 30-s VI 60-s). They concluded that the reversal in preference occurred because the shorter initial link produced the common terminal-link stimulus more frequently. In other words, the conditioned reinforcer occurred more frequently.

### Discrimination Learning

Although most studies of conditioned reinforcement have used free-operant procedures (which focus on response rate as the dependent variable), Williams and Dunn (1991b) examined the effects of conditioned reinforcement on the acquisition of a discrimination task. Rats were presented with trials in which one of two  $S^D$ s, a light or a noise, was presented. Responding to one lever was reinforced in the presence of the light, and responding to the other lever was reinforced in the presence of noise. After a learning criterion was reached, the contingencies for the  $S^D$ s were reversed. In baseline conditions, food always followed correct responses. In other conditions, food followed correct responses only 50% or 30% of the time. The critical manipulation involved presenting a conditioned reinforcer on the unreinforced correct trials. When the conditioned reinforcer was presented on those trials, learning occurred more quickly. When the conditioned reinforcer was not presented, learning occurred more slowly. Williams and Dunn's results show that conditioned reinforcers could substitute for primary reinforcers in a discrimination learning task.

Despite the weaknesses inherent in individual approaches, taken together, these findings offer strong support for the notion that conditioned reinforcers are just that, stimuli that acquire reinforcement strength of their own. Next, we consider the determinants of conditioned reinforcement.

### Determinants of Conditioned Value

The exact quantitative relationship that best describes the critical factors determining conditioned reinforcement is still a matter of some debate (see Preston & Fantino, 1991, for a review). Fantino's (1977; Fantino, Preston, & Dunn, 1993) *delay-reduction theory*, for example, proposes that the value of a conditioned reinforcer is

a function of the reduction in the waiting time to reinforcement signaled by a stimulus. More specifically, delay-reduction theory posits that choice responding in the initial link of a concurrent chains schedule is determined by the relative improvement in time to reinforcement signaled by the terminal-link onsets (Fantino, 1969), as shown in Equation 1.  $B_L$  and  $B_R$  represent the number of responses to the left and right initial links,  $T$  represents the average time to primary reinforcement from the onset of the initial links, and  $t_L$  and  $t_R$  represent the average durations of the left and right terminal links.

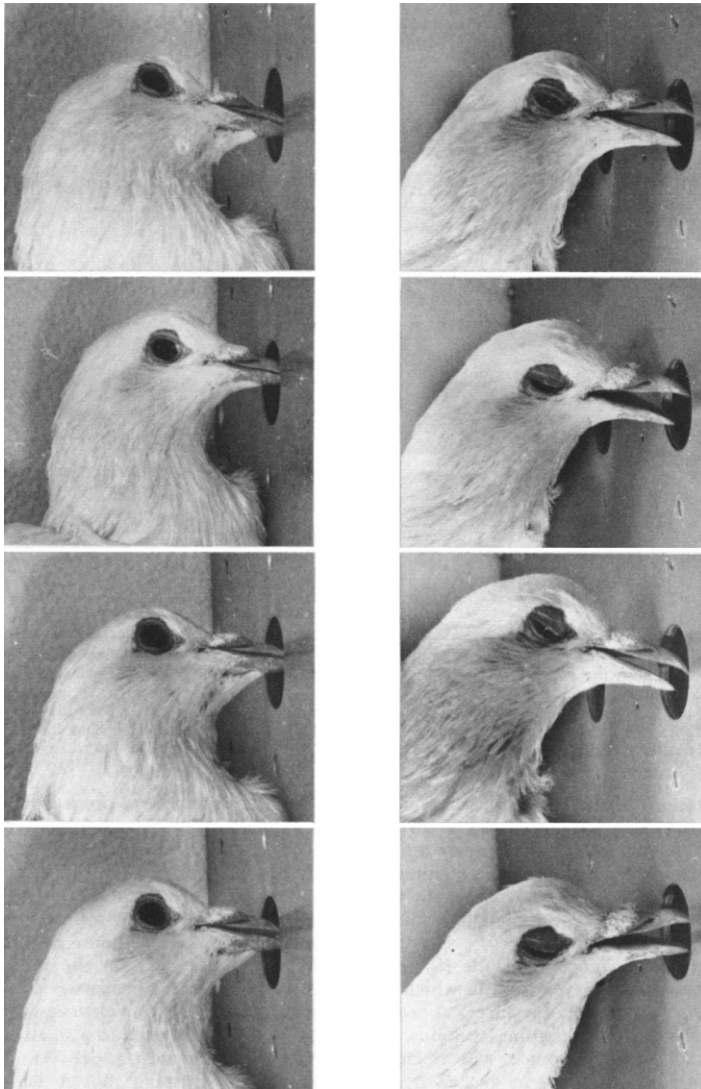
$$\frac{B_L}{B_R} = \frac{T - t_L}{T - t_R} \quad (1)$$

Others have emphasized the information value of the conditioned reinforcer—in other words, the degree to which a stimulus signals availability of primary reinforcement and reduces the uncertainty about the availability of reinforcement (referred to as either the *information hypothesis* or the *uncertainty-reduction hypothesis*; Berlyne, 1960; Bloomfield, 1972; Rachlin, 1976). Fantino (1977) provided a summary of the literature focused on discriminating between these accounts, and suggested that delay-reduction theory, and thus a conditioned reinforcement account, better accounts for the data.

Autoshaping (sometimes referred to as sign-tracking) procedures (Brown & Jenkins, 1968) provide some evidence that neutral stimuli paired with food and water acquire value. In the typical autoshaping procedure, naïve pigeons are trained to eat from a food hopper. They are then presented with an illuminated response key for 10 s followed by 4-s access to food which is then followed by a 60-s intertrial interval. In this procedure, the illuminated response key functions as a conditioned stimulus (CS) and the food as an unconditioned stimulus (US). Even though there is no requirement that the pigeons respond, most will come to peck at the illuminated response key.

Jenkins and Moore (1973) provided compelling evidence that the CS acquires some of the properties of the associated US. Across a series of experiments, they exposed pigeons to an autoshaping procedure. The critical manipulation was the type of US employed. In some cases the US was food and in others it was water. The difference in the US is important because of the different beak positions: when pigeons eat, they do so with an open gape and when they drink it is with a closed gape. Jenkins and Moore reported that the pecking topography toward the key (the CR) differed as a function of the US used. When food was used as the US, the conditioned pecks were open-gaped. When the US was water, the conditioned pecks were close-gaped (see Figure 10.5). This finding suggests that some of the properties of the US have transferred to the CS as indicated by the type of conditioned pecks that occur. It is important to note that a CR need not be similar to a UR as shown by these results. In fact, in many cases of Pavlovian conditioning the UR and CR differ dramatically.

Autoshaping suggests that conditioned stimuli acquire some of the properties of the unconditioned stimuli that they are paired with as evidenced by pigeons pecking at keys and doing so in a very specific way even though there is no contingency for doing so. If stimuli functioning as conditioned reinforcers acquire their function



**Figure 10.5** Typical food and water autoshaped responses for pigeons. Photos were taken at the moment of key contact. The left column shows responses to the water-paired key. The right column shows responses to the food-paired key. Reprinted from Jenkins and Moore (1973, Figure 2).

through Pavlovian conditioning process, it follows that the development of conditioned reinforcers should conform to the rules of Pavlovian conditioning. Furthermore, it suggests that the properties of those conditioned stimuli could be similar to the properties of the associated primary reinforcers.

### Behavioral Momentum Theory

One example of a theory that explicitly connects conditioned reinforcement value to Pavlovian contingencies is *behavioral momentum theory* (Nevin, 1992; Nevin &

Grace, 2000). Behavioral momentum theory describes behavior both under steady-state conditions and, more importantly, under conditions in which behavior is disrupted. The theory uses a metaphor of momentum from physics which says that responding in a discriminated operant is a function of two separate factors analogous to the factors that control the velocity of a body in motion. The first factor, resistance to change (persistence) under conditions of disruption, is likened to inertial mass in classical mechanics of physics. Resistance to change is said to be determined by the rate of food in the presence of a stimulus—in other words, by the Pavlovian stimulus-reinforcer relationship—and independent from any operant response-reinforcer contingencies. The second factor is the rate of occurrence of the operant response, likened to velocity of an object in motion and is well-described by the matching law (e.g., Herrnstein, 1961).

Nevin, Tota, Torquato, and Shull (1990) provide evidence for the role of Pavlovian contingencies in determining stimulus value. Their Experiment 1 presented pigeons with a multiple schedule in which two 3-min components alternated during experimental sessions. In both components (of most conditions) reinforcement was available according to equal VI 1-min schedules. Components were separated by a 1-min inter-component interval (ICI). The critical manipulation was the addition of free food (using a VT schedule) added to one of the VI 1-min components. This manipulation had the effect of both disrupting the operant (response-reinforcer) contingency and increasing the Pavlovian (stimulus-reinforcer) relationship. The increase in the Pavlovian relationship, according to behavioral momentum theory, should increase resistance to change, which is what they reported when they disrupted behavior using prefeeding and extinction manipulations. The findings are consistent with the general claim that Pavlovian contingencies function to determine the conditioned reinforcing properties of a stimulus (although note that other interpretations of these results are possible).

The Pavlovian conditioning and resistance to change relationship described in behavioral momentum theory is not the final word on the matter. Thus, while Nevin's theory attributes differences in resistance to change to Pavlovian conditioning, there is research challenging this conclusion (see, for example, Grace, Schwendiman, & Nevin, 1998 and Bell, 1999). Although beyond the scope of this chapter, alternative interpretations should be taken into consideration when considering the viability of using Pavlovian conditioning to explain resistance to change.

### Comparing Conditioned and Primary Reinforcers

A final issue to consider is whether or not conditioned reinforcers, once established, are qualitatively different from primary reinforcers. One way this has been assessed was to determine if variables affect primary reinforcement and conditioned reinforcement differently. We provide three examples of commonalities. First, early work with chain schedules of reinforcement (e.g., Ferster & Skinner, 1957) show, for example, that response patterns in earlier links (i.e., links not paired directly with primary reinforcement) of a chain schedule reflect response patterns seen in simple schedules with the same response contingencies. Second, Royalty et al. (1987), described above, showed effects of unsignaled delays for conditioned reinforcers. Williams (1976) showed a similar effect using unsignaled delays to primary reinforcement. Finally,

Shahan, Podlesnik, and Jimenez-Gomez (2006) presented pigeons a choice between two concurrently available observing responses. They showed that the ratio of delivery of conditioned reinforcers predicted the ratio of observing responses and that the data were well-described by the generalized matching law which was developed to describe the relationship between primary reinforcement ratios and response ratios (Baum, 1974; Baum & Rachlin, 1969). Taken together, these examples show that there are a number of shared features between conditioned and primary reinforcers.

There are of course differences between conditioned and primary reinforcers. However, it is not critical that conditioned reinforcers possess all of the functional capabilities of the primary reinforcer with which it is associated. In Pavlovian conditioning, CSs may elicit only a subset of the response that the associated US elicits (Mazur, 2006). Conditioned reinforcers may also show differences in function compared to the primary reinforcers with which they have been associated. For example, some researchers have found that conditioned reinforcers seem to be less “trans-situational” than primary reinforcers (Keehn, 1962; Schuster, 1969; Reinforcers are considered to have the property of transsituationality when they work in different situations. See Meehl, 1950 for more on this concept).

## Alternative Views: Challenges to Conditioned Reinforcement

Experimental evidence supporting conditioned reinforcement has often had alternative interpretations, and other theoretical constructs have been proposed as possible alternatives or supplements to the concept of conditioned reinforcement. We will briefly describe three: marking, bridging, and timing. All of these alternatives focus on the discriminative properties of a stimulus and how it signals distance to primary reinforcement.

### Marking and Bridging

Marking is a procedure in which a brief stimulus signals a delay that follows a choice response. The signal and the delay occur regardless of whether or not a correct choice was made. Bridging is similar to the marking procedure except that the stimulus is present during the entire delay. Both marking (Lieberman, McIntosh, & Thomas, 1979) and bridging (Rescorla, 1982) can facilitate acquisition of behavior relative to an unsignaled delay control group. Superficially, conditioned reinforcement, marking, and bridging procedures are quite similar: they all allow a connection between an event (stimulus or behavior) and its temporally distant consequence. Somehow they function to link the temporal gap between two events.

The critical distinction between conditioned reinforcement and marking/bridging procedures is that in conditioned reinforcement procedures the stimulus occurs only after a correct response (i.e. a response leading to primary reinforcement) has been made, whereas with marking and bridging the stimulus occurs regardless of the choice response. (Note that brief stimuli and stimuli that fill the entire delay are both considered in conditioned reinforcement accounts, with no distinguishing terminology.) One view is that marking effects “perceptually isolate the choice response, thus



making it more salient in memory at the time the response consequence is eventually delivered” (Williams, 1991, p. 264).

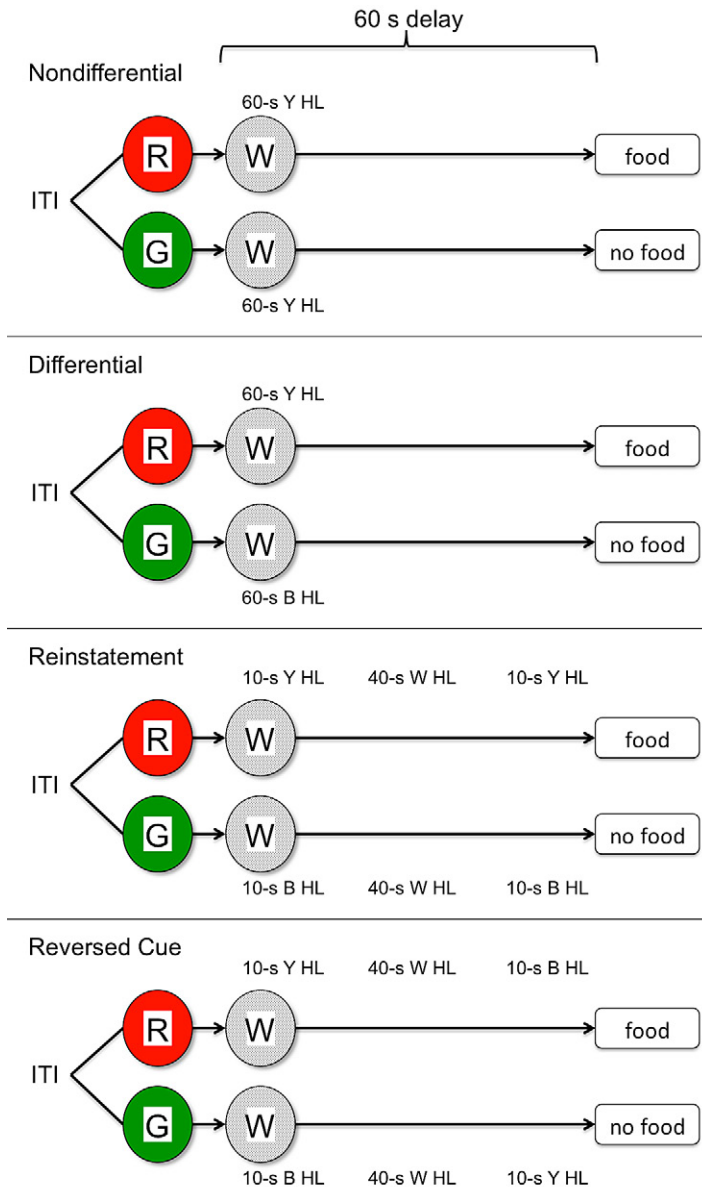
Williams (1991) directly compared marking and bridging procedures with a conditioned reinforcement procedure using rats in a two-choice conditional discrimination. Two different discriminative stimuli, a noise and a light, were used to indicate which response lever would result in food after a delay (3- to 12-s, depending upon the condition). Thus, the  $S^D$  was presented, and subjects chose. The  $S^D$  terminated, and a delay occurred. If subjects made a correct choice, the delay ended in food. If they made an incorrect choice, the delay ended without a food delivery. The critical manipulation was what occurred during the delay. In marking and bridging conditions, a tone stimulus followed both correct and incorrect responses. In marking conditions, the stimulus was presented briefly (0.5-1 s at the beginning of the delay) and in bridging conditions, the stimulus filled the entire delay. In the conditioned reinforcement conditions, a tone stimulus followed correct responses only. After subjects reached a training criterion, the contingencies were reversed (i.e., if the noise signaled that the left key was correct and the light that the right key was correct, in the subsequent condition, noise signaled the right key was correct and light signaled the left key was correct). Williams failed to find a marking or bridging effect, but did find that the conditioned reinforcement conditions resulted in fewer trials to criterion (i.e., faster learning) compared to all other conditions (no signal, bridging and marking conditions). Although marking and bridging procedures do appear to enhance learning under some conditions, they do not eliminate the need for the concept of conditioned reinforcement, as we will describe later.

### Timing

Timing explanations serve as an alternative approach to conditioned reinforcement and have attracted a good deal of attention. Timing accounts explain the effects of delays as a function of connections to primary reinforcement and rely upon the ability of a subject to maintain an accurate memory of the temporal relationship between events. In other words, rather than acquiring the functional properties of reinforcers, stimulus changes function to mark passage of time with respect to the primary reinforcer. One goal of this approach is to provide a unified theoretical approach to explaining operant behavior. As yet, these accounts have failed to provide a more parsimonious or complete account of operant conditioning compared to the theoretical construct of conditioned reinforcement (see Williams, 2003). Royalty et al. (1987), which was described above, provides some evidence to support this claim. Recall that they showed differences in responding within the link of a chain schedule as a function of the unsignaled delays associated with that link. The fact that they demonstrated this effect while holding time to primary reinforcement constant is problematic for timing theories. In other words, if time is the only controlling variable, then there should be no difference in responding between links with an unsignaled delay and links with no delay (for additional evidence challenging a timing account, see Bell & Williams, 2013).

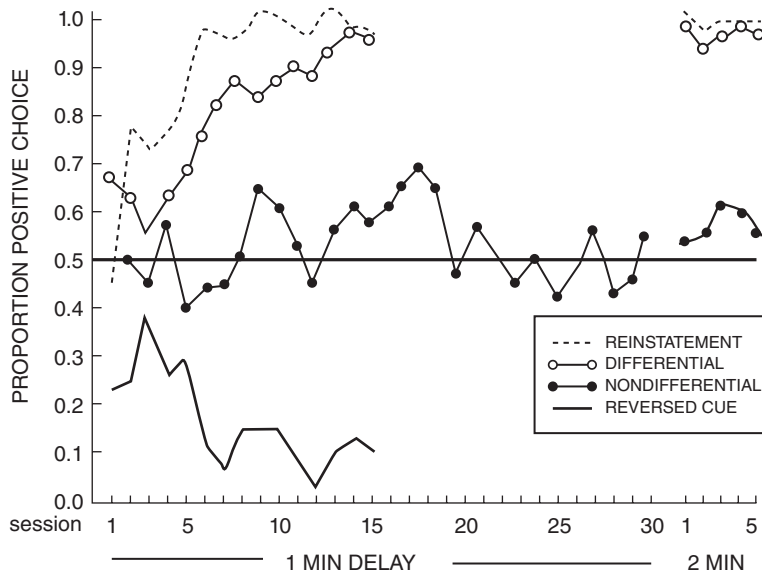
Cronin (1980) also provides a clear example of the difficulties encountered by accounts based solely on timing. The basic procedure involved a simple discrimination in which pigeons chose between two keys, one illuminated red and the other green





**Figure 10.6** Procedure schematic of Cronin (1980). Following a 2-min intertrial interval (ITI), a trial began with a choice between a red and green key light (KL). In this example, responses to red always led to food reinforcement and responses to green never did (i.e., extinction). Pecks to the red and green KL stimuli were followed by a change in KL color and house light (HL) color. Additional procedural details are provided in the text.

(see Figure 10.6). Responding to either stimulus was followed by a 60-s delay. A peck to the red stimulus led to food while pecks to the green stimulus resulted in no food. In the baseline condition, the delay period was signaled with a yellow house light and a white keylight. In the differential signal condition, the 60-s delay period was segmented as follows: Following a peck to the red stimulus a yellow houselight



**Figure 10.7** Mean choice proportion for the alternatives leading to food reinforcement for the four training groups in Cronin (1980, showing Figure 2 reprinted). The reinstatement and differential groups quickly learned to respond to the stimulus leading to food. The non-differential group remained at chance levels of performance, and never learned the task. Interestingly, the reverse cue group responded most to the stimulus that did not lead to food.

was illuminated for 10-s, followed by a white houselight for 40-s, and again followed by the yellow houselight for 10s and then food. Pecks to the green stimulus were followed by a blue houselight for 10s, the white houselight for 40s, and the blue houselight for 10s, then the end of the trial without food. In the reversed cue condition, the delays were also signaled, but choice of the red stimulus led to yellow, white, then blue, and choice of the green stimulus led to blue, white, and then yellow. The results (Figure 10.7) showed that birds in the baseline condition chose red about 50% of the time, and never learned the task. Birds in the differential signal condition quickly learned the task, approaching 100% correct. Birds in the reversed cue condition, however, went below chance to 9% correct (i.e., the pigeons were choosing the option that did not lead to food 91% of the time). This last finding offers compelling support of a conditioned reinforcement interpretation over a timing interpretation (or bridging or marking). If the stimuli merely mark the passage of time, one stimulus should be as good as any other—they should be functionally identical and completely interchangeable. The fact that responding went below chance accuracy in the inconsistent signal condition suggests that the blue stimulus acquired conditioned reinforcement value as a function of being paired, albeit infrequently, with reinforcement when birds chose the red alternative. Thus, choice between red and green essentially involved pitting a long delayed outcome of a conditioned reinforcer (blue) and food against an immediate conditioned reinforcer.

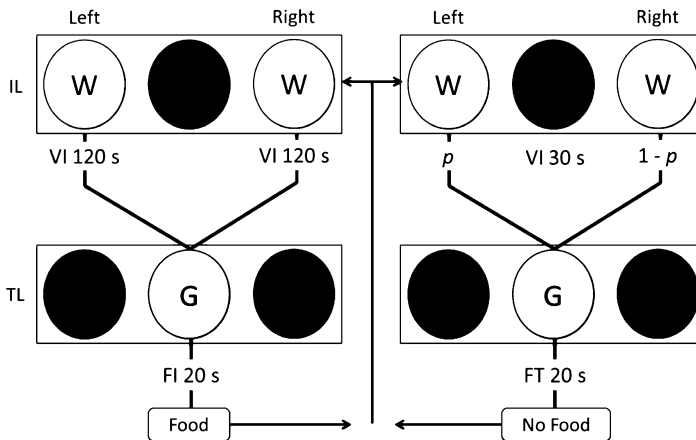
### Other Challenges

Schuster (1969) reported a series of studies often cited as evidence contradicting conditioned reinforcement and worth some consideration here. The first two experiments were multiple VI VI schedules in which a brief stimulus presentation was added to one schedule on a fixed ratio (FR) 11 schedule. FR schedules require a certain number of responses for re-inforcer delivery (Ferster & Skinner, 1957). The brief stimulus presentations were either paired or unpaired with primary reinforcement deliveries across conditions. He found that when the brief stimulus presentations were regularly paired with reinforcement, relative response rates increased for all subjects. Schuster claimed, however, that this finding did not constitute a conditioned reinforcement effect as the results lacked a clear contrast effect—in other words, there was not an accompanying decrease in responding to the schedule without the brief signal presentations. Although it is well known that an increase in reinforcement rate in one component can reduce responding in other components (Reynolds, 1961), this relationship is complex (see, for example, Williams, 1989, 1988, 2002; Williams & McDevitt, 2001).

Schuster (1969) also reported studies using a concurrent chains procedure in which an additional stimulus was presented on only one of two terminal links. Schuster reported no effects of the added stimulus presentations, and thus his results do not support a conditioned reinforcement interpretation. However, Williams and Dunn (1991a) raised several noteworthy criticisms of this procedure, and suggested that Schuster's results should be interpreted with caution. The most critical issue they identified was that the added stimulus presentations were never presented contingent upon choice responding. Thus, any conditioned reinforcement effect would have to have been mediated through the terminal-link stimulus and would require the stimulus to acquire conditioned value through second-order conditioning (as opposed to being directly paired with the primary reinforcer).

Williams and Dunn (1991a) carried out an extension of the Schuster (1969) procedures. In their procedure (Figure 10.8), pigeons chose between two equal VI initial-link schedules, each leading to an FI terminal link. The key manipulation was adding non-reinforced FT terminal-link presentations differentially to the two-choice alternatives. This manipulation makes terminal-link stimulus presentations contingent upon choice responding, addressing the flaw in Schuster's design. The distribution of additional terminal-link presentations was systematically varied across conditions. In some conditions the presentations were 50/50 across the two alternatives. In others, the distribution was 80/20. Initial-link preference varied as a function of the distribution of the additional terminal-link stimulus presentations, with higher responding occurring to the alternative paired with 80% of the additional stimulus presentations. In other words, subjects preferred the alternative that led to more terminal-link presentations and, as a result, proportionally fewer primary reinforcers. These results demonstrate that the additional terminal-link presentations functioned as conditioned reinforcers.

On another front, Staddon (1983, 2010; see also Staddon & Ettinger, 1989; Staddon & Cerutti, 2003) suggests that there are other problems with a traditional conditioned reinforcement account, which hinge upon a few specific issues. First,



**Figure 10.8** Schedules of center key light presentations. The left panel shows the concurrent chain schedule leading to food reinforcement. Initial links were equal VI 120-s schedules followed by equal FI 20-s terminal links, after which food was delivered. The conditioned reinforcement schedule, shown on the right, was superimposed on the food schedule. Here initial-link responding moved subjects to the terminal link according to a single VI 30-s schedule where entry to the terminal link from the left or right keys was varied with a probability  $p$  across conditions. When subjects entered this terminal link, the center key light terminated on a fixed-time (FT) 20-s schedule (and did not end in food). Reprinted from Williams and Dunn (1991a).

some experiments fail to demonstrate conditioned reinforcement effects. For example, Catania, Yohalem, and Silverman (1980) used pigeons to compare responding on three-link chain schedules (with FI 30-s links) with several yoked control conditions, including a tandem schedule control. As responding did not differ between the chain and tandem schedules, they concluded that the stimulus changes in the chain schedule did not function as conditioned reinforcers. This conclusion has been challenged by Royalty et al. (1987), described earlier, who suggested that the control conditions were not appropriate because, in part, they failed to control for overall time to reinforcement and, as we noted earlier, the response rates drop in early links of long chains thus extending the overall delay to primary reinforcement. In other words, the overall rate of primary reinforcement decreased in the chain schedules.

Second, Staddon (1983, 2010) discusses the finding that increasing the number of links in a chain can eliminate responding. Typically chain schedules with four or more links will not maintain much, if any, responding in early links. Staddon suggests this failure challenges a conditioned reinforcement account and that mere proximity to primary reinforcement is the key variable controlling responding. Contrary to this claim, low rates of responding in early links are compatible with a conditioned reinforcement account. Again, stimuli in the early links of a chain schedule are distant from reinforcement and therefore signal extinction. Any conditioned reinforcement effects would have to be mediated through the following links. (Davison & Baum, 2010, provide an alternative perspective.) The related Pavlovian literature also shows that conditioning beyond the second order is difficult, at best, to achieve (e.g.,

Murphy & Miller, 1957). Finally, Staddon claims that simple temporal proximity to food is sufficient to explain responding in chain schedules. Once again, Royalty et al. (1987) provided clear evidence that the contingencies between links of the chain play a key role.

Recently, it has been argued that the very notion of conditioned reinforcement needs to be reconsidered (Shahan, 2010, 2013), and this argument is certainly not new. In fact, Hendry (1969) suggested that it would perhaps be “more prudent and objective to avoid the connotations of the word ‘reinforcer’” (p. 401). However, no superior concept has yet been proposed to replace it. In any case, research into conditioned reinforcement will continue. The concept of conditioned reinforcement appears indispensable. For any theory of learning to be complete, it must include an explanation of how a (previously) neutral stimulus acquires the ability to reinforce responding.

## Applications of Conditioned Reinforcement

Notwithstanding the theoretical issues that have surrounded conditioned reinforcement, its importance as a key concept in the development of techniques to change and maintain both human and nonhuman behavior has been enormous. Our daily lives consist of sequences of events and responses and chains of behavior, and conditioned reinforcement propels us through those chains (and lack of it impedes our progress).

Conditioned reinforcement can also make delayed reinforcement more effective. There are times when it is not possible (or desirable) for a primary reinforcer to immediately follow our behavior, but a conditioned reinforcer can be applied quickly and effectively. For example, a simple “thank you” might be all that is necessary to strengthen a behavior. Of course, life often involves layers of conditioned reinforcers. For example, our work might evoke a “good job” from a colleague, but is also reinforced by the conditioned reinforcement associated with a direct deposit notice from our employer, which then leads to a succession of other conditioned and primary reinforcers. Learning to use conditioned reinforcers effectively is a vital part of learning to effectively change and maintain behavior.

### Animal Training

Animal trainers rely heavily on conditioned reinforcers. In fact, the recent popularity of “clicker training” attests to this (Pryor, 2004). The first step in these programs is to pair the clicker or other sound with a particularly delicious treat. The pet owner “clicks,” and then immediately gives the pet a small morsel of food. Once the sound of the clicker has become an established conditioned reinforcer, it can then be used to reinforce particular behaviors.

Conditioned reinforcers have a number of benefits compared to primary reinforcers. First, they are often more portable, and easily dispensed. A parent can easily apply a verbal conditioned reinforcer, such as the verbal response “Nice sharing!” to a child, regardless of where they are, as the response needs no direct environmental support, except the periodic pairing with primary or other reinforcers. Conditioned reinforcers

can also be more effective than primary reinforcers in terms of pinpointing the portion or characteristic of a response that was responsible for producing reinforcement. For example, a clicker that has developed conditioned reinforcing properties can be presented at the point at which a dog's hind legs clear a bar during a jump, indicating the part of the response that was deemed favorable by the trainer. Finally, conditioned reinforcers can effectively lengthen animal training sessions, as they do not result in satiation the way food reinforcers do.

Interestingly, in spite of the widespread popularity of clicker training, we were unable to find much research reported on its use with pets. Meyer and Ladewig (2008), for example, report what appears to be one of the first systematic studies of dog clicker training. For more information on training, see Pryor's chapter in this volume.

### Token Economies

To see the enormous power of conditioned reinforcement, one must only consider the potent effects of money on human behavior. A "token" is an object that can be exchanged for other reinforcers. It is also called a "generalized reinforcer," as it does not need to rely on a single reinforcer in order to be effective. Money is one such token and maintains much of our behavior, but other kinds of tokens can be used to establish or maintain specific behaviors.

Token economies have been used in a variety of work environments in order to support particular behavioral goals. For example, Fox, Hopkins, and Anger (1987) used a token economy to successfully increase safety at two open-pit mines. Employees at the mines earned stamps for working without accidents and injuries, and could exchange their stamps for tangible items. Token economies have also been used effectively to maintain appropriate classroom behavior (e.g., Boniecki & Moore, 2003), treat drug addiction (Eriksson, Götestam, Melin, & Ost, 1975), reduce disruptive noise in a group living situation (Meyers, Artz, & Craighead, 1976), and modify countless other behaviors. See Hackenberg (2009) and Kazdin & Bootzin (1972) for reviews of the research related to and application of token economies, as well as a discussion of the challenges faced in implementing them.

### Future Directions

As we have seen in this chapter, the concept of conditioned reinforcement holds a central place in the field of learning. The power of operant conditioning lies in the use of consequences to change future behavior. Because much of what reinforces our behavior is not primary reinforcement, conditioned reinforcement plays an even more important role in understanding human behavior although interest in conditioned reinforcement appears to have dropped off in recent decades (Williams, 1994). However, some issues relating to our understanding of conditioned reinforcement remain unresolved.

One issue relates to how to best estimate the value of a conditioned reinforcer and model the effects of conditioned reinforcement on choice behavior. There are several possible approaches, including models based on the degree to which stimuli

function to signal a reduction in the waiting time to reinforcement (Fantino, 1969, 1977, 1981) or models that estimate the conditioned value based on time to reinforcement (Mazur, 1993). While the field surely benefits from the variety of approaches, there is no theoretical approach that can be considered a clear winner.

There also appears to be quite a divide between the basic research on conditioned reinforcement and applications based on that construct. The utility of clicker training and other token economies is clear, but the applications and the basic research are not clearly connected. This disconnect is a problem, as applications that are not based on a clear connection to the basic science are at best incomplete and at worst antithetical to the fundamental properties of behavior analysis.

## Summary

Conditioned reinforcement has an esteemed position in the experimental analysis of behavior, appearing in the earliest texts in the field. The early procedures provided evidence for conditioned reinforcement but were plagued with small, quickly-disappearing effects, confounding variables, and alternative interpretations. These early findings led to refined procedures for establishing the importance of conditioned reinforcement. The literature, taken as a whole, provides substantial and converging evidence that conditioned reinforcement continues to play an important role in learning theory. Some work has suggested alternative interpretations, including explanations focused exclusively on timing. However, timing-based explanations as yet do not provide a more parsimonious explanation for the literature than does conditioned reinforcement. Conditioned reinforcement continues to have enormous utility, providing a straightforward explanation for a variety of experimental results. Furthermore, the usefulness of the construct is shown in the ubiquity and central importance of conditioned reinforcement in applications such as animal training and token economies.

## Author Note

One challenge we faced in writing this chapter was the sheer corpus of work on this topic. Conditioned reinforcement has a long and varied history warranting serious consideration. It is an important and necessary conceptual construct (although, as you see, some disagree). A single chapter cannot adequately cover the topic in depth; that is not our goal. Rather, we want to provide an introduction to the topic, touching on key issues and present examples of important research in the area to provide a broad overview of the topic. We have provided a few review references at the end of the chapter for anyone interested in pursuing the topic. We thank Roger Dunn and Ben A. Williams for comments on an earlier version of this chapter.

## Notes

- 1 However, as Williams (1994) notes, neither Skinner (1938) nor Wolfe (1936) are definitive demonstrations of conditioned reinforcement.



- 2 There is a substantial literature on the concurrent chains procedure and various theories related to that procedure that is beyond the scope of this chapter (see, e.g., Fantino, 1977 and Davison & McCarthy, 1987, for reviews; see also Williams, 1988, 1994). Our purpose is to present this procedure specifically in the context of conditioned reinforcement.

## References

- Baum, W. M. (1974). On two types of deviation from the Matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 22, 231–242.
- Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, 12, 861–874.
- Belke, T. W., & Spetch, M. L. (1994). Choice between reliable and unreliable reinforcement alternatives revisited: Preference for unreliable reinforcement. *Journal of the Experimental Analysis of Behavior*, 62, 353–366.
- Bell, M. C. (1999). Pavlovian contingencies and resistance to change in a multiple schedule. *Journal of the Experimental Analysis of Behavior*, 72, 81–96.
- Bell, M. C., & Williams, B. A. (2013). Conditioned reinforcement in chain schedules when time to reinforcement is held constant. *Journal of the Experimental Analysis of Behavior*, 99, 179–188. doi: 10.1002/jeab.10.
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw Hill.
- Bloomfield, T. M. (1972). *Reinforcement schedules: contingency or contiguity?* In R. M. Gilbert & J. R. Millenson (Eds.), *Reinforcement* (pp. 65–208). New York: Academic Press.
- Boniecki, K. A., & Moore, S. (2003). Breaking the silence: Using a token economy to reinforce classroom participation. *Teaching of Psychology*, 30, 224–227. doi:10.1207/S15328023TOP3003\_05.
- Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's keypeck. *Journal of the Experimental Analysis of Behavior*, 11, 1–8.
- Catania, A. C., Yohalem, R., & Silverman, P. J. (1980). Contingency and stimulus change in chained schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 33, 213–219.
- Cronin, P. B. (1980). Reinstatement of postresponse stimuli prior to reward in delayed-reward discrimination learning by pigeons. *Animal Learning & Behavior*, 8, 352–358.
- Davison, M., & Baum, W. M. (2010). Stimulus effects on local preference: Stimulus-response contingencies, stimulus-food pairing, and stimulus-food correlation. *Journal of the Experimental Analysis of Behavior*, 93, 45–59. doi:10.1901/jeab.2010.93-45.
- Davison, M., & McCarthy, D. (1987). The interaction of stimulus and reinforcer control in complex temporal discrimination. *Journal of the Experimental Analysis of Behavior*, 48, 97–116.
- Duncan, B., & Fantino, E. (1970). Choice for periodic schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 14, 73–86.
- Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior*, 53, 201–218.
- Eriksson, J. H., Götestam, K. G., Melin, L., & Ost, L.-G. (1975). A token economy treatment of drug addiction. *Behaviour Research and Therapy*, 13, 113–225.
- Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, 12, 723–730.
- Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313–339). Englewood Cliffs, NJ: Prentice Hall.



- Fantino, E. (1981). Contiguity, response strength, and the delay-reduction hypothesis. In P. Harzem & M. D. Zeiler (Eds.), *Advances in analysis of behaviour: Vol. 2. Predictability, correlation, and contiguity* (pp. 169–201). Chichester, UK: John Wiley & Sons, Inc.
- Fantino, E. J., & Logan, C. A. (1979). *The experimental analysis of behavior: A biological perspective*. San Francisco: Freeman.
- Fantino, E., Preston, R. A., & Dunn, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior*, 60, 159–169.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
- Fox, D. K., Hopkins, B. L., & Anger, W. K. (1987). The long-term effects of a token economy on safety performance in open-pit mining. *Journal of Applied Behavior Analysis*, 20, 215–224. doi:10.1901/jaba.1987.20-215.
- Gollub, L. R. (1958). *The chaining of fixed-interval schedules*. Unpublished doctoral dissertation, Harvard University.
- Gollub, L. R. (1977). Conditioned reinforcement: Schedule effects. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 288–312). Englewood Cliffs, NJ: Prentice Hall.
- Grace, R. C., Schwendiman, J. W., & Nevin, J. A. (1998). Effects of unsignaled delay of reinforcement on preference and resistance to change. *Journal of the Experimental Analysis of Behavior*, 69, 247–261.
- Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior*, 91, 257–286.
- Hendry, D. P. (1969). *Conditioned reinforcement*. Oxford: Dorsey.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 267–272.
- Jenkins, H. M., & Boakes, R. A. (1973). Observing stimulus sources that signal food or no food. *Journal of the Experimental Analysis of Behavior*, 20, 197–207.
- Jenkins, H. M., & Moore, R. A. (1973). The form of the autoshaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior*, 20, 163–181.
- Kazdin, A. E., & Bootzin, R. R. (1972). The token economy: An evaluative review. *Journal of Applied Behavior Analysis*, 5, 343–372.
- Keehn, J. D. (1962). The effect of post-stimulus conditions on the secondary reinforcing power of a stimulus. *Journal of Comparative and Physiological Psychology*, 55, 22–26. doi:10.1037/h0042005.
- Kelleher, R. T., & Gollub, L. R. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 5, 543–597.
- Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology: A systematic text in the science of behavior*. New York: Appleton-Century-Crofts.
- Kendall, S. B. (1973a). Effects of two procedures for varying information transmission on observing responses. *Journal of the Experimental Analysis of Behavior*, 20, 73–83.
- Kendall, S. B. (1973b). Redundant information in an observing-response procedure. *Journal of the Experimental Analysis of Behavior*, 19, 81–92.
- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, 21, 463–473.
- Lieberman, D. A., McIntosh, D. C., & Thomas, G. V. (1979). Learning when reward is delayed: A marking hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 224–242.
- Mazur, J. E. (1993). Predicting the strength of a conditioned reinforcer: Effects of delay and uncertainty. *Current Directions in Psychological Science*, 2, 70–74.
- Mazur, J. E. (2006). *Learning & behavior* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

- McDevitt, M. A., Spetch, M. L., & Dunn, R. (1997). Contiguity and conditioned reinforcement in probabilistic choice. *Journal of the Experimental Analysis of Behavior*, 68, 317–327.
- McDevitt, M. A., & Williams, B. A. (2010). Dual effects on choice of conditioned reinforcement frequency and conditioned reinforcement value. *Journal of the Experimental Analysis of Behavior*, 93, 147–155. doi:10.1901/jeab.2010.93-147.
- Meehl, P. E. (1950). On the circularity of the law of effect. *Psychological Bulletin*, 47, 52–75. doi: 10.1037/h0058557.
- Meyer, I., & Ladewig, J. (2008). The relationship between number of training sessions per week and learning in dogs. *Applied Animal Behaviour Science*, 111, 311–320.
- Meyers, A. W., Artz, L. M., & Craighead, W. E. (1976). The effects of instructions, incentive, and feedback on a community problem: dormitory noise. *Journal of Applied Behavior Analysis*, 9, 445–457. doi: 10.1901/jaba.1976.9-445.
- Murphy, J. V., & Miller, R. E. (1957). Higher-order conditioning in the monkey. *The Journal of General Psychology*, 56, 67–72.
- Myers, J. L. (1958). Secondary reinforcement: A review of recent experimentation. *Psychological Bulletin*, 55, 284–301.
- Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, 57, 301–316.
- Nevin, J. A., & Grace, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences*, 23, 73–130.
- Nevin, J. A., Tota, M. E., Torquato, R. D., & Shull, R. L. (1990). Alternative reinforcement increases resistance to change: Pavlovian or operant contingencies? *Journal of the Experimental Analysis of Behavior*, 53, 359–379.
- Preston, R. A., & Fantino, E. (1991). Conditioned reinforcement value and choice. *Journal of the Experimental Analysis of Behavior*, 55, 155–175.
- Pryor, K. (2004). *Getting started: Clicker training for dogs*. Dorking, UK: Interpet Publishing.
- Rachlin, H. (1976). *Behavior & learning*. San Francisco: W H. Freeman.
- Rescorla, R. A. (1982). Effect of a stimulus intervening between CS and US in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, 8, 131–141.
- Reynolds, G. S. (1961). An analysis of interactions in a multiple schedule. *Journal of the Experimental Analysis of Behavior*, 4, 107–117.
- Royalty, P., Williams, B. A., & Fantino, E. (1987). Effects of delayed conditioned reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior*, 47, 41–56.
- Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior*, 93, 269–289.
- Shahan, T. A. (2013). Attention and conditioned reinforcement. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, K. A. Lattal (Eds.), *APA handbook of behavior analysis, Vol. 1: Methods and principles* (pp. 387–410). Washington, DC: American Psychological Association.
- Shahan, T. A., Podlesnik, C. A., & Jimenez-Gomez, C. (2006). Matching and conditioned reinforcement rate. *Journal of the Experimental Analysis of Behavior*, 85, 167–180. doi:10.1901/jeab.2006.34-05.
- Schoenfeld, W. N., Antonitis, J. J., & Bersh, P. J. (1950). A preliminary study of training conditions necessary for secondary reinforcement. *Journal of Experimental Psychology*, 40, 40–45.
- Schuster, R. H. (1969). A functional analysis of conditioned reinforcement. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 192–235). Homewood, IL: The Dorsey Press.

- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Free Press.
- Spetch, M. L., Belke, T. W., Barnett, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure: Effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior*, 53, 219–234.
- Staddon, J. E. R. (1983). *Adaptive behavior and learning*. New York: Cambridge University Press.
- Staddon, J. E. R. (2010). *Adaptive behavior and learning* (Internet ed.) New York: Cambridge University Press.
- Staddon, J. E. R., & Cerutti, D. T. (2003). Operant behavior. *Annual Review of Psychology*, 54, 115–144.
- Staddon, J. E. R., & Ettinger, R. H. (1989). *Learning: An introduction to the principles of adaptive behavior*. Orlando, FL: Harcourt.
- Williams, B. A. (1976). The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, 26, 441–449.
- Williams, B. A. (1988). The effects of stimulus similarity on different types of behavioral contrast. *Animal Learning & Behavior*, 16, 206–216.
- Williams, B. A. (1989). Component duration effects in multiple schedules. *Animal Learning & Behavior*, 17, 223–233.
- Williams, B. A. (1991). Marking and bridging versus conditioned reinforcement. *Animal Learning & Behavior*, 19, 264–269.
- Williams, B. A. (1994). Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review*, 1, 457–475.
- Williams, B. A. (2002). Behavioral contrast redux. *Animal Learning & Behavior*, 30, 1–20.
- Williams, B. A. (1997). Conditioned reinforcement dynamics in three-link chained schedules. *Journal of the Experimental Analysis of Behavior*, 67, 145–159.
- Williams, B. A. (2003). Challenges to timing-based theories of operant behavior. *Behavioural Processes*, 62, 115–123.
- Williams, B. A., & Dunn, R. (1991a). Preference for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 55, 37–46.
- Williams, B. A., & Dunn, R. (1991b). Substitutability between conditioned and primary reinforcers in discrimination acquisition. *Journal of the Experimental Analysis of Behavior*, 55, 21–35.
- Williams, B. A., & McDevitt, M. A. (2001). Competing sources of stimulus value in anticipatory contrast. *Animal Learning & Behavior*, 29, 302–310.
- Williams, B. A., Ploog, B. O., & Bell, M. C. (1995). Stimulus devaluation and extinction of chain schedule performance. *Animal Learning & Behavior*, 23, 104–114.
- Wolfe, J. B. (1936). Effectiveness of token rewards for chimpanzees. *Comparative Psychological Monographs*, 12, 1–72.
- Wyckoff Jr., L. B. (1952). The role of observing responses in discrimination learning: Part I. *Psychological Review*, 59, 431–442.
- Wyckoff Jr., L. B. (1969). The role of observing responses in discrimination learning. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 237–260). Homewood, IL: Dorsey.
- Wyckoff Jr., L. B., Sidowski, J., & Chambliss, D. J. (1958). An experimental study of the relationship between secondary reinforcing and cue effects of a stimulus. *Journal of Comparative Physiological Psychology*, 51, 103–109.
- Zimmerman, D. W. (1957). Durable secondary reinforcement: Method and theory. *Psychological Review*, 64, 373–383.

Zimmerman, J., Hanford, P. V., & Brown, W. (1967). Effects of conditioned reinforcement frequency in an intermittent free-feeding situation. *Journal of the Experimental Analysis of Behavior*, 10, 331–340.

### **Additional Reading**

There are several excellent reviews (e.g., Myers, 1958; Kelleher & Gollub, 1962; Gollub, 1977; Fantino, 1977; Williams, 1994) and a text (Hendry, 1969) describing early work on conditioned reinforcement.