The Behavior Analyst 1994, 17, 261-285 No. 2 (Fall)

Tutorial

Conditioned Reinforcement: Experimental and Theoretical Issues

Ben A. Williams
University of California, San Diego

The concept of conditioned reinforcement has received decreased attention in learning textbooks over the past decade, in part because of criticisms of its validity by major behavior theorists and in part because its explanatory function in a variety of different conditioning procedures has become uncertain. Critical data from the major procedures that have been used to investigate the concept (second-order schedules, chain schedules, concurrent chains, observing responses, delay-of-reinforcement procedures) are reviewed, along with the major issues of interpretation. Although the role played by conditioned reinforcement in some procedures remains unresolved, the results taken together leave little doubt that the underlying idea of conditioned value is a critical component of behavior theory that is necessary to explain many different types of data. Other processes (marking, bridging) may also operate to produce effects similar to those of conditioned reinforcement, but these clearly cannot explain the full domain of experimental effects ascribed to conditioned reinforcement and should be regarded as complements to the concept rather than theoretical competitors. Examples of practical and theoretical applications of the concept of conditioned reinforcement are also considered.

Key words: conditioned reinforcement, behavior theory, observing behavior, chain schedules, delay of reinforcement, concurrent chains

A general assumption in contemporary behavior analysis is that human behavior is best understood in terms of the contingencies of reinforcement operating on that behavior. Yet much, if not most, human behavior has little immediate impact on satisfying the biological motives that underlie the reinforcement contingencies commonly studied in the laboratory. People are not born with a tendency to work for money, to like the taste of alcohol or coffee, or to discover laws of behavior. We are also not born with the motivation to engage in compulsive hand washing or to be fearful of speaking in public. Such motives, both positive and negative, are learned, and a major task of any behavior theory is to specify how such learning occurs, both in order to have a complete theory of behavior and for rectifying motives that are maladaptive.

The theoretical concept that has long served the primary burden for understanding acquired motives is conditioned reinforcement (also called secondary reinforcement, especially in the early literature). Its essential idea is that an initially neutral event acquires value because of its relation to primary reinforcement, and subsequently can serve as an effective reinforcer in its own right. The phrase, "its relation to primary reinforcement," is sufficiently general to encompass several different more specific accounts of precisely how conditioned reinforcers are established. This issue will be considered in greater detail later. To foreshadow that discussion, the current consensus is that conditioned reinforcers acquire value in their own right because of Pavlovian conditioning. That is, a stimulus paired with a Pavlovian reinforcer acquires not only the ability to elicit the responses appropriate to the Pavlovian reinforcer but its reinforcing value

---

Preparation of this paper was supported by grants from the National Science Foundation. Correspondence regarding the paper should be addressed to the author, Psychology Department, UCSD, La Jolla, CA 92093-0109.

as well. The implication is that variables that govern Pavlovian conditioning also determine when initially neutral stimuli will become conditioned reinforcers.[1]

Although the present discussion will consider only events that assume positive value because of their conditioning history, it is important to recognize that events may assume negative value as well, thus serving as conditioned aversive stimuli that may motivate various types of avoidance and escape behavior that have served as models of neurotic behavior (see Levis, 1989, for a recent rendition of this approach). We will confine ourselves to positive conditioned reinforcement, both because of the need to limit the range of discussion and because the experimental literatures for the acquisition of positive and negative value have evolved into essentially separate enterprises. It is important to appreciate, nevertheless, that the same conceptual issues occur in both arenas.

At first glance, discussion of the concept of conditioned reinforcement seems to be superfluous. It has long been accepted as a valid explanatory concept, and a variety of different experimental procedures have been used to demonstrate its effects on behavior. In fact, however, considerable disagreement has developed over the past two decades about its applicability to several different learning procedures, and several of the most eminent contemporary learning theorists have directly challenged whether putative conditioned reinforcers really do serve as reinforcers in the same sense as do primary reinforcers. For example, Rachlin's (1976) widely used textbook questioned whether conditioned reinforcers actually acquire conditioned value, and suggested that the effects of conditioned reinforcement contingencies, although often potent, occur for reasons other than the process of reinforcement. A similar challenge has been offered by Staddon (1983; also see the subsequent textbook of Staddon & Ettinger, 1989). Other major behavior theorists have continued to view conditioned reinforcement as a valid concept, but have argued that several situations in which it has been invoked do not really involve conditioned reinforcement after all. For example, Catania (1992, p. 175) has argued that conditioned reinforcement effects are relatively unimportant in chain schedules. The result of these challenges is that textbook writers now seem uncertain about how to portray the concept, with diminishing attention given to it. For example, the recent third edition of the introductory textbook on learning and behavior by Chance (1994) contains only a single paragraph dealing with the concept.

This reduced attention to the concept of conditioned reinforcement in recent years stands in marked contrast to its treatment during the heyday of behavior theory, when major behavior theorists such as Neal Miller, Robert Sears, John Dollard, and O. H. Mowrer used the concept as the essential explanatory idea for such diverse topics as personality theory, social behavior, psychopathology, and critical aspects of human cognition (e.g., Dollard & Miller, 1950; Mowrer, 1950, 1960). This enthusiasm for the concept was not limited to Hullian behavior theory, as is evident from the writings of Skinner (1953) and Keller and Schoenfeld (1950), in which the concept of conditioned reinforcement also plays an essential explanatory role for extrapolating the concept of reinforcement to human

---

[1] The idea that conditioned reinforcers are created by Pavlovian conditioning is only one of several hypotheses that have been advanced over the years. For example, Keller and Schoenfeld (1950) argued that a stimulus became a conditioned reinforcer as the result of being a discriminative stimulus. Others have argued that stimuli that provide information about the primary reinforcement contingency acquire conditioned reinforcement properties. The reason for not considering these alternative hypotheses here is that the consensus among researchers is that they have been discredited by considerable research (see Fantino, 1977, for a review). It should also be emphasized, however, that many instances of the acquisition of conditioned value cannot be reduced simply to the history of simple Pavlovian conditioning. Pavlov himself (1927) recognized the need for a "second signaling system" (the major function served by language) that produced mediated conditioning effects even when the stimulus was not itself directly paired with the Pavlovian reinforcer.

behavior in real-world situations. What then are we to make of the change in emphasis seen in the past two decades? An answer to this question requires an analysis of the specific explanatory functions served by the concept, the nature of the evidence supporting these functions, and the subsequent evidence that has inspired the questioning of the concept by writers such as Rachlin and Staddon. As a starting point in this endeavor, it is helpful to provide an overview of the extensive experimental literature on the topic.

METHODS FOR STUDYING CONDITIONED REINFORCEMENT

There are two general categories of procedures that have been used to study how conditioned reinforcement contingencies affect behavior. The first is when the stimulus that has been previously paired with the primary reinforcer is made contingent on some behavior without the accompanying presence of a primary reinforcer contingency. The second is when both a conditioned reinforcement and primary reinforcement contingency are applied to the same behavior, typically with a primary reinforcement contingency that in isolation will maintain the behavior only at a low level.

Conditioned Reinforcement Effects Isolated from Primary Reinforcement Effects

Resistance to extinction. The most common method used in early research to demonstrate the importance of conditioned reinforcement was to assess how resistance to extinction was prolonged when a conditioned reinforcer remained contingent on the response after the primary reinforcer had been removed. One of the earliest such experiments was reported by Bugelski (1938), in which he trained two groups of rats to press a bar for food pellets. After this initial training, responding of both groups was extinguished, but one of them continued to receive the click of the pellet dispenser contingent on bar pressing. Bugelski found that presentations of the click produced approximately 30% more bar presses during extinction than occurred when extinction was presented without the click.

Numerous other experiments have followed the basic design just described with similar results (see Kimble, 1961, for a review). Unfortunately, a major difficulty in interpreting these experiments is that it is unclear whether the increased number of responses during extinction is actually due to conditioned reinforcement. An alternative interpretation is in terms of generalization decrement: Subjects without the click during extinction had their stimulus situation changed more at the start of extinction than did subjects with the click present, and this difference in the degree of stimulus change may be solely responsible for the lesser resistance to extinction because it allows the onset of the extinction contingencies to be discriminated more readily. Although it is possible to reduce the degree of generalization decrement (e.g., by presenting the clicks independently of responding for a different control group), it can never be eliminated completely, with the result that the general procedure of using resistance to extinction as a measure of conditioned reinforcement effectiveness has been largely abandoned.

Conditioning a new response. Perhaps the clearest demonstration of the conditioned reinforcement properties of a stimulus is to use that stimulus to train a new response in the absence of the primary reward. Skinner (1938, p. 82) provides an early example of this method. Rats were first trained to approach the food cup at the sound of the pellet dispenser, and the food was then removed from the situation when a lever was introduced into the chamber. Lever presses produced the sound of the pellet dispenser but no food. The rate of bar pressing increased over the first 5 to 10 min, then decreased, and finally was reduced to near-zero levels after 30 to 45 min of training. Approximately 40 to 80 total bar presses occurred over that period of time.

Skinner's (1938) results are typical of many others that have used the new-response procedure. The response-contingent presentation of the stimulus previously paired with the primary reward[2] usually produces some initial level of acquisition, but then loses its ability to maintain the behavior as training continues. Presumably this loss of control over responding reflects the extinction of the conditioned value of the conditioned reinforcer, in that its presentation in the absence of the primary reinforcer removes the conditioned value that was originally established by the pairings of the initially neutral stimulus and the primary reward. Because this extinction process often occurs rapidly, the use of the new-response procedure with the primary reinforcer removed from the situation is of limited utility as a practical method of using conditioned reinforcement contingencies.

It is also important to recognize that Skinner's (1938) original demonstration cannot be interpreted unambiguously as evidence for the reality of conditioned reinforcement. Any convincing demonstration requires that the effects of the stimulus on the acquisition of the new response be due to (a) the contingency between the stimulus and primary reinforcer during the initial phase of training and (b) the operant contingency between the response and conditioned reinforcer during the response-acquisition phase. Each of these conditions requires control conditions to exclude alternative interpretations. For example, it is possible that mere stimulus change contingent on a response might produce some degree of response acquisition, in which case the Pavlovian conditioning during Phase 1 would be irrelevant. Similarly, the presentations of a Pavlovian positive conditioned stimulus (CS+) might produce increases in arousal, resulting in an increase in the level of the response being measured, quite apart from the conditioned reinforcer being contingent on the response. The great majority of experiments that have used the new-response procedure have failed to exclude both of these possibilities. This failure to include appropriate control conditions appears not to be crucial, however, because experiments in which appropriate control conditions have been employed have yielded an outcome essentially similar to that originally reported by Skinner (e.g., Hyde, 1976).

A variation of the new-response procedure that attempts to avoid the rapid loss of conditioned value during the extinction presentations of the stimulus has been to maintain separately the correlation between the stimulus and primary reward and the correlation between the response and conditioned reinforcing stimulus. This general approach was pioneered in a series of experiments by Zimmerman and his collaborators over the period of 1959 to 1967, and will be exemplified here by a description of one of their later studies (Zimmerman, Hanford, & Brown, 1967). Food was presented to pigeons on a background schedule of response-independent food presentations delivered after variable time periods (a variable-time, or VT, schedule). In addition, any scheduled food presentation was delayed until 6 s had elapsed without a response, in order to ensure that pecking behavior was not adventitiously followed by food. Conditioned reinforcers, contingent on key pecking, consisted of brief presentations (0.5 s) of the stimulus complex that ordinarily accompanied food presentation. Different frequencies of these brief stimuli were then presented across different experimental conditions. Response rate tracked the frequency of conditioned reinforcement, with a typical rate of five to seven pecks per minute with the highest rate of conditioned reinforcement, a variable-interval (VI) 1-min schedule. Training conditions were continued for 15 to 20 sessions for each schedule value, with no indication that response rate decreased with continued training. In addition, previous studies by Zimmerman's research group had shown that peck rate declined to a near-zero level when the stimuli contingent on responding had not been paired with food and when food

---

[2] When talking about a stimulus independent of its role in a response contingency, I try to use the term *reward*; when talking about a stimulus that is involved in a response contingency, I use the term *reinforcer*.

deliveries were discontinued, both effects being due, presumably, to the disruption of the pairing of the brief stimuli with food.

The results of Zimmerman et al. (1967) make a prima facie case that conditioned reinforcers can maintain operant behavior over a sustained period of time without that behavior ever being followed by the primary reinforcer. However, two caveats must be noted regarding the sufficiency of their control procedures. Their finding that pecking was not maintained by stimuli unpaired with the food delivery may not be taken as decisive evidence for the critical role of pairing, because the two types of stimuli were not equivalent, in that the unpaired stimuli were changes in the keylight stimulus but the paired stimuli consisted of a complex array of food-magazine sounds and houselight and keylight changes. Their finding that pecking was not maintained when the food deliveries were suspended also may not be crucial evidence for the role of the pairing, because it is possible that the sensory reinforcement properties of stimulus change per se may depend on food being present in the situation, as argued by Herrnstein and Loveland (1972).

The finding of Zimmerman et al. (1967) that the pecking behavior was maintained over a sustained period of time is somewhat surprising, given the likelihood of conditional discrimination of the food contingencies correlated with the two types of brief-stimulus presentations. That is, stimulus presentations followed by food always occurred after periods without key pecking, whereas their presentations when not followed by food always occurred immediately after a key peck. Given that pigeons can clearly discriminate whether their pecks do or do not precede stimulus change (e.g., Killeen, 1981), the issue is why a similar discrimination did not occur here as well. If such a discrimination did occur, presumably behavior would not be maintained by the brief stimuli in the absence of food, because such stimuli after a response would functionally serve as a signal for nonreinforcement.

The issue of discrimination is central to determining when behavior will or will not be maintained by conditioned reinforcement contingencies. A critical concern, therefore, is to define the conditions that retard the likelihood that such discrimination will occur. One important determinant is the schedule by which the conditioned reinforcing stimulus and the primary reward are initially paired. Rashotte, Marshall, and O'Connell (1981) have extensively investigated this issue using a second-order conditioning preparation with autoshaping. The basic autoshaping procedure involves presentations of a keylight stimulus, usually for 5 to 6 s, that is followed by food regardless of the pigeon's behavior. Pecking then occurs to the keylight despite the absence of any response contingency, due to the Pavlovian contingency between the keylight and food. Second-order conditioning may then be studied by first conditioning pecking to one stimulus (S1) and then randomly interspersing S1–food pairings with pairings of a second stimulus (S2) with S1 in the absence of food. Conditioning to S2 is then assumed to be the result of the conditioned reinforcement properties of S1. Note that with this procedure S2 serves two discriminative functions: It signals that S1 will follow and that food will not occur.

The issue addressed by Rashotte et al. (1981) was the rate of pecking during S2 and how that behavior was affected by the schedule of S2–S1 pairings. They noted that the acquisition of pecking to S2 was a function of the number of pairings between S2 and S1, regardless of the percentage of S2 presentations followed by S1. For example, 100 presentations of S2 in which 75 were followed by nothing and 25 were followed by S1 produced the same amount of conditioning to S2 as did 25 pairings of S2 all of which were followed by S1. In contrast to the absence of an effect of the percentage of S2 presentations followed by S1 during acquisition, this percentage had dramatic effects on whether S2 pecking was sustained. When 100% of the S2 presentations were followed by S1, S2 behavior quickly declined; when only 25% of the S2 presentations were followed by S1, S2 behavior continued, although with some decline, over many sessions of training.

Related effects with conditioned reinforcement contingencies have been reported by Klein (1959). Rats were trained initially in a straight alley ending in a distinctive goal box. Subsequently, the alley was converted to a T maze in which the goal box used during the initial training was one choice alternative and a different goal box not used in the initial training was the other choice alternative. Preference tests between the two goal boxes in the T maze occurred with no food presented. All rats received 20 trials per day of training on the straight alley before the preference test, with different groups receiving different percentages of food at the end of the alley (ranging from 100% to 20%). Consequently, the number of pairings between the distinctive goal box and food was five times as great for the 100% group as for the 20% group. Despite this difference in the number of pairings, the degree of preference for the goal box used during the original training was substantially greater for the groups with the smaller percentages of reinforcement. In other words, the conditioned reinforcement value of the goal box was more persistent the more intermittent the relation between the goal box and food.

The percentage of stimulus presentations followed by reinforcement is only one of several variables that affect the persistence of conditioned reinforcement effects (see, e.g., Astley & Perkins, 1985). Although such variables may prolong the period over which conditioned reinforcement contingencies will effectively control behavior, it must be recognized that conditioned reinforcement contingencies that are separate from primary reinforcement contingencies eventually will allow the subject to discriminate the true contingencies in the situation. This should not be surprising, because the procedures used to study conditioned reinforcement effects free of any contribution from primary reinforcement are identical with procedures used to create conditioned inhibitors. This is most obvious in the case of second-order conditioning with Pavlovian procedures, such as that used by Rashotte et al. (1981) described above. Presentations of a stimulus paired with the primary reinforcer (S1–unconditioned stimulus [US] pairings) occur alternately with a second stimulus paired with S1 (S2–S1 pairings). Despite the value of S1 being maintained by continued pairings with the primary reinforcer and the fact that some significant amount of second-order conditioning may initially occur to S2, the responding established to S2 ultimately ceases, as does responding to S1 on trials in which it is preceded by S2. In other words, the subject learns a discrimination between the two types of trials. Moreover, reacquisition of responding when S2 itself is paired with the primary reinforcer will be retarded; this is a standard test for determining when a stimulus has acquired conditioned inhibitory properties. Exactly what determines the transition from second-order conditioning to conditioned inhibition to S2 remains an important question in conditioning theory that has not been adequately answered. The critical observation for our present purposes is that an analogous set of dynamics also occurs with operant procedures in which the response is maintained by a conditioned reinforcer during one phase of the procedure and the stimulus used as a conditioned reinforcer is paired with the primary reward during an alternate phase of the procedure. With this procedure, the occurrence of the response eventually will serve as a negative predictor of the availability of the primary reward, and presumably, therefore, will cease to occur. It is critical to appreciate, however, that this finding does not itself challenge the validity of the concept of conditioned reinforcement. Instead, the finding only reveals the competing effects of other principles of conditioning that counteract the response-strengthening effects that otherwise would occur.

Contingencies Combining Conditioned and Primary Reinforcers

These competing effects are the primary reason that research on conditioned reinforcement has shifted to procedures in which the conditioned and primary reinforcement procedures are positively, rather than negatively, correlated. The major feature of these procedures is that the contingency for the primary reinforcer is typically weak, occurring only infrequently or after long time intervals, and the conditioned reinforcement contingency is then added to it. By then comparing the difference in the strength of behavior when the conditioned reinforcement contingency is present versus when it is not, the effects of conditioned reinforcement can be isolated. The most well known of such procedures are second-order schedules and chain schedules.

Second-order schedules. A typical example of a second-order schedule is provided by Kelleher (1966), who presented pigeons with a fixed-interval (FI) 4-min schedule of brief-stimulus presentations (0.7 s of a white keylight), 15 completions of which were required for food delivery, for a total interfood interval of 60 min. In general, behavior was well maintained under the brief-stimulus schedules, with response rates of approximately 35 pecks per minute despite the extremely low food rate. Response rate in the absence of the brief-stimulus presentations but with the same interfood interval (a tandem schedule) was 10 to 15 pecks per minute.

There is little doubt that second-order schedules of brief-stimulus presentations may produce major enhancements of response rates, and these may often be very useful in situations in which the frequency of presentation of the primary reinforcer must be limited for various reasons (e.g., when large doses of a drug are given that require substantial recovery time or other motivational effects independent of the response contingency; Katz, 1979). Nevertheless, these rate enhancements cannot be interpreted unambiguously as being due to conditioned reinforcement. One reason for this uncertainty is that often (but not always) the rate-enhancement effects are just as large when the brief stimuli are themselves not paired with the primary reinforcer just prior to its delivery (e.g., Stubbs, 1971). If such pairings are not required, it seems unlikely that the enhancements of response rate seen with such stimuli are in fact due to the conditioned value of the brief-stimuli, which presumably depends on their Pavlovian association with the primary reward. The experimental literature on the effects of paired versus unpaired brief-stimulus presentations is extremely complex, with no clear rationale for when the pairing operation will be critical (see Gollub, 1977, pp. 302–305 for a review; also see Cohen, Calisto, & Lentz, 1979, for an important subsequent study).

Even when the increase in response rate is larger with paired than with unpaired stimuli, this difference cannot be ascribed unambiguously to the conditioned value of the stimuli, because it is possible that the pairing operation may increase the salience of the brief stimulus independent of any effect on conditioned value. This possibility is important because both Fantino (1977) and Staddon (1983) have argued that the discriminative effects of the brief stimulus, not their reinforcing effects, are the causes of the increases in response rate that are typically observed. The assumption behind their interpretations is that the brief-stimulus presentations that occur near the end of the overall food interval are sufficiently close to the food presentation to signal its imminent availability, thus resulting in a subsequent increase in response rate. Brief-stimulus presentations earlier in the interfood interval serve a similar signaling function, in effect confusing the subject into responding as if it were nearer to food availability than it actually is, with an increase in response rate as the result. In other words, the rate increase that is seen depends on the subject not being able to discriminate the occurrences of the brief stimulus early in the interfood interval from those later in the interval when they are in fact proximal to food presentations. For this reason, Fantino (1977) has argued that the effects of brief-stimulus presentations are better characterized as being due to "conditioned confusion" than to conditioned reinforcement.

Chain schedules. In a chain schedule, a single primary reinforcer follows the completion of a sequence of individual schedule requirements, each of which is accompanied by a characteristic stimulus. These schedules are well known to almost all beginning students of behavior analysis, because they are commonly used for training animals to perform apparently complex behavioral feats. One example is a common laboratory exercise for students, who are provided with individual rats and a box-like apparatus with several detachable components: a hoop, a stick, a lever, and a marble. The task is then to train to rats to perform the sequence of running through the hoop, pulling the stick, pressing the lever, and pushing the marble out of the box, all components of which have to be completed for the rat to receive food. As is well known, such behavioral sequences are best trained in reverse order, such that the stimulus cueing the last component of the chain (the presentation of the marble) is used as the reinforcer contingent on the behavior in the penultimate component, the stimulus for the penultimate component (insertion of the lever) serves as the contingent reinforcer for the behavior in the preceding component (stick pulling), and so forth. The critical role of backward chaining[3] invites the interpretation that the stimulus for each link of the chain serves as a conditioned reinforcer for the behavior in the just-preceding link, and thus has acquired conditioned value because of its own relation to the food at the end of the chain. Given that chain schedules have often been taken as models of much complex human behavior (e.g., Keller & Schoenfeld, 1950), the implication is that conditioned reinforcement is a crucial explanatory concept for extrapolating laboratory principles to real-world behavior.

Given the efficacy of backward chaining as a training procedure (it continues to be used by commercial animal training enterprises such as Sea World), it is perhaps surprising that extended chain schedules studied in the laboratory have been found to maintain behavior very poorly in the initial links of the chain. For example, Gollub (1958, described in Gollub, 1977) presented pigeons with a five-component chain with FI 30-s schedules in each link of the chain, and found that response rate in the initial link of the chain was below 0.3 responses per minute, far below the response rate maintained during the first 30 s of an FI 150-s schedule correlated with the same stimulus throughout. This difficulty in maintaining behavior with extended chains has caused some investigators to question the role played by conditioned reinforcement in chain schedules. For example, Catania (1992) has argued that conditioned reinforcement effects operate only for the later links of the schedule relatively near food presentation. Staddon (1983) has made the stronger argument that no role is played by conditioned reinforcement contingencies in chain schedules; instead, the stimulus correlated with each link of the chain signals its own distinctive time to food, and it is this time to food that determines the response rate maintained by the stimulus, not the value of the following stimulus immediately contingent on responding during that link of the chain.

One approach to assessing the role of conditioned reinforcement in chain schedules is to determine whether the general laws of reinforcement apply to the contingency between responding and the stimulus change signifying access to the succeeding link of the chain. One such general law is that reinforcement effects are weaker when the reinforcer is delayed rather than when it immediately follows a response. To evaluate whether delay-of-reinforcement effects apply to chain-schedule behavior, Royalty, Williams, and Fantino (1987) imposed unsignaled delays of reinforcement on the different

---

[3] It should be noted that backward chaining is not necessarily the optimal procedure for training human subjects on similar sequentially structured tasks. Research with retarded subjects has shown, for example, that the "total task presentation," in which the subject practices all the steps from the beginning to the end of the chain on each attempt, may produce faster acquisition than backward chaining (see Martin & Pear, 1992, chap. 11, for a discussion). However, it remains unclear to what extent such alternative methods of creating behavioral chains rely on verbal instructions. With nonverbal subjects, backward chaining is clearly the method of choice, unless the sequence of chain components is very brief.

transitions between the links of a three-link chain. The rationale of the study was that similar delay-of-reinforcement contingencies have been shown to produce large decrements in behavior when imposed on primary reinforcement contingencies (e.g., Williams, 1976), so that any conditioned reinforcement effect in chain schedules should be similarly diminished when the delay contingency is added with respect to stimulus transition. In the baseline phase of training, pigeons were presented with a three-link chain in which all three links were VI 33 s. A delay contingency was then imposed on the transition between the initial and middle links of the chain, or on the transition between the middle and terminal links of the chain. The contingency between responding in the initial and middle links of the chain and food at the end of the chain was held constant throughout the experiment by shortening the VI schedule by 3 s whenever a 3-s unsignaled delay-of-reinforcement contingency was in effect (becoming a VI 30-s schedule with a 3-s delay). In other words, the delay contingency postponed the presentation of the stimulus correlated with the next link of the chain by 3 s but did not alter the relation between responding in the earlier links of the chain and food delivery. The results were that the delay contingency with respect to stimulus change produced effects on responding that were essentially like those previously found for food contingencies. When the delay contingency was in effect for the advancement from the initial to the middle link of the schedule, response rate in the initial link was decreased by 75% to 80%, with no effect on responding in the middle and terminal links of the chain. Similarly, when the delay occurred with respect to the advancement from the middle to the terminal link of the chain, responding in the middle link was reduced by 60% to 70%, again with no effects on responding in either of the other links of the chain. The specificity of the delay variable leaves little doubt that the contingency between responding and the onset of the stimulus of the succeeding link of the chain was crucial for maintaining the behavior.

Given that conditioned reinforcement contingencies do indeed play an essential role in maintaining behavior in the early links of a chain schedule, why then is behavior so difficult to maintain in the early links of an extended chain such as that studied by Gollub (1958)? One possible answer is that the conditioned reinforcement value of the stimuli correlated with the early links of the chain depend on higher order conditioning, such that the transfer of conditioned value back through the links of the chain is diminished with each additional link of the chain. For example, in the original studies of higher order conditioning by Pavlov (1927, Lecture III), conditioning did not occur with anything above the level of second-order conditioning using the standard conditioned salivation preparation, whereas third-order conditioning, but no higher, was possible with conditioned defense reactions. This cannot be the entire story, however, because the training of complex behavior sequences (such as the class exercise described above) often involves chains with a substantially larger number of links.

A possible alternative reason for failures to maintain behavior in extended chain schedules is the particular type of schedules used in the individual links of the chain. As was the case for Gollub (1958), the great majority of studies failing to maintain significant amounts of behavior have used FI schedules. The problem with this procedure is that the onset of an FI schedule is known to be a cue for nonreinforcement, as indexed by the substantial pauses that typically occur. Given that the stimulus onset is the event immediately contingent on behavior in the preceding link of the schedule, it should thus not be surprising that such stimulus onset has limited conditioned reinforcement effectiveness. The schedule during the initial link of the schedule itself also plays a role in its own right, as demonstrated by Williams and Royalty (1990), who showed that three-link chains with an initial VI component maintained a much higher rate of behavior in the initial link than did chains with an initial FI component, when all other aspects of

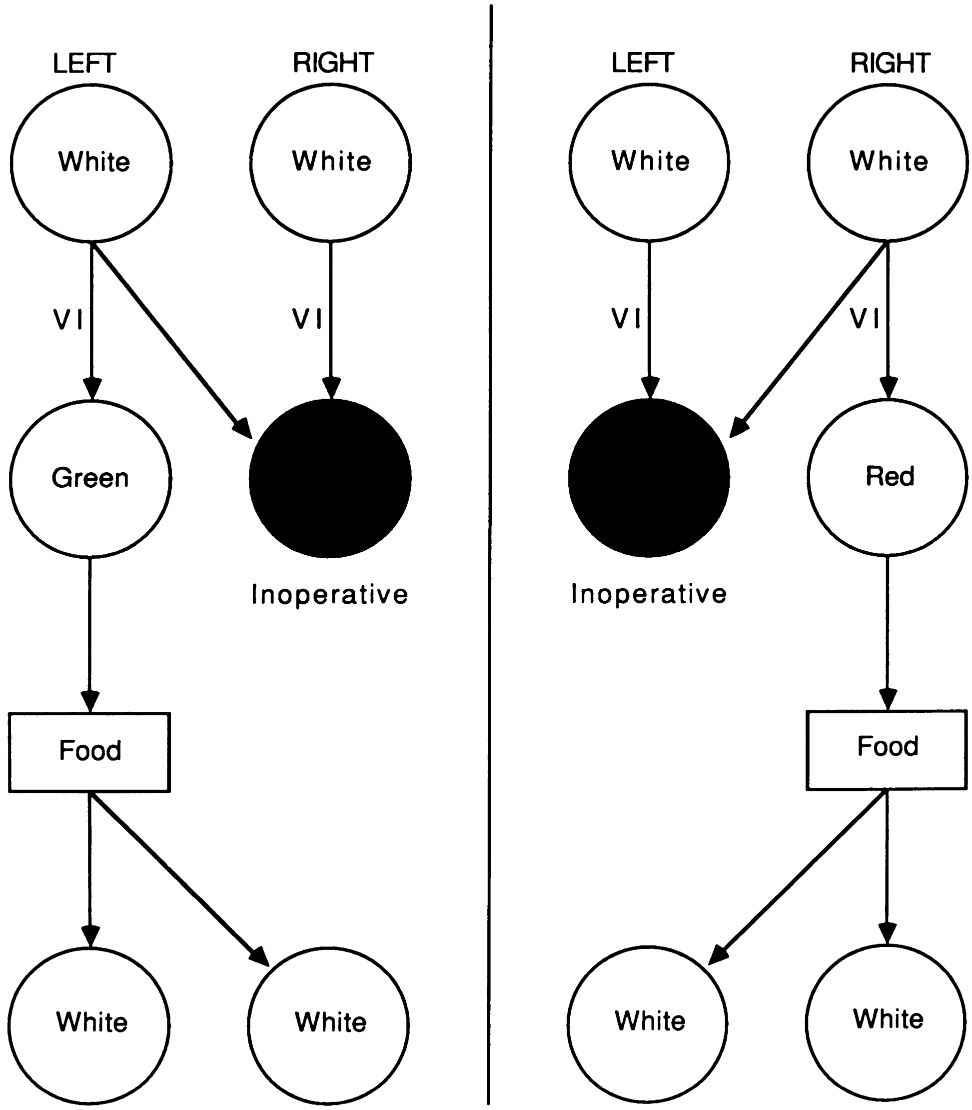

Figure 1. Schematic of the concurrent-chains procedure. Time runs from top to bottom. The left side of the figure shows the sequence of events if the left response key is chosen during the initial (choice) link of the schedule; the right side shows the sequence of events if the right response key is chosen during the initial link. After food is obtained in one or the other terminal-link stimulus, the procedure reverts to the choice phase in which both initial-link stimuli are available.

the schedule were equivalent. The difference in response rate in favor of the VI schedule was much greater than is typically the case when the comparison between VI and FI schedules involves simple schedules of primary reinforcement. Why the schedule type should be more critical for early links of a chain is unclear.

Concurrent chains. This procedure, diagrammed in Figure 1, has become perhaps the most widely used method for studying conditioned reinforcement effects, in part because it allows a quantitative analysis of the determinants of conditioned reinforcement value. As the name implies, the procedure involves a choice between chain schedules. The subject initially works on two concurrently available alternatives, each associated with its own reinforcement schedule (typically equal-valued VIs). Whenever either of these initial-link schedules is completed, a new stimulus appears, along with the terminal link of the chain appropriate to the selected initial-link response. During either terminal link, the other key becomes dark and inoperative, which means that choice is possible only during the initial link of the chains, not during the terminal links. Responses in the presence of the terminal-link stimuli earn primary reinforcement on some schedule. In most experiments, a single primary reinforcement reinstates the initial link, and the entire procedure recycles. The assumption underlying the procedure is that choice proportions during the initial links reflect the relative value of the two terminal-link stimuli.

Although several different variables have been specified that affect the degree of preference in concurrent-chain schedules (see Williams, 1988, for a review), we will briefly consider only two major generalizations. First, the major determinant of conditioned reinforcement value is not simply the average rate of reinforcement associated with the different terminal-link stimuli. For example, if one terminal link is an FI schedule and the other terminal link is a VI schedule with the same average reinforcement rate, preference is strongly in favor of the VI schedule. The reason this is true is that the appropriate measure of value is the harmonic mean of the different times to reinforcement correlated with a terminal-link stimulus, measured from stimulus onset (e.g., Killeen, 1968). The harmonic mean corresponds to the average immediacy of reinforcement associated with the stimulus, which must be distinguished from the average rate of reinforcement. Why immediacy should be more fundamental than rate is not obvious, but its dominance has been confirmed by several different procedures (e.g., Mazur, 1986; Shull, Mellon, & Sharp, 1990; Shull, Spear, & Bryson, 1981).

The second major generalization to be noted is that the value of a given schedule during a terminal link depends not just on its absolute time parameters, but is also relative to the overall context of reinforcement in which the stimulus occurs. As is the case in Pavlovian conditioning, such context effects are fundamental to understanding conditioned reinforcement. Delay-reduction theory (Fantino, 1977) provides perhaps the most influential attempt to provide an explanatory framework for capturing context effects. Its basic premise is that the value of a stimulus is determined by how much the onset of the stimulus reduces the expected time to reinforcement, relative to the average time to reinforcement in the situation independent of the stimulus. To see the implications of this idea, it is helpful to work through an example provided by the results of Fantino (1969). He presented pigeons with a concurrent-chains procedure in which the schedules correlated with the two terminal-link stimuli were always VI 30 s and VI 90 s. Although the ratio of times to food associated with these two stimuli is 3:1, delay reduction implies that their relative value will vary strongly as a function of the average interreinforcement interval. To vary the average interval between food presentations, Fantino used three different initial-link schedules: VI 600, VI 120, and VI 40 s. The average interreinforcement time is calculated as the sum of the average time in the initial links and the average time in the terminal links of the schedule. For example, if equal VI 120-s schedules are operative in the initial links, the average time spent in the initial link will be 60 s (because either schedule may produce access to its terminal link). Given the terminal-link values of VI 30 and VI 90 s, the average time spent in the terminal links is 60 s, because the two different terminal links are presented equally often. Thus, the average reinforcement times corresponding to the three different initial-link schedule values used by Fantino were 360, 120, and 80 s. The delay reduction signaled by the onset of the terminal link is then given by what percentage of this average time remains at the point of the terminal link. For example, with an average interreinforcement time of 120 s, the onset of a 30-s terminal link signals a 75% reduction in time to food, whereas the onset of a 90-s terminal link signals a 25% reduction in time. Given the three different initial-link schedules used by Fantino (1969), corresponding delay reductions for the VI 30-s terminal-link schedule were .92, .75, and .63, whereas those for the VI 90-s terminal-link schedule were .75, .25, and −.11 (negative values refer to the stimulus onset that predicts an increase in the time to reinforcement relative to the average). Choice proportions for the shorter terminal-link schedule were then assumed to be determined by its proportion of the total of the delay reductions summed over both terminal links, and were .55, .75, and 1.0 for the different initial-link schedules. The obtained choice proportions corresponded closely to these predictions. Thus, with long initial-link schedules, preference was near indifference, but with the short initial-link schedule, preference was exclusively in favor of the shorter terminal link, despite the absolute values of the terminal-link schedules remaining unchanged.

A variety of data support the essential idea underlying delay-reduction theory (see especially Fantino & Dunn, 1983, and Preston & Fantino, 1991). Given that it was developed entirely within the concurrent-chains methodology (although note that it applies equally well to the observing response procedure; see below), it is noteworthy that its basic assumptions are essentially similar to those of a prominent account of Pavlovian conditioning—the scalar expectancy theory of Gibbon and Balsam (1981)—that was developed completely independently. This similarity provides further support for the previously presented assumption that conditioned reinforcers are established via the Pavlovian contingency between the stimulus and the primary reinforcer.

Observing responses. The procedure about which there is perhaps the most agreement that conditioned reinforcement contingencies can maintain behavior indefinitely is the observing response procedure pioneered by Wyckoff (1952). In his initial study, periods of reinforcement availability on an interval schedule, contingent on pecking a response key, were alternated with periods of extinction. Responses to a lever then changed the stimulus conditions from the same stimulus being continuously present on the key at all times to different stimuli being correlated with the different schedules (i.e., the observing response changed a mixed schedule into a multiple schedule). Responses to the lever had no effect on food delivery itself. The result was that a high rate of the "observing response" occurred. Numerous other studies have confirmed the generality of this finding (see Fantino, 1977, for a review).

The typical contingency used in an observing response experiment involves the presentation of both the positive and negative stimuli (S+ and S−) contingent on the performance of the observing response. Early research with the procedure was thus concerned with which aspect of the contingency was crucial to maintaining the behavior. Several investigators proposed that either stimulus should be sufficient because the critical ingredient was the information about which reinforcement schedule on the main operant key was in effect. This "information hypothesis" was contrasted with the predictions of conditioned reinforcement, which maintained that the reinforcing event was only the S+ stimulus because of its pairing with the food delivery. Many different experiments were conducted to separate these two accounts, with the result that the conditioned reinforcement hypothesis has been generally supported (Dinsmoor, 1983, provides the most recent review). Specifically, observing behavior appears not to be maintained by presentations of the S− alone (in fact, considerable evidence indicates that S− presentations are aversive and actually suppress the observing response), although some controversy still remains with respect to experiments involving human subjects (see Case, Fantino, & Wixted, 1985).

Although the great majority of observing response experiments have used standard operant chambers, usually with pigeons as subjects, several studies have used an E maze in which the start box is in the center and separate goal boxes are located in either arm. An important feature of the maze is that the stimulus conditions associated with the different goal boxes are not seen until the animal is beyond the choice point. The issue is which side of the maze will be chosen as a function of the stimulus conditions correlated with different outcomes in the different goal boxes. For example, when food is presented in each goal box on 50% of the trials, one side may always have the same stimulus present during both reinforcement and nonreinforcement while the other side will have different stimuli correlated with the two outcomes.

The majority of E-maze studies have produced results similar to those with free-operant methodology: The side associated with differential stimuli for the two outcomes is preferred over the side without differential stimuli. However, there are important exceptions to this general pattern of results. Daly (1985, 1989) has identified several conditions in which preference favors the side with the unsignaled outcome. These include the combination of small rewards, minimal delays between the choice response and access to the goal box, and high percentages of rewarded versus nonrewarded trials (e.g., preference for the unsignaled side when 75% of the trials end in food, but preference for the signaled side when 25% or 50% of trials end in food). In addition, the administration of alcohol increases preference for the unsignaled side. Daly interprets these results in terms of Amsel's (1992) frustration theory: In most situations, preference favors the signaled outcome because a greater amount of conditioned frustration occurs to the stimulus of the unsignaled alternative, when both reward and extinction are presented during the same stimulus, than to the stimuli of the signaled alternative, when nonreinforcement occurs during a stimulus that otherwise is not paired with food. Whenever the degree of aversiveness caused by nonreward is reduced (e.g., by using small rewards or by alcohol ingestion), the role of conditioned frustration is diminished. Preference may then occur for the unsignaled alternative because of the greater frequency of conditioned reinforcement that occurs with unsignaled outcomes (i.e., the subject receives a stimulus on all trials with a history of food pairings rather than on only half of the trials with the signaled outcome). This interpretation is notably different from that based on conditioned reinforcement, and if valid, suggests that observing response procedures have complexities not heretofore appreciated. As yet, however, there have been no attempts using conventional free-operant procedures to test whether the variables specified by Daly have effects in the typical free-operant observing response procedure that are similar to her E-maze results.

Delay-of-Reinforcement Contingencies in Discrimination Learning

The set of procedures that have been described above generally have involved free-operant schedules of reinforcement in which response rate, or relative rate, has been the measure of conditioned reinforcement effectiveness. A somewhat simpler, and historically earlier, method of studying the role of conditioned reinforcement has been the insertion of stimuli during delay-of-reinforcement intervals in a discrimination procedure. In the absence of such stimuli, behavioral control by the delayed reinforcement contingency is often very weak. Insertion of the stimuli into the delay interval typically strengthens that control, often to an extent not notably less than the degree of control by immediate reinforcement contingencies.

An especially instructive example of the dynamics of conditioned reinforcers occurring during delay-of-reinforcement intervals comes from Cronin (1980), who trained pigeons on a simultaneous discrimination between line orientations. When a 60-s delay separated the choice response from the food delivery after correct choices, acquisition of the discrimination failed to occur even after 30 ses-

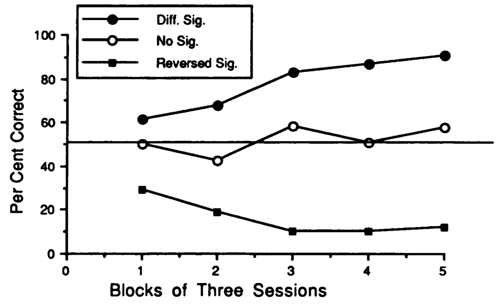

Figure 2. Data for three of the experimental conditions of Cronin (1980). Different groups of subjects received the same simultaneous discrimination with a 60-s delay-of-reinforcement interval, but with different stimulus events during the delay. The straight line in the middle of the figure is the level of performance expected by chance (50% correct). See text for details of the procedure for the different groups.

sions of training. These results are shown for the first 15 sessions of training as the "no sig" condition in Figure 2. When choice of the S+ versus S− was followed by different-colored houselights extending throughout the delay interval ("diff. sign" condition in Figure 2), rapid acquisition occurred. This major facilitation in the rate of learning is typical of many other studies of discrimination learning under delayed reinforcement contingencies, including the classic study of Grice (1948). The issue has been how to characterize the facilitory effect of these differential stimuli during the delay interval. Strong support for the conditioned reinforcement interpretation of these stimulus effects was shown in Cronin's "reversed cue" condition, in which birds received one color of houselight (yellow) for the first 10 s after an S+ and then a different houselight color (blue) during the last 10 s of the delay prior to food. After an S− choice, the houselight presentations were reversed: The blue houselight occurred immediately after an S− choice, and the yellow houselight occurred during the last 10 s of the 60-s delay before the onset of the next intertrial interval. As shown in Figure 2, this "reversed-sig" condition not only failed to produce discrimination acquisition, but it also resulted in a consistent choice of the S− stimulus. Presumably this was because the choice of the S− had the immediate effect of producing the stimulus (blue houselight) that preceded food on the trials on which the S+ was chosen. Thus, preference for the conditioned reinforcer was sufficiently strong that it controlled behavior at the expense of primary reinforcement; this result can only be interpreted as showing that the immediate conditioned reinforcer possessed considerable value in its own right.

Reprise

The preceding abbreviated review of the major procedures used to study conditioned reinforcement has been intended to convey some sense of the complexity of the issues and difficulties in interpretation. Although some of the classic demonstrations of conditioned reinforcement (e.g., Skinner, 1938; Zimmerman et al., 1967) lack the necessary control conditions to be interpreted unambiguously, subsequent work generally supports the generality of their findings. On the other hand, some of the major procedures still used to study conditioned reinforcement (e.g., second-order schedules) suffer from significant ambiguities of interpretation, and procedures that for the most part provide strong evidence for the concept (e.g., observing responses) have complexities that are not yet understood (which may eventually cause a rethinking of the role of conditioned reinforcement in these procedures). On the other hand, some procedures for which the role of conditioned reinforcement has been widely disputed (e.g., chain schedules) clearly do require the concept in light of more recent evidence. The strongest support for the concept comes from the classic procedure of using conditioned reinforcers to fill delays of reinforcement, where critical experiments (e.g., Cronin, 1980) leave little doubt that conditioned reinforcement is an essential explanatory concept.

CONDITIONED REINFORCEMENT VERSUS CONDITIONED FRUSTRATION

Given the preceding discussion, the issue becomes why there has been so much dubiety regarding the validity of conditioned reinforcement as a key concept in behavior theory, with the result that it has been increasingly ignored in recent textbooks. The answer to this question perhaps lies with changes in the theoretical zeitgeist, and may not be explained easily in objective terms. For example, the rise in influence of molar maximization models of reinforcement (e.g., Rachlin, Battalio, Kagel, & Green, 1981) has deemphasized the importance of immediate response consequences as the primary determinant of behavior, and one of the key functions of conditioned reinforcement as an explanatory tool has been to provide such immediate response consequences. Quite apart from such overarching perspectives, significant substantive problems raised by the concept still remain unresolved, as can be seen from considering the observations of one of its most perceptive critics, Rachlin (1976).

Rachlin does not dispute that conditioned reinforcement contingencies may have strong effects on behavior, but he argues that such effects occur for reasons other than the stimuli regarded as conditioned reinforcers having acquired conditioned value in their own right. He suggests two alternative mechanisms for their effect: to provide information that a response has registered and to serve as a signal that the primary reinforcer will eventually occur. Neither type of effect depends on the stimulus having acquired conditioned value.

To demonstrate the implausibility of the concept of conditioned value, Rachlin offers an analogy involving a comparison between two hotels. In one, a dinner bell (the putative conditioned reinforcer) is sounded prior to mealtimes. In the other, which is otherwise identical, the same dinner bell is sounded prior to mealtimes and at other times as well. If the bell has acquired conditioned value, argues Rachlin, we should prefer to stay in the hotel with the additional bell presentations. But intuition suggests that exactly the opposite preference would occur.

Empirical support for Rachlin's intuition comes from the results of Schuster (1969), who presented pigeons with a concurrent-chains procedure (see Figure 1 and discussion above) in which the conditioned reinforcer (the combination of brief stimuli correlated with food presentation) was presented on a fixed-ratio (FR) 11 schedule during one of the two terminal links. The initial-link and terminal-link schedules for the two choice alternatives were otherwise identical. Response rate during the terminal link with the brief-stimulus presentations was higher than in the alternative terminal link without the brief-stimulus presentations, indicating that the contingency had been effective. However, preference for that terminal link, as measured by the choice proportion during the initial link of the schedule, was substantially reduced. Thus, rather than the additional presentations of the food-paired cues enhancing the value of the terminal link in which they occurred, they actually seemed to lessen its value.

It should be apparent that the rationale behind Rachlin's hotel example and the supporting results of Schuster (1969) is very close to the concept of frustration, as developed and elaborated by Amsel (1992). A commonplace observation is that a stimulus that has been paired with reward in the past will elicit emotional reactions when presented unaccompanied by the reward. For example, rats finding a goal box empty which in the past contained food will become aggressive and attempt to escape from the goal box. Thus, increases in response rate, such as those observed by Schuster during the terminal link in which extra brief-stimulus presentations occurred, may not necessarily reflect an increase in conditioned value but instead may be only the rate-enhancement effects of frustration-induced arousal. One of the major conceptual paradoxes involved in the study of conditioned reinforcement is that the operation used to detect conditioned reinforcement effects—the presentation of a stimulus alone that previously has been paired with reward—is identical to the operation presumed to cause frustration. It is for that reason that Rachlin's example has intuitive appeal.

How the concepts of conditioned frustration and conditioned reinforcement will be reconciled remains uncertain. Nevertheless, research subsequent to Schuster (1969) has shown that contrary to his earlier findings and Rachlin's intuition, additional nonreinforced presentations of a stimulus otherwise paired with reward do in fact possess conditioned value. Williams and Dunn (1991a) presented pigeons with a concurrent schedule in which both choice alternatives led to the onset of a green color on the center key, which itself was correlated with an FI 20-s food schedule. During the baseline phase of the experiment, the frequency with which pecking of either choice alternative produced this green stimulus and the presentation of the food at the end of the 20-s interval were equal, as determined by an independent VI 2-min schedule operative for each choice alternative. The critical manipulation was the superimposition on these equal food schedules of an independent VI 30-s schedule of presentations of the green stimulus, but now without food at the end of the 20-s stimulus period (i.e., additional extinction presentations). These additional extinction periods were assigned differentially to the two choice alternatives, 80% for one alternative and 20% for the other. Because the two choice alternatives were treated identically except for these additional extinction presentations, whether the animals responded more to the alternative associated with the more frequent presentations provides a test of how preference is determined by the frequency of conditioned reinforcement, and thus of whether the stimulus had acquired conditioned value. The results were unequivocal: All subjects strongly preferred the choice alternative associated with 80% of the additional extinction periods. Thus, unlike the results of Schuster (1969), our findings show that stimuli paired with food are valued even on occasions when they occur without terminating in food presentation. Why Schuster failed to obtain this effect is uncertain, although possible interpretations have been offered (Williams & Dunn, 1991a; Gollub, 1970).

MARKING AND BRIDGING AS ALTERNATIVES TO CONDITIONED REINFORCEMENT

The results of Williams and Dunn (1991a) make a strong case that the concept of conditioned value has validity. This of course should not be surprising, given that Pavlovian conditioning is generally agreed to be the basis of conditioned reinforcement, and value is one of the attributes of the Pavlovian reinforcer that is transferred to the CS during conditioning. Nevertheless, one may not infer that all conditioned reinforcement procedures control behavior because of the effects of conditioned value on responding. The reason is that the mechanisms suggested by Rachlin (1976) as alternatives to conditioned reinforcement have been shown to have merit in their own right.

Marking. The first mechanism suggested by Rachlin (1976) as an alternative to conditioned value is that a stimulus presentation contingent on a response may define the response as a significant event. Support for this general notion has come from research on the concept of *marking,* in procedures involving delays of reinforcement between the choice and outcome in a simultaneous discrimination procedure. The critical feature of the marking procedure is that a brief distinctive stimulus is presented after the choice response, regardless of whether the choice was for the S+ or S−. Thus, any effect of conditioned value should strengthen correct and incorrect choices equally and not have any differential effect on the rate of acquisition of the discrimination. Nevertheless, marking procedures have been shown to facilitate the rate of learning substantially, in comparison to when no signal is presented during the delay-of-reinforcement interval (Lieberman, Davidson, & Thomas, 1985; Lieberman, McIntosh, & Thomas, 1979). The interpretation of this facilitation has been that the brief stimulus perceptually isolates the choice response, making it more salient at the time the response consequence is eventually delivered at the end of the delay interval, thus facilitating the associative connection between the response and the contingent reinforcer despite their temporal separation. The critical feature of this account is that the ability of the stimulus to facilitate control by the delayed reinforcement contingency does not depend on an association between the response and stimulus itself, nor on the stimulus having value in its own right.

As an example of how the concept of marking might be applied to procedures conventionally interpreted in terms of conditioned reinforcement, consider the results of Royalty et al. (1987), described above, regarding the effects of a delay contingency between responding in the early links of a chain schedule and stimulus onset of the succeeding link. The fact that the delay contingency greatly decreased the response rate may tell us nothing about whether the effect of the stimulus onsets was due to their having acquired conditioned value. An alternative interpretation in terms of the concept of marking is that the function of the contingent stimulus presentations in the absence of any delay may be to cause the animal to better learn that responding in the early links of the chain is necessary for reaching the end of the chain (i.e., the delayed food contingency is made more effective). When stimulus onset is then delayed, the behavior that is marked is no longer restricted to pecking the response key, but may include behavior that competes with key pecking.

A further example of effects potentially interpretable in terms of marking comes from the study of delay-of-reinforcement contingencies used with simple VI schedules. It is well known that delay procedures with unsignaled delay intervals maintain much lower levels of behavior than when the delay interval is filled with a differential signal (e.g., Richards, 1981), and the usual interpretation of the signal effects is in terms of conditioned reinforcement. However, Schaal and Branch (1988, 1990) demonstrated that a brief signal (0.5 s) at the onset of the delay interval was also sufficient to maintain the behavior at a high level, even when the signal itself was not paired with food. Little difference in response rate was evident for conditions in which only the brief stimulus occurred versus conditions in which a continuous signal extended throughout the delay for delay values up to 10 s, although differences did occur with delays as long as 30 s. It is important to recognize that a delay of 10 s meant that the end of the signal was separated by 9 s from the food delivery, so that it seems plausible that a considerable reduction in the degree of conditioned value to the stimulus should be evident. The failure to find a corresponding reduction in response rate suggests that marking, rather than conditioned reinforcement, may be the proper interpretation.

Bridging. The second mechanism suggested by Rachlin (1976) as an alternative to the conditioned value interpretation of conditioned reinforcement effects is that of *bridging,* the idea that the stimulus somehow serves to connect the response with the eventual delayed contingent reinforcer. This concept has resisted a precise definition, but its differentiation from the conventional concept of conditioned reinforcement can perhaps be appreciated by considering a simple example. Suppose a teacher says "good" to a young child and then gives the child candy. According to the concept of conditioned value, the child will feel a warm glow of positive affect whenever he or she hears the word, and it is this reaction that makes the word an effective reinforcer in subsequent occurrences. According to the bridging concept, however, the word "good" is not valued for its own sake, but instead is only a means to an end. It signals that the candy will eventually be delivered. Such discriminative effects of the event then increase the level of behavior by making the delayed primary reinforcement contingency more effective.

It should be evident that such bridging effects of the signal are not easily separated empirically from effects based on conditioned value. Accordingly, the evidence in favor of the concept is considerably less robust than that supporting the concept of marking. Rescorla (1982) has provided several procedures that may potentially differentiate the notion of bridging from that of conditioned reinforcement, but other investigators have disputed whether the procedures really do allow an unambiguous interpretation (Honey, Schachtman, & Hall, 1987; Thomas, Robertson, & Cunniffe, 1989). For present purposes, it suffices to acknowledge the possibility of marking as a separate process, with the implication that many effects previously ascribed to conditioned reinforcement may have an alternative interpretation.

Despite the availability of the alternative concepts of marking and bridging, it is clear that neither can displace the concept of conditioned reinforcement in its entirety. The results of Williams and Dunn (1991a) on the role of conditioned value in a choice procedure, described above, provide one example that lies clearly outside the domain of these alternative theoretical notions, as do the results of the reversed-cue condition of Cronin (1980), also described above. Moreover, when the different alternative concepts have been pitted directly against each other (Williams, 1991), the concept of conditioned reinforcement has clearly been shown to be the preferred explanation. Nevertheless, it is also apparent that any particular behavioral situation will require its own analysis in order to determine exactly which processes are the critical determinants of behavior.

DETERMINANTS OF CONDITIONED VALUE

Given that conditioned reinforcement is a valid behavioral concept, the issue of central importance is how the value of a conditioned reinforcer is determined. The most obvious variable that should determine conditioned value is the rate of reinforcement in the presence of the stimulus. However, the earlier discussion of delay-reduction theory has shown this notion to be simplistic, in that the value of a given rate of reinforcement is relative to the context of reinforcement in which it is embedded. Rate of reinforcement in the presence of the stimulus also fails to predict the degree of conditioned value even when the context of reinforcement is held constant. Also as noted above, the critical variable appears to be the immediacy of reinforcement predicted by the stimulus onset, rather than rate per se. An example of the differences between immediacy and rate is provided by Shull et al. (1981), who presented pigeons with a choice between a constant VI schedule during a green key color and a different schedule correlated with a red key color that could be produced by a response to a changeover key. Once the changeover occurred, they remained in the presence of the red color for a predetermined period of time (e.g., 3 min). Because the immediate effect of a response to the changeover key was the onset of the red keylight, the frequency of such responses was presumably a function of the conditioned reinforcement value of that stimulus. The independent variables were the number and locations of the reinforcers during red. The frequency of changeover-key responding was found to be well predicted by the sum of the reciprocals of the delays to each individual reinforcer during red, with no evidence of any independent effect of rate of reinforcement itself. For example, a single reinforcer occurring 30 s after the choice response was more strongly preferred than two reinforcers occurring at 60 and 90 s, respectively. The reason for this can be seen readily by considering the values of the reciprocals of the delay values, which correspond to the immediacies of the reinforcers relative to stimulus onset. For the one reinforcer after 30 s, the reciprocal is simply $1/30 = 0.33$; for the combined effects of the reinforcers at 60 and 90 s, the sum of their immediacies is $.017 + .011 = .028$. The validity of summed immediacies of reinforcement as the critical determinant of stimulus value has now been supported by numerous studies using a variety of different procedures (e.g., Killeen, 1968; Mazur, 1986; Shull et al., 1990). Mazur (1993) has provided an interesting and readable discussion of this issue and its ramifications for understanding the effects of probability of reinforcement.

Numerous other variables play significant roles in determining the degree of conditioned value possessed by a conditioned reinforcer. We will not review them here (see Nevin, 1973, and Wike, 1966, for thorough reviews of the early literature), but instead will only note that all of the variables that have been implicated as determinants of the degree of Pavlovian conditioning affect the degree of conditioned value as well. Recent research has gone beyond the identification of the critical variables to focus more on their quantitative properties, with the result that various different theoretical models have been proposed (see Williams, 1988, and Davison & McCarthy, 1988, for reviews). Rather than consider this issue further, we now turn to the more mundane, but also important, concern of how conditioned reinforcement contingencies can be practically utilized.

PRACTICAL USES OF CONDITIONED REINFORCEMENT

The most well-known use of conditioned reinforcement contingencies is the token economy, in which poker chips or some comparable item are received by the subject for appropriate behavior of various types, and these tokens may then later be traded for commodities of value to the subject. There is no doubt that such procedures are highly useful in many different applied settings (see Kazdin, 1977, 1985, for reviews). It is also important to recognize, however, that the process by which token contingencies operate may or may not be interpreted in terms of reinforcement, because the use of verbal subjects allows myriad alternative interpretations. For example, token delivery may highlight the rules applicable to a behavior situation, thus increasing the role of rule-governed behavior (see Ferster, 1972, for a brief but insightful discussion of other possible effects of token contingencies). Perhaps because of the complexity of most applied settings, little research has been devoted to analyzing the actual causal mechanisms of how token contingencies facilitate performance.

Interpretation of the effects of token contingencies is often complex even when nonverbal subjects are involved. Perhaps the most widely cited investigations of conditioned reinforcement effects are the classic token-reinforcement studies of Wolfe (1936) and Cowles (1937), in which chimpanzees learned various tasks with poker chips as rewards. These tokens could then be inserted into a vending machine for various types of rewards, usually food. Behavior was maintained successfully with these token contingencies even when the exchange of the tokens for food was delayed until the end of the experimental session. However, subsequent work by Kelleher (1958) demonstrated that the vigor of behavior was strongly affected by the time between token delivery and when they were exchangeable. In general, when exchange was delayed until the end of the session, response rate was low at the start of the session, often for prolonged periods of time. For example, when tokens were earned on an FR 125 schedule and 50 tokens were required before the tokens could be exchanged for food, the chimpanzees typically paused for more than 2 hr at the start of each session (Kelleher, 1958). These pauses were shown to be under the discriminative control of the number of tokens in the animal's possession, in that delivery of a number of tokens at the start of the session independent of the animal's behavior resulted in vigorous behavior from the start of the session. The apparent reason for this is that the animal had learned that no food was available until a substantial number of tokens had been accumulated. Thus, rather than the tokens being a true substitute for food, their discriminative status as a cue for the availability of food determined their ability to maintain behavior. To what extent the reinforcing properties of tokens depend upon similar discriminative functions in human token economies is uncertain.

Quite apart from the substitutability of conditioned for primary reinforcers, it is worthwhile to consider the circumstances under which conditioned reinforcement contingencies have significant benefits over primary reinforcement contingencies. One general type of situation is when it is difficult to provide the primary reinforcer immediately contingent on the behavior, for any of a number of different reasons. For example, the primary reinforcer may disrupt the behavior because it is physically debilitating, as is the case with large dosages of some drugs (e.g., Katz, 1979). An interesting example of such usage has been reported by Everitt, Fray, Kostarczyk, Taylor, and Stacy (1987) in their investigation of the reinforcing properties of access to a receptive female rat for lever pressing by a male rat. Noting that the delivery of the primary reinforcer substantially disrupted lever pressing for substantial periods after its occurrence (because of the inherent difficulties of delivering small portions of the reward) and that previous attempts to condition operant behavior using sexual access as a reward had generally failed to maintain significant amounts of behavior, they presented their female rat contingent on bar pressing on an FI 15-min schedule, with presentations of the female at the end of the interval paired with a 30-s signal. In addition, brief presentations of the signal were contingent on lever pressing on an FR 10 schedule. Rate of lever pressing throughout the interval averaged 8 to 10 responses per minute. In the absence of the second-order schedule of conditioned reinforcement but with the FI contingency for access to the female still in effect, response rate decreased markedly and the pattern of behavior became erratic. Thus, conditioned reinforcement was an essential ingredient for maintaining sexually motivated behavior, allowing that behavior to be assessed under a variety of physiological and pharmacological variables.

Similar uses of conditioned reinforcement occur whenever the number of primary reinforcers per session must be limited, either because of satiation effects or because of expense. Consider, for example, the optimal procedure for producing the acquisition of a discrimination given that only 100 food pellets are allowed each day. A variety of data has shown that the rate of discrimination learning is a function of the number of stimulus–reinforcer (or response–reinforcer) pairings independent of the percentage of pairings, an effect that has been labeled *invariance in reinforcements to acquisition* (see Williams, 1989). This means that 200 trials with a 50% reinforcement schedule will produce the same amount of learning as 100 trials with a 100% schedule. However, when nonreinforced (by the primary reinforcer) correct trials are followed by a conditioned reinforcer, this invariance effect no longer holds. Instead, the conditioned reinforcer serves as a partial substitute for the primary reinforcer, so that the amount of learning from 200 trials with a 50% schedule of primary reinforcement in combination with a 100% conditioned reinforcement schedule will substantially exceed either of the other two conditions (Williams & Dunn, 1991b), despite the same number of primary reinforcers being delivered in all cases.

A second practical use of conditioned reinforcement occurs when the act of procuring the primary reinforcer interferes with the stream of behavior that is desired, as for example in animal acts such as those at Sea World in which the fish reward is provided on the side of the tank. The result is that some auditory signal (usually audible only to the whale or porpoise) is provided as feedback for the various segments of the performance, which allows the behavioral segments to be reinforced without the requirement that the subject procure the reward by approaching the food site.

A similar advantage for conditioned reinforcement contingencies in simple animal training was suggested by Skinner (1951). Suppose you wish to train your dog to roll over using a food reward. A typical problem with such attempts is that the dog will quickly discriminate that you are the source of the reward and will attend closely to you, often approaching and begging for the food at the expense of the behavior that you are attempting to shape. Although it is possible to extinguish this competing behavior, a more expeditious procedure is first to establish a discriminative contingency such that the reward is delivered only after the sound of a clicker, and then use the clicker as a conditioned reinforcer contingent on the behavior being shaped. The underlying principle is that Pavlovian contingencies may often compete with the operant contingencies, so that a judicious choice of the stimulus signaling reward availability is often necessary to minimize competition from the "sign tracking" that so often occurs when signals for food are physically localized in the environment.

A final example of the beneficial effects of conditioned reinforcement contingencies occurs when conditioned reinforcers produce faster learning than primary reinforcers because they induce lesser degrees of perseverative responding that are inconsistent with the discriminative contingencies. Luck, Colgrove, and Neuringer (1988) trained pigeons to generate fixed sequences of response on three keys (e.g., middle-left-right). One group received a 2-s food reward after each correct response except the terminal response, which produced a 6-s food reward. A second group received conditioned reinforcement from a 2-s presentation of an overhead houselight for each correct response plus the 6-s food reward for the terminal correct response. More accurate discrimination occurred with the houselight than with brief food rewards, and the primary basis of this difference was that the food rewards induced a higher rate of perseverative responding (e.g., responding to the middle key on the second position of the above sequence). Luck et al. noted that the superiority of conditioned reinforcement contingencies in similar conditional discrimination problems had been reported by several previous investigators, and that the perseverative effects of primary reinforcement should be considered to be another example of the negative side effects of reward that have been highlighted by Balsam and Bondy (1983). By reducing the interfering effects of such perseveration, conditioned reinforcement contingencies apparently allowed the conditional relations between current behavior and previous behavior to be discriminated more readily.

THEORETICAL UTILITY OF CONDITIONED REINFORCEMENT

In addition to the various procedures that have been considered above in which the concept of conditioned reinforcement has been studied explicitly, various other behavioral phenomena not directly linked to the concept may be elucidated by considering the role conditioned reinforcement plays in them. As illustrations, we will consider one laboratory phenomenon and one phenomenon from modern culture.

Contrafreeloading. An empirical phenomenon that generated considerable interest two decades ago is that animals will engage in operant behavior even when the food reinforcer contingent on the behavior is also freely available in the conditioning chamber itself (Jensen, 1963). Moreover, animals will learn, without prior shaping, to produce the operant response even when the free food is available from the onset of training (Neuringer, 1969). Such behavior has been observed with several different species, different schedules of reinforcement, and reinforcers other than food (see Osborne, 1977, for a review). The behavior appears paradoxical because it violates the law of least effort, which has been otherwise upheld in many different settings. This paradox disappears, however, once the role of conditioned reinforcement is appreciated. As shown by several different investigators (see Osborne, 1977), the critical ingredient appears to be the stimulus change contingent on the operant behavior that does not occur during the procurement of the free food itself. When the stimulus change is equated for the two sources of food, preference for free food quickly develops. Conditioned reinforcement, rather than stimulus change per se, appears to be essential, in that the operant behavior is not maintained when only stimulus change is contingent on the behavior; instead, the stimulus change must accompany the availability of the response-contingent primary reward. Osborne (1977) has also reviewed a variety of related phenomena for which conditioned reinforcement contingencies appear to be critical for proper understanding, including preference for signaled reinforcement, autoshaping, and self-reinforcement. Numerous other examples could be provided as well, including the much-neglected analysis of imprinting by Hoffman and Ratner (1973).

Implicit shaping contingencies. One of the most fascinating examples of the importance of conditioned reinforcement has been provided by Mowrer (1950, chap. 24) in his analysis of how myna birds acquire comprehensible vocal behavior. Mowrer argues that a prerequisite for vocal behavior is that the bird develop an attachment to its caretaker, so that attributes of the caretaker acquire positive value via their pairings with the primary rewards provided by the caretaker. One of these attributes is the sound of the caretaker's voice, including specific verbal utterances. The bird then increasingly approximates these sounds, because the similarity between the bird's own sounds and those of the caretaker cause the bird's own sounds to have positive value as well. Thus, the closer the bird approximates its caretaker's verbal utterances, the greater the degree of conditioned reinforcement contingent on vocalizing. Close matches will eventually occur because that is when the conditioned reinforcement is at a maximum.

Mowrer's analysis has strong similarities to current accounts of the acquisition of bird song, in that there is a consensus that the birds essentially shape themselves to sound increasingly like the models they have heard when young. Sounds different from the models are tried and rejected; those similar to the models receive increasing practice. The major theoretical difference concerns the learning histories necessary for the particular model's sound to have reinforcing properties. Whereas Mowrer's analysis followed a simple conditioning model, current accounts of bird-song acquisition have argued that only some models are possible to be learned, because the combination of genetic prewiring and experience allows only a limited range of auditory templates to be formed, and it is matching of the song to this auditory template that serves as the reinforcing agent for shaping the bird song later in life. The data relevant to this issue are complex and cannot be reviewed here. However, it is important to note that the social interactions between the young bird and the model have received increasing emphasis as a determinant of which bird songs will be acquired (e.g., Baptista & Petrinovich, 1984, 1986; Mann, Slater, Eales, & Richards, 1991).

To illustrate the explanatory power of Mowrer's (1950) analysis for domains other than bird vocalizations, consider the results of Neuringer and Neuringer (1974) on "social learning" in pigeons. Young pigeons were hand-fed by the experimenters, who subsequently placed the birds in an apparatus with a standard operant response key. Then, while the birds were closely watching the hand of the experimenter, he or she pressed the response key with his or her index finger. Subsequently the birds were allowed access to the response key, with the measure being how quickly they acquired the pecking response, relative to various control conditions. The observation by the young birds of the experimenter's finger touching the key significantly facilitated the acquisition of the response. The apparent interpretation of this finding is that the hand of the experimenter had acquired conditioned value via the history of hand feeding, and that this conditioned value was then transferred to the location of the key because of the spatial and temporal contiguity with the finger.

Mowrer's (1950) analysis can also be extended easily to a considerable amount of human behavior. The most obvious is the development of the phonetic structure of infant babbling, which evolves from including the entire spectrum of sound to being restricted to only those phonemes in the immediate linguistic community. A more far-reaching example is provided by the ubiquitous changes in the language patterns of teenagers. To the chagrin of many parents, an epidemic of speech insertions that serve no semantic function has developed: "you know," "like," and so forth. Much like venereal disease, these patterns of verbal behavior continue to spread among teenagers despite the disapproval and disdain of the adult linguistic community. Informal observations suggest that most teenagers are unaware of the nature of their verbal utterances on most occasions and will acknowledge their inappropriateness when queried. The concept of conditioned reinforcement offers a ready explanation for why such behavior is maintained. Teenagers are reinforced by assuming the characteristics of their peers, characteristics that include speech patterns as well as clothes and social customs. The robustness of the behavior in face of adult censure provides striking confirmation of the power of the conditioned reinforcement value of peer-modeled behavior.

If conditioned reinforcement is accepted as the explanation for why various types of behavior are imitated by children, the nature of that behavior provides an important perspective on the evolution of social values. The speech patterns of adults and teenagers are often in competition as sources of conditioned reinforcement for imitative behavior, and the fact that speech insertions have increased greatly in frequency implies that the conditioned value of imitating teenage peers is considerably stronger than that of imitating one's parents. This is hardly surprising, but it illustrates that the nature of imitative contingencies provides a window on changing patterns of human socialization.

The preceding example, although obviously conjectural, has strong ties to the views of the Hullian behavior theorists of the 1940s and 1950s (see introduction), who regarded the concept of conditioned reinforcement as a rich framework for a large segment of human behavior (also see Keller and Schoenfeld, 1950, for many other insightful examples of interpretation). Unfortunately, its role as an explanatory concept has considerably eroded over the past two decades. A major purpose of the present essay has been to reacquaint the reader with the critical evidence supporting the necessity of the concept and its ubiquitous influence on behavior in both laboratory and real-world situations. Despite its detractors, it remains one of the most fundamentally important psychological concepts.

REFERENCES

Amsel, A. (1992). *Frustration theory: An analysis of dispositional learning and memory.* Cambridge: Cambridge University Press.

Astley, S. L., & Perkins, C. C. (1985). Stimulus duration and conditioned reinforcing value measured by a learning-tests procedure. *Animal Learning & Behavior, 13,* 18–24.

Balsam, P., & Bondy, A. (1983). The negative side effects of reward. *Journal of Applied Behavior Analysis, 16,* 283–296.

Baptista, L. F., & Petrinovich, L. (1984). Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Animal Behaviour, 32,* 172–181.

Baptista, L. F., & Petrinovich, L. (1986). Song development in the white-crowned sparrow: Social factors and sex differences. *Animal Behaviour, 34,* 1359–1371.

Bugelski, R. (1938). Extinction with and without sub-goal reinforcement. *Journal of Comparative Psychology, 26,* 121–134.

Case, D. A., Fantino, E., & Wixted, J. (1985). Human observing: Maintained by negative informative stimuli only if correlated with improvement in response efficiency. *Journal of the Experimental Analysis of Behavior, 43,* 285–300.

Catania, A. C. (1992). *Learning* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Chance, P. (1994). *Learning & behavior* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Cohen, S. L., Calisto, G., & Lentz, B. E. (1979). Separating the reinforcing and discriminative properties of brief-stimulus presentations in second-order schedules. *Journal of the Experimental Analysis of Behavior, 32,* 149–156.

Cowles, J. T. (1937). Food tokens as incentives for learning by chimpanzees. *Comparative Psychology Monographs, 14,* 1–96.

Cronin, P. B. (1980). Reinstatement of postresponse stimuli prior to reward in delayed-reward discrimination learning by pigeons. *Animal Learning & Behavior, 8,* 352–358.

Daly, H. B. (1985). Observing response acquisition: Preference for unpredictable appetitive rewards obtained under conditions predicted by DMOD. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 294–316.

Daly, H. B. (1989). Preference for unpredictable food rewards occurs with high proportion of reinforced trials or alcohol if rewards are not delayed. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 3–13.

Davison, M. C., & McCarthy, D. (1988). *The matching law.* Hillsdale, NJ: Erlbaum.

Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences, 6,* 693–728.

Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy: An analysis in terms of learn-*

ing, thinking and culture. New York: McGraw-Hill.

Everitt, B. J., Fray, P., Kostarczyk, R., Taylor, S., & Stacy, P. (1987). Studies of instrumental behavior with sexual reinforcement in male rats (*Rattus norvegicus*): I. Control by brief visual stimuli paired with a receptive female. *Journal of Comparative Psychology, 101,* 395–406.

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., & Dunn, R. (1983). The delay-reduction hypothesis: Extension to three-alternative choice. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 132–146.

Ferster, C. B. (1972). The experimental analysis of clinical phenomena. *Psychological Record, 22,* 1–16.

Gibbon, J., & Balsam, P. (1981). Spreading association in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219–253). New York: Academic Press.

Gollub, L. R. (1958). *The chaining of fixed-interval schedules.* Unpublished doctoral dissertation, Harvard University.

Gollub, L. R. (1970). Information on conditioned reinforcement: A review of *Conditioned Reinforcement* edited by D. Hendry. *Journal of the Experimental Analysis of Behavior, 14,* 361–372.

Gollub, L. (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.

Grice, G. R. (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *Journal of Experimental Psychology, 38,* 1–16.

Herrnstein, R. J., & Loveland, D. H. (1972). Food-avoidance in hungry pigeons and other perplexities. *Journal of the Experimental Analysis of Behavior, 18,* 369–383.

Hoffman, H. S., & Ratner, A. M. (1973). A reinforcement model of imprinting: Implications for socialization in monkeys and man. *Psychological Review, 80,* 527–546.

Honey, R. C., Schachtman, T. R., & Hall, G. (1987). Partial reinforcement in serial autoshaping: The role of attentional and associative factors. *Learning and Motivation, 18,* 288–300.

Hyde, T. S. (1976). The effect of Pavlovian stimuli in the acquisition of a new response. *Learning and Motivation, 7,* 223–239.

Jensen, G. D. (1963). Preference for bar pressing over "freeloading" as a function of number of rewarded presses. *Journal of Experimental Psychology, 65,* 451–454.

Katz, J. L. (1979). A comparison of responding maintained under second-order schedules of intramuscular cocaine injection or food presentation in squirrel monkeys. *Journal of the Experimental Analysis of Behavior, 32,* 419–431.

Kazdin, A. E. (1977). *The token economy: A review and evaluation.* New York: Plenum.

Kazdin, A. E. (1985). The token economy. In R. M. Turner & L. M. Ascher (Eds.), *Evaluating behavior therapy outcome.* New York: Springer.

Kelleher, R. T. (1958). Fixed-ratio schedules of conditioned reinforcement with chimpanzees. *Journal of the Experimental Analysis of Behavior, 1,* 281–289.

Kelleher, R. T. (1966). Conditioned reinforcement in second-order schedules. *Journal of the Experimental Analysis of Behavior, 9,* 475–485.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology.* New York: Appleton-Century-Crofts.

Killeen, P. (1968). On the measure of reinforcement frequency in the study of preference. *Journal of the Experimental Analysis of Behavior, 11,* 263–269.

Killeen, P. R. (1981). Learning as causal inference. In M. L. Commons & J. A. Nevin (Eds.), *Quantitative analyses of behavior: Discriminative properties of reinforcement schedules* (pp. 89–112). Cambridge, MA: Ballinger.

Kimble, G. A. (1961). *Hilgard and Marquis' conditioning and learning.* Englewood Cliffs, NJ: Prentice-Hall.

Klein, R. M. (1959). Intermittent primary reinforcement as a parameter of secondary reinforcement. *Journal of Experimental Psychology, 58,* 423–427.

Levis, D. J. (1989). The case for a return to a two-factor theory of avoidance: The failure of non-fear interpretations. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 227–277). Hillsdale, NJ: Erlbaum.

Lieberman, D. A., Davidson, F. H., & Thomas, G. V. (1985). Marking in pigeons: The role of memory in delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 611–624.

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.

Luck, S., Colgrove, M., & Neuringer, A. (1988). Response sequence learning as a function of primary versus conditioned reinforcement. *Animal Learning & Behavior, 16,* 8–14.

Mann, N. I., Slater, P. J. B., Eales, L. A., & Richards, C. (1991). The influence of visual stimuli on song-tutor choice in the zebra finch, *Taeniopygia guttata. Animal Behaviour, 42,* 285–293.

Martin, G., & Pear, J. (1992). *Behavior modification: What it is and how to do it* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Mazur, J. E. (1986). Choice between single and multiple delayed reinforcers. *Journal of the Experimental Analysis of Behavior, 46,* 67–77.

Mazur, J. E. (1993). Predicting the strength of a conditioned reinforcer: Effects of delay and uncertainty. *Current Directions in Psychological Science, 2,* 70–74.

Mowrer, O. H. (1950). *Learning theory and personality dynamics.* New York: The Ronald Press.

Mowrer, O. H. (1960). *Learning theory and the symbolic processes.* New York: Wiley.

Neuringer, A. J. (1969). Animals respond for food in the presence of free food. *Science, 166,* 399–401.

Neuringer, A., & Nueringer, M. (1974). Learning by following a food source. *Science, 184,* 1005–1008.

Nevin, J. A. (1973). Conditioned reinforcement. In J. A. Nevin & G. S. Reynolds (Eds.), *The study of behavior* (pp. 154–198). Glenview, IL: Scott-Foresman.

Osborne, S. R. (1977). The free food (contrafreeloading) phenomenon: A review and analysis. *Animal Learning & Behavior, 5,* 221–235.

Pavlov, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). London: Oxford University Press.

Preston, R. A., & Fantino, E. (1991). Conditioned reinforcement value and choice. *Journal of the Experimental Analysis of Behavior, 55,* 155–175.

Rachlin, H. (1976). *Behavior and learning.* San Francisco: Freeman.

Rachlin, H. C., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences, 4,* 371–417.

Rashotte, M. E., Marshall, B. S., & O'Connell, J. M. (1981). Signaling functions of the second-order CS: Partial reinforcement during second-order conditioning of the pigeon's keypeck. *Animal Learning & Behavior, 9,* 253–260.

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.

Richards, R. W. (1981). A comparison of signaled and unsignaled delay of reinforcement. *Journal of the Experimental Analysis of Behavior, 35,* 145–152.

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.

Schaal, D. W., & Branch, M. N. (1988). Responding of pigeons under variable-interval schedules of unsignaled, briefly signaled, and completely signaled delays to reinforcement. *Journal of the Experimental Analysis of Behavior, 50,* 33–54.

Schaal, D. W., & Branch, M. N. (1990). Responding of pigeons under variable-interval schedules of signaled-delayed reinforcement: Effects of delay-signal duration. *Journal of the Experimental Analysis of Behavior, 53,* 103–121.

Schuster, R. H. (1969). A functional analysis of conditioned reinforcement. In D. Hendry (Ed.), *Conditioned reinforcement* (pp. 192–234). Homewood, IL: Dorsey Press.

Shull, R. L., Mellon, R. C., & Sharp, J. A. (1990). Delay and number of food reinforcers: Effect on choice and latencies. *Journal of the Experimental Analysis of Behavior, 53,* 235–246.

Shull, R. L., Spear, D. J., & Bryson, A. E. (1981). Delay or rate of food delivery as a determiner of response rate. *Journal of the Experimental Analysis of Behavior, 35,* 129–143.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1951). How to teach animals. *Scientific American, 185,* 26–29.

Skinner, B. F. (1953). *Science and human behavior.* New York: Free Press.

Staddon, J. E. R. (1983). *Adaptive learning and behavior.* Cambridge: Cambridge University Press.

Staddon, J. E. R., & Ettinger, R. H. (1989). *Learning: An introduction to the principles of adaptive behavior.* San Diego: Harcourt Brace Jovanovich.

Stubbs, A. (1971). Second-order schedules and the problem of conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 16,* 289–313.

Thomas, G. V., Robertson, D., & Cunniffe, G. (1989). Facilitation of responding in a filled-delay trace autoshaping procedure: An occasion-setting hypothesis. *Learning and Motivation, 20,* 225–241.

Wike, E. L. (1966). *Secondary reinforcement: Selected experiments.* New York: Harper & Row.

Williams, B. A. (1976). The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 26,* 441–449.

Williams, B. A. (1988). Reinforcement, choice, and response strength. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., pp. 167–244). New York: Wiley.

Williams, B. A. (1989). Partial reinforcement effects on discrimination learning. *Animal Learning & Behavior, 17,* 418–432.

Williams, B. A. (1991). Marking and bridging versus conditioned reinforcement. *Animal Learning & Behavior, 19,* 264–269.

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., & Royalty, P. (1990). Conditioned reinforcement versus time to reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior, 53,* 381–393.

Wolfe, J. B. (1936). Effectiveness of token rewards for chimpanzees. *Comparative Psychology Monographs, 12,* 1–72, No. 60.

Wyckoff, L. B., Jr. (1952). The role of observing responses in discrimination learning. *Psychological Review, 59,* 431–442.

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.