Psychological Review
2000, Vol. 107, No. 2, 289–344

0033-295X/00/$5.00 DOI: 10.1037//0033-295X.107.2.289

Time, Rate, and Conditioning

C. R. Gallistel
University of California, Los Angeles

John Gibbon
New York State Psychiatric Institute and Columbia University

The authors draw together and develop previous timing models for a broad range of conditioning phenomena to reveal their common conceptual foundations: First, conditioning depends on the learning of the temporal intervals between events and the reciprocals of these intervals, the rates of event occurrence. Second, remembered intervals and rates translate into observed behavior through decision processes whose structure is adapted to noise in the decision variables. The noise and the uncertainties consequent on it have both subjective and objective origins. A third feature of these models is their timescale invariance, which the authors argue is a very important property evident in the available experimental data. This conceptual framework is similar to the psychophysical conceptual framework in which contemporary models of sensory processing are rooted. The authors contrast it with the associative conceptual framework.

Pavlov (1928) recognized that the timing of the conditioned response (CR; e.g., salivation) in a well-conditioned subject depended on the reinforcement delay, or latency. The longer the interval was between the onset of the conditioned stimulus (CS; e.g., the ringing of a bell) and the delivery of the unconditioned stimulus (US; e.g., meat powder), the longer the latency was between CS onset and the onset of salivation. An obvious explanation is that the dogs in Pavlov's experiment learned the reinforcement latency and did not begin to salivate until they judged that the delivery of food was more or less imminent. This is not the kind of explanation that Pavlov favored because it lacks a clear basis in reflex physiology. Similarly, Skinner (1938) observed that the timing of operant responses was governed by the intervals in the schedule of reinforcement. When his pigeons pecked keys to obtain reinforcement on fixed-interval (FI) schedules, the longer the fixed interval imposed between the obtaining of one reinforcement and the availability of the next, the longer the pigeons waited after each reinforcement before beginning to peck the key to obtain the next reinforcement. An obvious explanation is that the pigeons learned the duration of the interval between a delivered reinforcement and the next arming of the key and did not begin to peck until they judged that the opportunity to obtain another reinforcement was more or less imminent. Again, this is not the sort of explanation that Skinner favored, although for reasons different than Pavlov's.

In this article, we take the interval-learning assumption as the point of departure in the analysis of conditioned behavior. We assume that the subjects in conditioning experiments do, in fact, store in memory the durations of interevent intervals and subsequently recall those remembered durations for use in the decisions that determine their conditioned behavior. An extensive experimental literature on timed behavior has developed in the past few decades (for reviews, see Fantino, Preston, & Dunn, 1993; Gallistel, 1989; Gibbon & Allan, 1984; Gibbon, Malapani, Dale, & Gallistel, 1997; Killeen & Fetterman, 1988; Miller & Barnet, 1993; Staddon & Higa, 1993). Consequently, it is now widely accepted that the subjects in conditioning experiments do in some sense learn the intervals in the experimental protocols. But those aspects of conditioned behavior that seem to depend on knowledge of the temporal intervals are often seen as adjunctive to the process of association formation (e.g., Miller & Barnet, 1993), which is commonly assumed to be the core process mediating conditioned behavior. We argue that it is the learning of temporal intervals and their reciprocals (event rates) that is the core process in both Pavlovian and instrumental conditioning.

It is our sense that most contemporary associative theorists no longer assume that the association-forming process itself is fundamentally different in Pavlovian and instrumental conditioning. Until the discovery of autoshaping (Brown & Jenkins, 1968), a now widely used Pavlovian procedure for teaching what used to be regarded as instrumental responses (pecking a key or pressing a lever for food), it was assumed that there were two fundamentally different association-forming processes. One, which operated in Pavlovian conditioning, required only the temporal contiguity of a

---

C. R. Gallistel, Department of Psychology, University of California, Los Angeles; John Gibbon, Division of Biopsychology, New York State Psychiatric Institute, and Department of Psychology, Columbia University.

We gratefully acknowledge support from National Institutes of Health Award MH41649 and National Science Foundation Award IBN-9306283.

We are grateful to our many colleagues who read and critiqued parts of earlier versions of this article. We particularly thank Gustavo Stolovitzky for helping in the derivation of acquisition variability, Peter Dayan for calling attention to a serious error in the way that rate estimation theory computed the odds that two rates differed, and Stephen Fairhurst for showing that taking variability into account leads to the prediction that the slope of the function relating reinforcements to acquisition to the ratio of intertrial duration to trial duration is less than one. We also thank Ralph Miller and three self-identified reviewers, Russ Church, Peter Killeen, and John Pearce, for their thorough and helpful critiques of earlier versions of this article.

Correspondence concerning this article should be addressed to C. R. Gallistel, Department of Psychology, University of California, Los Angeles, Box 951563, Los Angeles, California 90095-1563. Electronic mail may be sent to randy@psych.ucla.edu.

CS and a US. The other, which operated in instrumental conditioning, required that a reinforcer stamp in the latent association between a stimulus and a response. In this older conception of the association-forming process in instrumental conditioning, the reinforcer was not itself part of the associative structure; it merely stamped in the stimulus–response association. Recently, however, it has been shown that Pavlovian response–outcome or outcome–response associations are important in instrumental conditioning (Adams & Dickinson, 1981; Colwill & Delamater, 1995; Colwill & Rescorla, 1986; Mackintosh & Dickinson, 1979; Rescorla, 1991). Our reading of the most recent trends in associative theorizing is that, for these and other reasons (e.g., Williams, 1982), the two kinds of conditioning paradigms are no longer thought to tap fundamentally different association-forming processes. Rather, they are thought to give rise to different associative structures through a single association-forming process.

In any event, the underlying learning processes in the two kinds of paradigms are not fundamentally different from the perspective of timing theory. The paradigms differ only in the kinds of events that mark the start of relevant temporal intervals or alter the expected intervals between reinforcements. In Pavlovian paradigms, the animal's behavior has no effect on the delivery of reinforcement. The conditioned behavior is determined by the rate and timing of reinforcement when a CS is present relative to the rate and timing of reinforcement when that CS is not present. In instrumental paradigms, the animal's behavior alters the rate and timing of reinforcement. Reinforcements occur at a higher rate when the animal pecks the key (or presses the lever, etc.) than when it does not. And the time of delivery of the next reinforcement may depend on the interval since a response-elicited event such as the previous reinforcement. In both cases, the essential underlying process from a timing perspective is the learning of the contingency between the rate of reinforcement (or expected interval between reinforcements) and some state of affairs (bell ringing vs. bell not ringing, key being pecked vs. key not being pecked, or rapid key pecking vs. slow key pecking). Thus, we do not treat these conditioning paradigms separately. We move back and forth between them.

We develop our argument around models that we ourselves have elaborated because we are more intimately familiar with them. We emphasize, however, that there are several other timing models (e.g., Church & Broadbent, 1990; Fantino et al., 1993; Grossberg & Schmajuk, 1991; Killeen & Fetterman, 1988; Miller & Barnet, 1993; Staddon & Higa, 1993). We do not imagine our own models to be the last word. In fact, we call attention at several points to difficulties and lacunae in these models. Our goal is to make clear essential features of a conceptual framework that differs quite fundamentally from the framework in which conditioning is most commonly analyzed. We expect that as this framework becomes more widely used, the models rooted in it will become more sophisticated, more complete, and ever broader in scope.

We also use the framework to call attention to quantitative features of conditioning data that we believe have far-reaching theoretical implications, most notably the many manifestations of timescale invariance. A conditioning result is timescale-invariant if the graph of the result looks the same when the experiment is repeated at a different timescale, by changing all the temporal intervals in the protocol by a common scaling factor, and the scaling factors on the data graphs are adjusted so as to offset the change in timescale. Somewhat more technically, conditioning data are timescale-invariant if the normalized data plots are superimposable. Normalization takes out the timescale. Superimposability means that the normalized curves (or, in the limit, individual points) fall on top of each other. We give several examples in this article, beginning with Figures 1A and 2. An extremely important empirical consequence of the new conceptual framework is that it stimulates research to test the limits of fundamentally important principles such as this.

Conditioned-Response Timing

The learning of temporal intervals in the course of conditioning is most directly evident in the timing of the CR in protocols in which reinforcement occurs at some fixed delay after a marking event. In what follows, this delay is called the reinforcement latency, *T*.

Some well-established facts of CR timing are as follows:

- The CR is maximally likely at the reinforcement latency: When there is a fixed latency between a marking event (e.g., placement in the experimental chamber, the delivery of a previous reinforcement, the sounding of a tone, the extension of a lever, or the illumination of a response key), then the probability that a

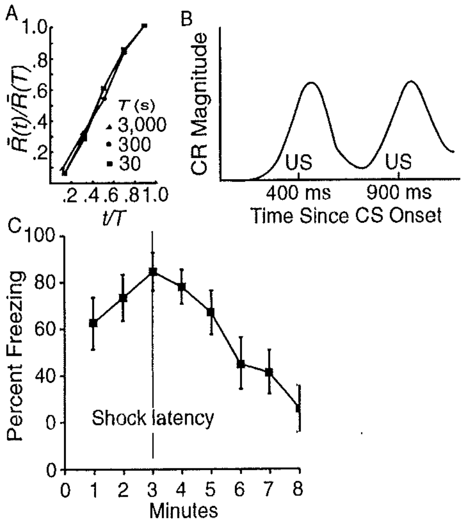

Figure 1. A: Normalized rate of responding as a function of the normalized elapsed interval, for pigeons responding on fixed-interval schedules, with interreinforcement intervals (*T*) ranging from 30 to 3,000 s. $\bar{R}(t)$ is the average rate of responding at elapsed interval *t* since the last reinforcement. $\bar{R}(T)$ is the average terminal rate of responding. The data are from Dews (1970). The figure is from "Scalar Expectancy Theory and Weber's Law in Animal Timing," by J. Gibbon, 1977, *Psychological Review, 84,* p. 280. Copyright 1977 by the American Psychological Association. Reprinted with permission. B: The time course of the conditioned double blink on a single representative trial in an experiment in which rabbits were trained with two different unconditioned-stimulus (US) latencies (400 and 900 ms). The data are from Kehoe, Graham-Clarke, and Schreurs (1989). CR = conditioned response; CS = conditioned stimulus. C: Percent of subjects freezing as a function of the interval since placement in the experimental chamber after a single conditioning trial in which rats were shocked 3 min after being placed in the chamber. Vertical bars represent ±1 *SE*. The data are from Fanselow and Stote (1995).

Psychological Review
2000, Vol. 107, No. 2, 289–344

0033-295X/00/$5.00 DOI: 10.1037//0033-295X.107.2.289

Time, Rate, and Conditioning

C. R. Gallistel
University of California, Los Angeles

John Gibbon
New York State Psychiatric Institute and Columbia University

The authors draw together and develop previous timing models for a broad range of conditioning phenomena to reveal their common conceptual foundations: First, conditioning depends on the learning of the temporal intervals between events and the reciprocals of these intervals, the rates of event occurrence. Second, remembered intervals and rates translate into observed behavior through decision processes whose structure is adapted to noise in the decision variables. The noise and the uncertainties consequent on it have both subjective and objective origins. A third feature of these models is their timescale invariance, which the authors argue is a very important property evident in the available experimental data. This conceptual framework is similar to the psychophysical conceptual framework in which contemporary models of sensory processing are rooted. The authors contrast it with the associative conceptual framework.

Pavlov (1928) recognized that the timing of the conditioned response (CR; e.g., salivation) in a well-conditioned subject depended on the reinforcement delay, or latency. The longer the interval was between the onset of the conditioned stimulus (CS; e.g., the ringing of a bell) and the delivery of the unconditioned stimulus (US; e.g., meat powder), the longer the latency was between CS onset and the onset of salivation. An obvious explanation is that the dogs in Pavlov's experiment learned the reinforcement latency and did not begin to salivate until they judged that the delivery of food was more or less imminent. This is not the kind of explanation that Pavlov favored because it lacks a clear basis in reflex physiology. Similarly, Skinner (1938) observed that the timing of operant responses was governed by the intervals in the schedule of reinforcement. When his pigeons pecked keys to obtain reinforcement on fixed-interval (FI) schedules, the longer the fixed interval imposed between the obtaining of one reinforcement and the availability of the next, the longer the pigeons waited after each reinforcement before beginning to peck the key to obtain the next reinforcement. An obvious explanation is that the pigeons learned the duration of the interval between a delivered reinforcement and the next arming of the key and did not begin to peck until they judged that the opportunity to obtain another reinforcement was more or less imminent. Again, this is not the sort of explanation that Skinner favored, although for reasons different than Pavlov's.

In this article, we take the interval-learning assumption as the point of departure in the analysis of conditioned behavior. We assume that the subjects in conditioning experiments do, in fact, store in memory the durations of interevent intervals and subsequently recall those remembered durations for use in the decisions that determine their conditioned behavior. An extensive experimental literature on timed behavior has developed in the past few decades (for reviews, see Fantino, Preston, & Dunn, 1993; Gallistel, 1989; Gibbon & Allan, 1984; Gibbon, Malapani, Dale, & Gallistel, 1997; Killeen & Fetterman, 1988; Miller & Barnet, 1993; Staddon & Higa, 1993). Consequently, it is now widely accepted that the subjects in conditioning experiments do in some sense learn the intervals in the experimental protocols. But those aspects of conditioned behavior that seem to depend on knowledge of the temporal intervals are often seen as adjunctive to the process of association formation (e.g., Miller & Barnet, 1993), which is commonly assumed to be the core process mediating conditioned behavior. We argue that it is the learning of temporal intervals and their reciprocals (event rates) that is the core process in both Pavlovian and instrumental conditioning.

It is our sense that most contemporary associative theorists no longer assume that the association-forming process itself is fundamentally different in Pavlovian and instrumental conditioning. Until the discovery of autoshaping (Brown & Jenkins, 1968), a now widely used Pavlovian procedure for teaching what used to be regarded as instrumental responses (pecking a key or pressing a lever for food), it was assumed that there were two fundamentally different association-forming processes. One, which operated in Pavlovian conditioning, required only the temporal contiguity of a

---

C. R. Gallistel, Department of Psychology, University of California, Los Angeles; John Gibbon, Division of Biopsychology, New York State Psychiatric Institute, and Department of Psychology, Columbia University.

We gratefully acknowledge support from National Institutes of Health Award MH41649 and National Science Foundation Award IBN-9306283.

We are grateful to our many colleagues who read and critiqued parts of earlier versions of this article. We particularly thank Gustavo Stolovitzky for helping in the derivation of acquisition variability, Peter Dayan for calling attention to a serious error in the way that rate estimation theory computed the odds that two rates differed, and Stephen Fairhurst for showing that taking variability into account leads to the prediction that the slope of the function relating reinforcements to acquisition to the ratio of intertrial duration to trial duration is less than one. We also thank Ralph Miller and three self-identified reviewers, Russ Church, Peter Killeen, and John Pearce, for their thorough and helpful critiques of earlier versions of this article.

Correspondence concerning this article should be addressed to C. R. Gallistel, Department of Psychology, University of California, Los Angeles, Box 951563, Los Angeles, California 90095-1563. Electronic mail may be sent to randy@psych.ucla.edu.

CS and a US. The other, which operated in instrumental conditioning, required that a reinforcer stamp in the latent association between a stimulus and a response. In this older conception of the association-forming process in instrumental conditioning, the reinforcer was not itself part of the associative structure; it merely stamped in the stimulus–response association. Recently, however, it has been shown that Pavlovian response–outcome or outcome–response associations are important in instrumental conditioning (Adams & Dickinson, 1981; Colwill & Delamater, 1995; Colwill & Rescorla, 1986; Mackintosh & Dickinson, 1979; Rescorla, 1991). Our reading of the most recent trends in associative theorizing is that, for these and other reasons (e.g., Williams, 1982), the two kinds of conditioning paradigms are no longer thought to tap fundamentally different association-forming processes. Rather, they are thought to give rise to different associative structures through a single association-forming process.

In any event, the underlying learning processes in the two kinds of paradigms are not fundamentally different from the perspective of timing theory. The paradigms differ only in the kinds of events that mark the start of relevant temporal intervals or alter the expected intervals between reinforcements. In Pavlovian paradigms, the animal's behavior has no effect on the delivery of reinforcement. The conditioned behavior is determined by the rate and timing of reinforcement when a CS is present relative to the rate and timing of reinforcement when that CS is not present. In instrumental paradigms, the animal's behavior alters the rate and timing of reinforcement. Reinforcements occur at a higher rate when the animal pecks the key (or presses the lever, etc.) than when it does not. And the time of delivery of the next reinforcement may depend on the interval since a response-elicited event such as the previous reinforcement. In both cases, the essential underlying process from a timing perspective is the learning of the contingency between the rate of reinforcement (or expected interval between reinforcements) and some state of affairs (bell ringing vs. bell not ringing, key being pecked vs. key not being pecked, or rapid key pecking vs. slow key pecking). Thus, we do not treat these conditioning paradigms separately. We move back and forth between them.

We develop our argument around models that we ourselves have elaborated because we are more intimately familiar with them. We emphasize, however, that there are several other timing models (e.g., Church & Broadbent, 1990; Fantino et al., 1993; Grossberg & Schmajuk, 1991; Killeen & Fetterman, 1988; Miller & Barnet, 1993; Staddon & Higa, 1993). We do not imagine our own models to be the last word. In fact, we call attention at several points to difficulties and lacunae in these models. Our goal is to make clear essential features of a conceptual framework that differs quite fundamentally from the framework in which conditioning is most commonly analyzed. We expect that as this framework becomes more widely used, the models rooted in it will become more sophisticated, more complete, and ever broader in scope.

We also use the framework to call attention to quantitative features of conditioning data that we believe have far-reaching theoretical implications, most notably the many manifestations of timescale invariance. A conditioning result is timescale-invariant if the graph of the result looks the same when the experiment is repeated at a different timescale, by changing all the temporal intervals in the protocol by a common scaling factor, and the scaling factors on the data graphs are adjusted so as to offset the change in timescale. Somewhat more technically, conditioning data are timescale-invariant if the normalized data plots are superimposable. Normalization takes out the timescale. Superimposability means that the normalized curves (or, in the limit, individual points) fall on top of each other. We give several examples in this article, beginning with Figures 1A and 2. An extremely important empirical consequence of the new conceptual framework is that it stimulates research to test the limits of fundamentally important principles such as this.

Conditioned-Response Timing

The learning of temporal intervals in the course of conditioning is most directly evident in the timing of the CR in protocols in which reinforcement occurs at some fixed delay after a marking event. In what follows, this delay is called the reinforcement latency, *T*.

Some well-established facts of CR timing are as follows:

- The CR is maximally likely at the reinforcement latency: When there is a fixed latency between a marking event (e.g., placement in the experimental chamber, the delivery of a previous reinforcement, the sounding of a tone, the extension of a lever, or the illumination of a response key), then the probability that a

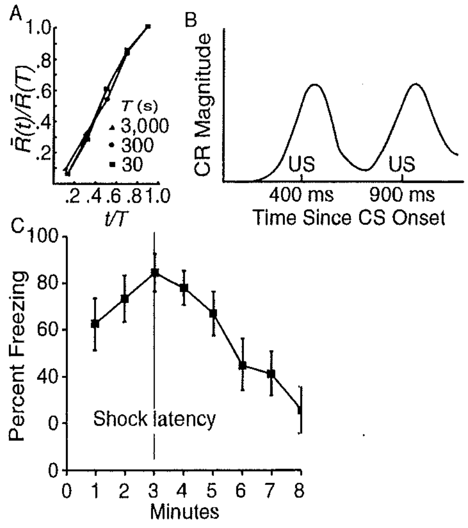

Figure 1. A: Normalized rate of responding as a function of the normalized elapsed interval, for pigeons responding on fixed-interval schedules, with interreinforcement intervals (*T*) ranging from 30 to 3,000 s. $\bar{R}(t)$ is the average rate of responding at elapsed interval *t* since the last reinforcement. $\bar{R}(T)$ is the average terminal rate of responding. The data are from Dews (1970). The figure is from "Scalar Expectancy Theory and Weber's Law in Animal Timing," by J. Gibbon, 1977, *Psychological Review, 84,* p. 280. Copyright 1977 by the American Psychological Association. Reprinted with permission. B: The time course of the conditioned double blink on a single representative trial in an experiment in which rabbits were trained with two different unconditioned-stimulus (US) latencies (400 and 900 ms). The data are from Kehoe, Graham-Clarke, and Schreurs (1989). CR = conditioned response; CS = conditioned stimulus. C: Percent of subjects freezing as a function of the interval since placement in the experimental chamber after a single conditioning trial in which rats were shocked 3 min after being placed in the chamber. Vertical bars represent ±1 *SE.* The data are from Fanselow and Stote (1995).

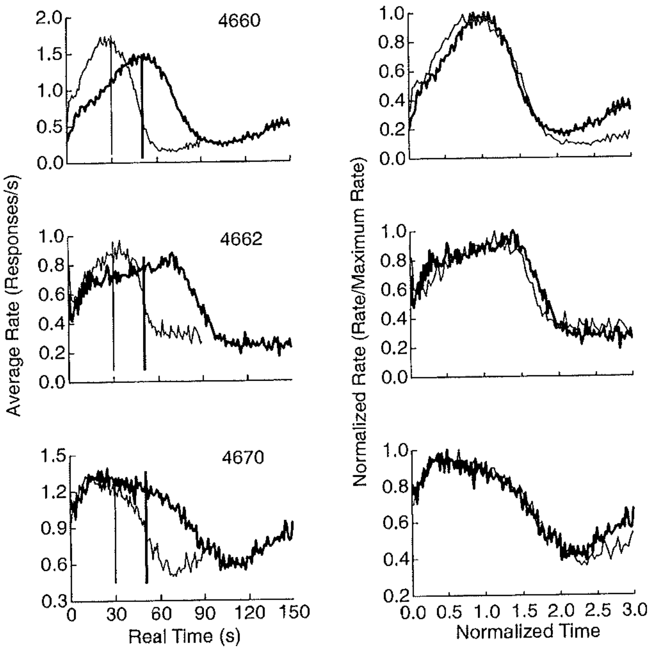

Figure 2. Scalar property: timescale invariance in the distribution of conditioned responses. The left panels show responding of 3 birds (4660, 4662, and 4670) on the peak procedure in blocked sessions at reinforcement latencies of 30 and 50 s (unreinforced conditioned-stimulus [CS] durations of 90 and 150 s, respectively). Vertical bars at the reinforcement latencies have heights equal to the peaks of the corresponding distributions. The right panels show the same functions normalized with respect to CS time and peak rate (so that vertical bars would superimpose). Note that although the distributions differ between birds, both in their shape and in whether they peak before or after the reinforcement latency (K^* error), they superimpose when normalized (rescaled). The data are replotted from data originally reported in Gibbon, Fairhurst, and Goldberg (1997).

well-trained subject will make a CR increases as the time of reinforcement approaches, reaching a maximum at the reinforcement latency (Figures 1 and 2).

- The distribution of CR onsets and offsets is scalar: There is a constant coefficient of variation in the distribution of response probability around the latency of peak probability; that is, the standard deviation of the distribution is proportionate to its mode. Thus, the temporal distribution of CR initiations (and terminations) is timescale-invariant: Scaling time in units proportional to the mode of the distribution renders the distributions obtained at different reinforcement latencies superimposable (see Figures 1A and 2).

Scalar Expectancy Theory

Scalar expectancy theory (SET) was developed to account for the aforementioned aspects of the CR (Gibbon, 1977). It is a model of what we call the "when decision," the decision that determines when the CR occurs in relation to a time mark such as CS onset or offset or the delivery of a previous reinforcement. The basic assumptions of SET and the components from which the model is constructed—a timing mechanism, a memory mechanism, sources of variability or noise in the decision variables, and a comparison mechanism adapted to that noise (see Figure 3)—appear in our explanation of all other aspects of conditioned behavior. The timing mechanism generates a signal, $\hat{t}_e$, which is proportional at every moment to the elapsed duration of the animal's current exposure to a CS. This quantity in the head is the animal's measure of the duration of an elapsing interval. The timer is reset to zero by the occurrence of a reinforcement, which marks the end of the interval that began with the onset of the CS. The magnitude of $\hat{t}_e$ at the time of reinforcement, $\hat{t}_T$, is written to memory through a multiplicative translation variable, k^*, whose expected value $[E(k^*) = K^*]$ is close to but not identically one. Thus, the reinforcement interval recorded in memory, $\hat{t}^* = k^*\hat{t}_T$, on average deviates from the timed value by some (generally small) percentage, which is determined by the extent to which the expected value (K^*) deviates from one. (See Table 1 for a list of the symbols and expressions used, together with their meanings.)

When the CS reappears (when a new trial begins), $\hat{t}_e$, the subjective duration of the currently elapsing interval of CS exposure, is compared with $\hat{t}^*$, which is derived by sampling (reading) the remembered reinforcement delay in memory. The comparison takes the form of a ratio, $\hat{t}_e/\hat{t}^*$, which we call the "decision variable." When this ratio exceeds a threshold, β, somewhat less than one, the animal responds to the CS—provided it has had sufficient experience with the CS to have already decided that it is a reliable predictor of the US (see the Acquisition section below).[1] The *when* decision threshold is somewhat less than one because the CR anticipates the US. If, on a given trial, reinforcement does not occur (e.g., in *The peak procedure* section below), then the CR ceases when this same decision ratio exceeds a second threshold somewhat greater than one. (The decision to stop responding when the reinforcement interval is past is not diagrammed in Figure 3, but see Gibbon and Church [1990].) In short, the animal begins to respond when it estimates the currently elapsing interval to be close to the remembered delay of reinforcement. If it does not get reinforced, it stops responding when it estimates the currently elapsing interval to be sufficiently past the remembered delay. The decision thresholds constitute the animal's criteria for "close" and "past." Its measure of closeness (or similarity) is the ratio between the currently elapsing interval and the remembered interval.

The interval timer in SET may be conceived as a clock system (pulse generator) feeding an accumulator (working memory), which continually integrates activity over time. The essential feature of such a mechanism is that the quantity in the accumulator grows as a linear function of time. By contrast, the reference memory system statically preserves the values of past intervals.

---

[1] The decision variable is formally a ratio of random variables and is demonstrably nonnormal in most cases. However, the decision rule $t_e/t^* > \beta$ is equivalent to $t_e > \beta t^*$, and the right-hand side of this inequality is approximately normal when t^* is normal. When the threshold β is itself variable, some nonnormality is induced in the right-hand side of the decision rule, introducing some positive skew in this composite variable. Gibbon and his collaborators (Gibbon, 1981b, 1992; Gibbon, Church, & Meck, 1984) have discussed the degree of nonnormality in this variate in considerable detail. It is shown that (a) the mean and variance of the decision variate are readily obtained in closed form and (b) the degree of skew in the composite variable is not large relative to other variance in the system. The behavioral performance often also shows a slight positive skew consistent with the formal analysis.

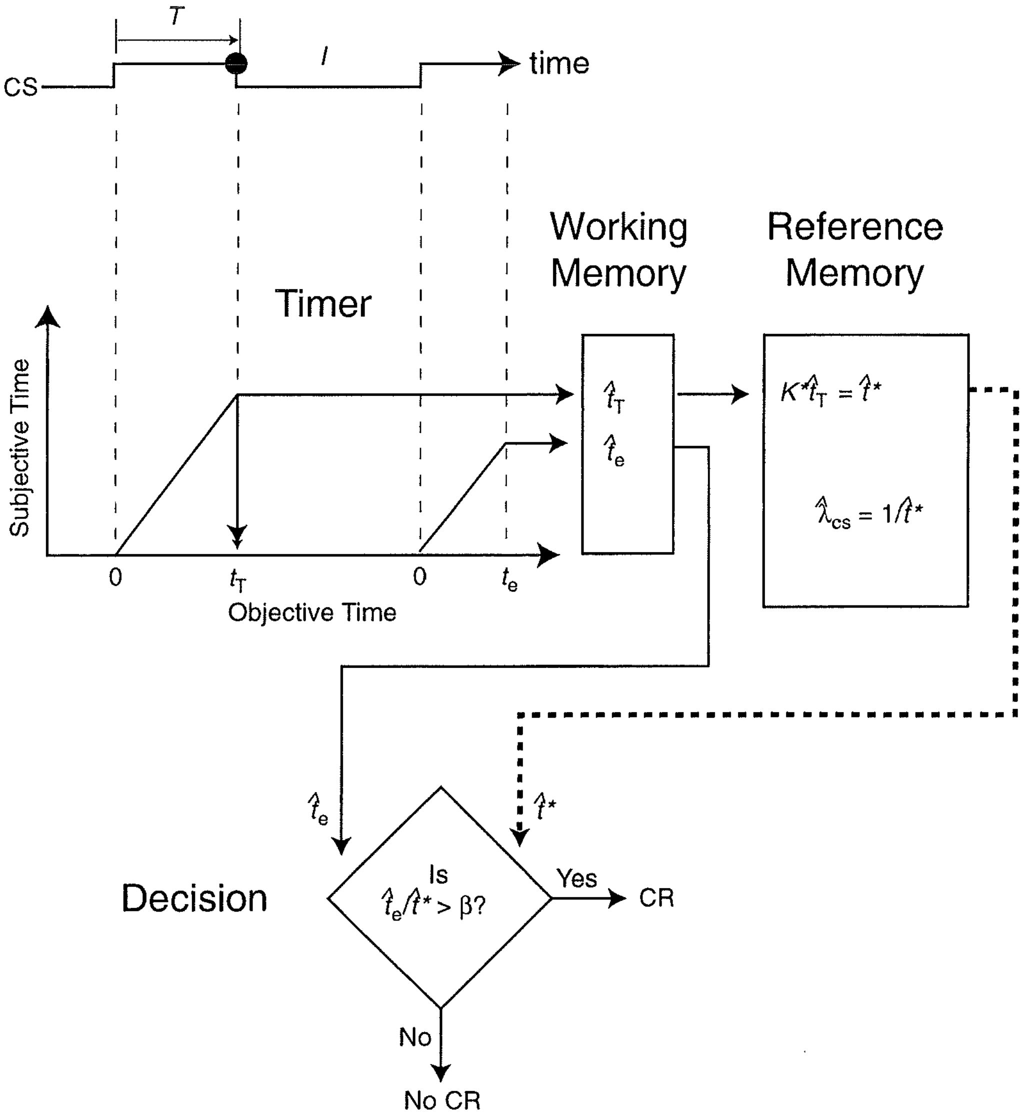

Figure 3. Flow diagram for the conditioned-response (CR) timing or the *when* decision. Two trials are shown: the first reinforced at T (filled circle on time line) and the second still elapsing at e. When the first trial is reinforced, the cumulated subjective time, $\hat{t}_T$, is stored in working memory and transferred to reference memory by a multiplicative variable, k^* ($\hat{t}^* = k^*\hat{t}_T$). The decision to respond is based on the ratio of the elapsing interval (in working memory) to the remembered interval (in reference memory). It occurs when this ratio exceeds a threshold (β) close to, but generally less than, 1. Note that the reciprocal of $\hat{t}^*$ is equal to $\hat{\lambda}_{CS}$, the estimated rate of conditioned-stimulus (CS) reinforcement, which plays a crucial role in the acquisition and extinction decisions described later (see the Acquisition section and the Extinction section of the text).

When accumulation is temporarily halted, for example, in paradigms when reinforcement is not delivered and the signal is briefly turned off and back on again after a short period (a gap), the value in the accumulator simply holds through the gap (working memory), and the integrator resumes accumulating when the signal comes back on.

Scalar variability, which is evident in the constant coefficient of variation in the distribution of the onsets, offsets, and peaks of conditioned responding, is a consequence of two fundamental assumptions. The first is that the comparison mechanism uses the ratio of the two values being compared, rather than, for example, their difference. The second is that subjective estimates of temporal durations, like subjective estimates of many other continuous variables (length, weight, loudness, etc.), obey Weber's law: The difference required to discriminate one subjective magnitude from another with a given degree of reliability is a fixed fraction of that magnitude (Gibbon, 1977; Killeen & Weiss, 1987). What this most likely implies—and what SET assumes—is that the uncertainty about the true value of a remembered magnitude is proportional to the magnitude. Both of these assumptions—the decision variable is a ratio and estimates of duration read from memory have scalar variability—are necessary to explain scale invariance in the dis-

Table 1
Symbols and Expressions in Scalar Expectancy Theory

Symbol or expression	Meaning
$\hat{t}_e$	Time elapsed since conditioned stimulus (CS) onset, the subjective measure of an elapsing interval
$\hat{t}_T$	Magnitude of $\hat{t}_e$ at time of reinforcement, the experienced duration of the CS–unconditioned stimulus (US) interval
$\hat{t}^*$	Remembered duration of CS–US interval
k*	Scaling factor relating $\hat{t}^*$ to $\hat{t}_T$: $\hat{t}^* = k^* \hat{t}_T$
K^*	Expected value of k*; close to but not equal to 1; the fact that the value is not equal to 1 explains the systematic discrepancy between average experienced duration and average remembered duration, the K^* error
$\hat{t}_e/\hat{t}^*$	The decision variable for the *when* decision, the measure of how similar the currently elapsed interval is to the remembered reinforcement latency
β	A decision threshold
T	Generally, a fixed reinforcement latency; in Pavlovian delay conditioning, the CS–US interval; in a fixed-interval operant schedule, the interval between reinforcements
λ_{CS}	Rate of reinforcement attributed to a CS, the reciprocal of the expected interval between reinforcements

Note. A hat on a variable indicates that it is a subjective estimate, a quantity in the head representing a physically measurable external variable. Variables without hats are measurable quantities outside the head, scaling constants (always symbolized by k), or decision thresholds (always symbolized by β).

tribution of CRs (Church & Gibbon, 1982; Gibbon & Fairhurst, 1994).

It has recently become clear that much of the observed trial-to-trial variability in response timing is due to the variability inherent in the signals derived from reading durations stored in long-term memory, rather than from variability in the timing process that generates inputs to memory. Even when there is only one such comparison duration in memory, a comparison signal, $\hat{t}^*$, derived from reading that one memory varies substantially from trial to trial. In some paradigms, the standard interval read from memory for comparison with a currently elapsing interval may be based on either a single standard (a single standard experienced repeatedly) or a double standard (two different values experienced in random intermixture). In the two-standards condition, the comparison value (expectation) recalled from memory is equal to the harmonic mean of the two standards. The trial-to-trial variability in timing performance observed in this two-standards condition is the same as in a one-standard condition with a standard equal to the harmonic mean of the standards in the two-standards condition. The variability in the input to memory is very different in the two conditions, but the output variability is the same. This implies that the trial-to-trial variability in the response latencies is largely due to noise in the memory reading operation rather than to variability in the values read into memory (see Gallistel, 1999, for a fuller discussion of the evidence for this conclusion).

The Timing of Appetitive Conditioned Responses

The fixed-interval scallop. An early application of SET was the explanation of the "FI scallop" in operant conditioning. An FI schedule of reinforcement delivers reinforcement for the first response made after a fixed interval has elapsed since the delivery of the last reinforcement. When responding on such a schedule, animals pause after each reinforcement and then resume responding after some interval has elapsed. It was generally supposed that the animals' rate of responding accelerated throughout the remainder of the interval leading up to reinforcement. In fact, however, conditioned responding in this paradigm, as in many others, is a two-state variable (slow, sporadic pecking vs. rapid, steady pecking), with one transition per interreinforcement interval (Schneider, 1969). The average latency to the onset of the high-rate state during the postreinforcement interval increases in proportion to the scheduled reinforcement interval over a very wide range of intervals (from 30 s to at least 50 min). The variability in this onset from one interval to the next also increases in proportion to the scheduled interval. As a result, averaging over many interreinforcement intervals results in the smooth increase in the average rate of responding that Dews (1970) called "proportional timing" (Figure 1A). The smooth, almost linear, increase in the average rate of responding seen in Figure 1A is the result of averaging across many different abrupt onsets. It could more appropriately be read as showing the probability that the subject has entered the high-rate state as a function of the time elapsed since the last reinforcement.

The peak procedure. The FI procedure allows one to see only the subject's anticipation of reinforcement. The peak procedure (Catania, 1970; Roberts, 1981) is a discrete-trials modification that enables one also to observe the cessation of responding when the expected time of reinforcement has passed without reinforcement. The beginning of each trial is marked by the illumination of the key (with pigeon subjects) or by the extension of the lever (with rat subjects). A response (key peck or leverpress) is reinforced only after a fixed interval has elapsed. However, a partial reinforcement schedule is used. On some trials, there is no reinforcement. On these trials (peak trials), the CS continues for three or four times the reinforcement latency, allowing the experimenter to observe the cessation of the CR when the expected time of reinforcement has passed. This procedure yielded the data shown in Figure 2, which came from the unreinforced trials.

The smoothness of the curves in Figure 2 is again an averaging artifact. On any one trial, there is an abrupt onset and an abrupt offset of steady responding (Church, Meck, & Gibbon, 1994; Church, Miller, Meck, & Gibbon, 1991; Gibbon & Church, 1992). The midpoint of the interval during which the animal responds (the CR interval) is proportionate to reinforcement latency and so is the average duration of this CR interval. However, there is considerable trial-to-trial variation in the onset and the offset of responding. The curves in Figure 2 are a consequence of averaging across the independently variable onsets and offsets of responding (Church et al., 1994). As in Figure 1A, these curves are more appropriately read as giving the probability that the animal will be responding at a high rate at any given fraction of the reinforcement latency.

The Timing of Aversive Conditioned Responses

Avoidance responses. The conditioned fear that is manifested in freezing behavior and other indices of a conditioned emotional response (CER) is classically conditioned in the operational sense that the reinforcement is not contingent on the animal's response.

Avoidance responses, by contrast, are instrumentally conditioned in the operational sense because their appearance depends on the contingency that the performance of the CR forestalls the aversive reinforcement. By responding, the subject avoids the aversive stimulus. We stress the purely operational, as opposed to the theoretical, distinction between classical and instrumental conditioning because, from the perspective of timing theory, the only difference between the two paradigms is in the events that mark the beginnings of expected and elapsing intervals. In the instrumental case, the expected interval to the next shock is longest immediately after a response, and the recurrence of a response resets the shock clock. Thus, the animal's response marks the onset of the relevant interval.

The timing of instrumentally conditioned avoidance responses is as dependent on the expected time of aversive reinforcement as the timing of classically conditioned emotional reactions, and it shows the same scale invariance in the mean and scalar variability around it (Gibbon, 1971, 1972). In shuttle box avoidance paradigms, in which the animal gets shocked at either end of the box if it stays too long, the mean latency at which the animal makes the avoidance response increases in proportion to the latency of the shock that is thereby avoided, and so does the variability in this avoidance latency. A similar result is obtained in free-operant avoidance paradigms, in which the rat must press a lever before a certain interval has elapsed in order to forestall for another such interval the shock that will otherwise occur (Gibbon, 1971, 1972, 1977; Libby & Church, 1974). As a result, the probability of an avoidance response at less than or equal to a given proportion of the mean latency is the same regardless of the absolute duration of the expected shock latency (see, e.g., Figure 1 in Gibbon, 1977). Scalar timing of avoidance responses is again a consequence of the central assumptions in SET—the use of a ratio to judge the similarity between the currently elapsed interval and the expected shock latency, and scalar variability (noise) in the shock latency durations read from memory.

When an animal must respond to avoid a pending shock, responding occurs long before the expected time of shock. One of the earliest applications of SET (Gibbon, 1971) showed that this early responding in avoidance procedures is nevertheless scalar in the shock delay (Figure 4). According to SET, the expectation of shock is maximal at the experienced latency between the onset of the warning signal and the shock, just as in other paradigms. However, a low decision threshold leads to responding at an elapsed interval equal to a small fraction of the expected shock latency. The result, of course, is successful avoidance on almost all trials. The low threshold compensates for trial-to-trial variability in the remembered duration of the warning interval. If the threshold were higher, the subject would more often fail to respond in time to avoid the shock. The low threshold ensures that responding almost always anticipates and thereby forestalls the shock.

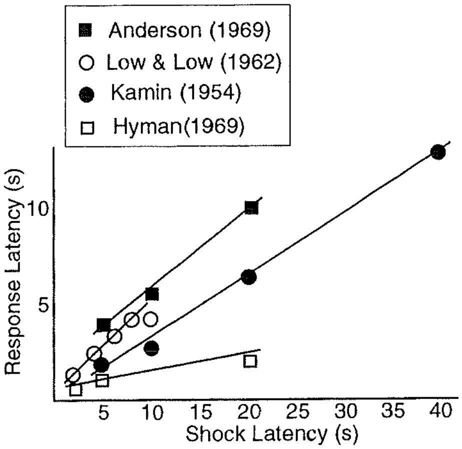

Figure 4. The mean latency of the avoidance response as a function of the latency of the shock (conditioned-stimulus/unconditioned-stimulus interval) in a variety of cued avoidance experiments with rats (Anderson, 1969; Kamin, 1954; Low & Low, 1962) and monkeys (Hyman, 1969). Note that although the response latency is much shorter than the shock latency, it is nonetheless proportional to the shock latency. The straight lines are drawn by eye. From "Scalar Timing and Semi-Markov Chains in Free-Operant Avoidance," by J. Gibbon, 1971, *Journal of Mathematical Psychology, 8,* p. 112. Copyright 1971 by Academic Press. Adapted with permission.

The conditioned emotional response. The CER is the suppression of appetitive responding that occurs when the subject (usually a rat) expects a shock to the feet (aversive reinforcement). The appetitive response is suppressed because the subject freezes in anticipation of the shock (Figure 1C). If shocks are scheduled at regular intervals, then the probability that the rat will stop its appetitive responding (pressing a bar to obtain food) increases with the fraction of the intershock interval that has elapsed. The suppression measures obtained from experiments using different intershock intervals are superimposable when they are plotted as a proportion of the intershock interval that has elapsed (LaBarbera & Church, 1974; see Figure 5). Put another way, the degree to which the rat fears the impending shock is determined by how close it is to the shock. Its subjective measure of closeness is the ratio between the interval elapsed since the last shock and the expected interval between shocks—a simple manifestation of scalar expectancy.

The immediate shock deficit. If a rat is shocked immediately after being placed in an experimental chamber (1–5-s latency), it shows very little CR (freezing) in the course of an 8-min test the next day. By contrast, if it is shocked several minutes after being placed in the chamber, it shows much more freezing during the subsequent test. The longer the reinforcement delay is, the more total freezing that is observed, up to several minutes (Fanselow, 1986). This has led to the suggestion that in conditioning an animal to fear the experimental context, the longer the reinforcement latency, the greater the resulting strength of the association will be (Fanselow, 1986, 1990; Fanselow, DeCola, & Young, 1993). This explanation of the *immediate-shock freezing deficit* rests on an ad hoc assumption made specifically to explain this phenomenon. Moreover, it is the opposite of the usual assumption about the effect of delay on the efficacy of reinforcement, namely, the shorter the delay, the greater the effect of reinforcement is.

From the perspective of SET, the immediate-shock freezing deficit is a manifestation of scalar variability in the distribution of the fear response about the expected time of shock. Bevins and Ayres (1995) varied the latency of the shock in a one-trial contextual-fear conditioning paradigm and showed that the later in the training session the shock was given, the later the observed peak in freezing behavior and the broader the distribution of this behavior throughout the session (Figure 6). The prediction of the

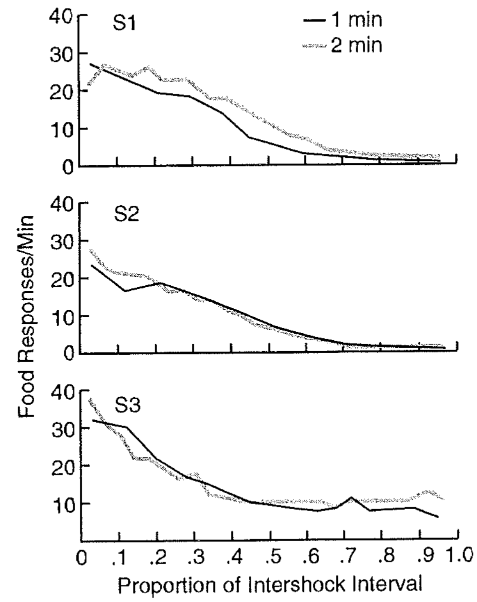

Figure 5. The strength of the conditioned emotional reaction to shock is measured by the decrease in appetitive responding when shock is anticipated (data from 3 rats—S1, S2, and S3). The decrease in responding for a food reward (a measure of the average strength of the fear) is determined by the proportion of the anticipated interval that has elapsed. Thus, the data from conditions using different fixed intershock intervals (1 and 2 min) are superimposable when normalized. This is timescale invariance in the fear response to impending shock. From "Magnitude of Fear as a Function of the Expected Time to an Aversive Event," by J. D. LaBarbera and R. M. Church, 1974, *Animal Learning and Behavior, 2,* p. 200. Copyright 1974 by the Psychonomic Society. Adapted with permission.

immediate shock deficit follows directly from the scalar variability of the fear response about the moment of peak probability. If the probability of freezing in a test session following training with a 3-min shock delay is given by the broad normal curve in Figure 7 (cf. freezing data in Figure 1C), then the distribution after a 3-s latency should be 60 times narrower (3-s curve in Figure 7). Thus, the amount of freezing observed during an 8-min test session following an immediate shock should be negligible in comparison with the amount observed following a shock delayed for 3 min.

It is important to note that our explanation of the failure to see significant evidence of fear in the chamber after the subjects have experienced short-latency shock does not imply that there is no fear associated with that brief delay. On the contrary, we suggest that the subjects fear the shock just as much in the short-latency condition as in the long-latency condition. But the fear begins and ends much sooner; hence, there is much less measured evidence of fear. Because the average breadth of the interval during which the subject fears shock grows in proportion to the remembered latency of that shock, the total amount of fearful behavior (number of seconds of freezing) observed is much greater with longer shock latencies.

The eyeblink. The conditioned eyeblink is often regarded as a basic or primitive example of a classically conditioned response to an aversive US. A fact well-known to those who have directly observed this CR is that the latency to the peak of the CR approximately matches the CS–US latency. Although the response is over literally in the blink of an eye, it is so timed that the eye is closed at the moment when the aversive stimulus is expected. Figure 1B provides an interesting example. In the experiment from which this representative plot of a double blink is taken (Kehoe, Graham-Clarke, & Schreurs, 1989), there was only one US on any given trial, but it occurred either 400 ms or 900 ms after CS onset, in a trial-to-trial sequence that was random (unpredictable). The rabbit learned to blink twice, once at about 400 ms and then again at 900 ms. Clearly, the timing of the eyeblink—the fact that longer reinforcement latencies produce longer latency blinks—cannot be explained by the idea that longer reinforcement latencies produce weaker associations. The fact that the blink latencies approximately match the expected latencies of the aversive stimuli to the

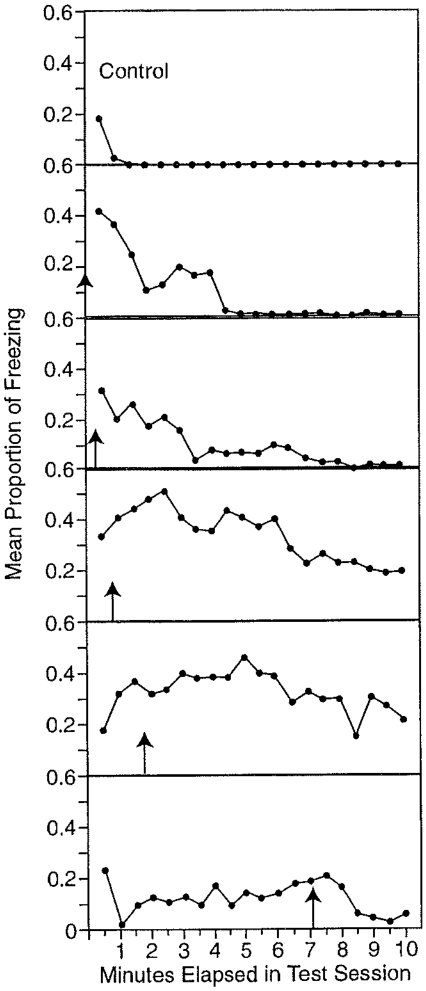

Figure 6. The distribution of freezing behavior in a 10-min test session following a single training session in which groups of rats were shocked once at different latencies (vertical arrows) after being placed in the experimental box (and removed 30 s after the shock). The control rats were shocked immediately after being placed in a different box (a different context from the one in which their freezing behavior was observed on the test day). From "One-Trial Context Fear Conditioning as a Function of the Interstimulus Interval," by R. A. Bevins and J. J. B. Ayres, 1995, *Animal Learning and Behavior, 23,* p. 403. Copyright 1995 by the Psychonomic Society. Adapted with permission.

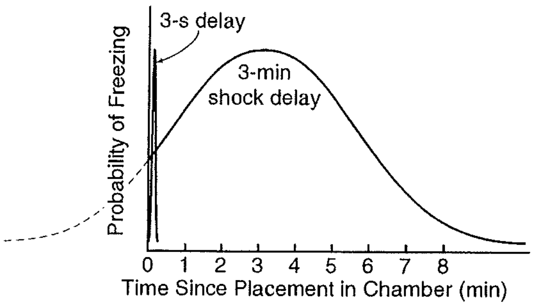

Figure 7. Explanation of the immediate-shock freezing deficit by scalar expectancy theory: Given the probability-of-freezing curve shown for the 3-min group (see Figure 1C), the scale invariance of conditioned-response distributions predicts the very narrow curve shown for subjects shocked immediately (3 s) after placement in the box. Scoring the percent of subjects freezing during the 8-min test period will show much more freezing in the 3-min group than in the 3-s group (about 60 times more).

eye is a simple indication that the learning of the temporal interval to reinforcement is a foundation of simple classically conditioned responding. Recent findings with this preparation further imply that the learning of the temporal intervals in the protocol is the foundation of the higher order effects called positive and negative patterning and occasion setting (Weidemann, Georgilas, & Kehoe, 1999).

The record in Figure 1B does not exhibit scalar variability because it is a record of the blinks on a single trial. Blinks, like pecks, have, we assume, a more or less fixed duration because they are ballistic responses programmed by the central nervous system. What exhibits scalar variability from trial to trial is the time at which the CR is initiated. In cases like pigeon pecking, in which the CR is repeated steadily for some while so that there is a stop decision as well as a start decision, the duration of conditioned responding shows the scalar property on individual trials. That is, the interval between the onset of responding and its cessation increases in proportion to the midpoint of the CR interval. In the case of the eyeblink, however, in which there is only one CR per expected US per trial, the duration of the CR may be controlled by the motor system itself rather than by higher level decision processes. The distribution of these CRs from repeated trials should, however, exhibit scalar variability. (John W. Moore and E. J. Kehoe have gathered data indicating a constant coefficient of variation in distributions of rabbit eyeblink latencies [J. W. Moore, personal communication, February 13, 2000].)

Timing the Conditioned Stimulus: Discrimination

The acquisition and extinction models to be considered shortly assume that the animal times the durations of the CSs it experiences and compares those durations with durations stored in memory. It is possible to directly test this assumption by presenting CSs of different duration and then asking the subject to indicate by a choice response which of two durations it just experienced. In other words, the duration of the just-experienced CS is made the basis of a discrimination in a successive discrimination paradigm, a paradigm in which the stimuli to be discriminated are presented individually on successive trials, rather than simultaneously in one trial. In the so-called bisection paradigm, the subject is reinforced for one choice after hearing a short-duration CS (say, a 2-s CS) and for the other choice after hearing a long-duration CS (say, an 8-s CS). After learning the reference durations (the "anchors"), the subject is probed with intermediate durations and required to make classification responses to these durations.

If the subject uses ratios to compare probe durations with the reference durations in memory, then the point of indifference, the probe duration that it judges to be equidistant from the two reference durations, will be at the geometric mean of the reference durations rather than at their arithmetic mean. SET assumes that the decision variable in the bisection task is the ratio of the similarities of the probe to the two reference durations. The similarity of two durations by this measure is the ratio of the smaller to the larger. Perfect similarity is a ratio of 1:1. Thus, for example, a 5-s probe is more similar to an 8-s probe than to a 2-s probe, because 5/8 is closer to 1 than is 2/5. If, by contrast, similarity were measured by the extent to which the difference between two durations approaches 0, then a 5-s probe would be equidistant (equally similar) to a 2-s and an 8-s referent, because $8 - 5 = 5 - 2$. Maximal uncertainty (indifference) should occur at the probe duration that is equally similar to 2 and 8. If similarity is measured by ratios rather than differences, then the probe is equally similar to the two anchors for T, such that $2/T = T/8$, or $T = 4$, the geometric mean of 2 and 8.

As predicted by the ratio assumption in SET, the probe duration at the point of indifference is in fact generally the geometric mean, which is the duration at which the ratio measures of similarity are equal, rather than the arithmetic mean, which is the duration at which the difference measures of similarity are equal (Church & Deluty, 1977; Gibbon et al., 1984; see Penney, Allan, Meck, & Gibbon, 1998, for a review and extension to human time discrimination) Moreover, the plots of the percent choice of one referent or the other as a function of the probe duration are scale-invariant, which means that the psychometric discrimination functions obtained from different pairs of reference durations are superimposed when time is normalized by the geometric mean of the reference durations (Church & Deluty, 1977; Gibbon et al., 1984).

Acquisition

Acquisition of Responding to the Conditioned Stimulus

The conceptual framework that we propose for the understanding of conditioning is, essentially, the decision-theoretic conceptual framework, which has long been used in psychophysical research and which has informed SET from its inception. In the psychophysical decision-theoretic framework, there is a stimulus whose strength may be varied by varying relevant parameters. The stimulus might be, for example, a light flash whose detectability is affected by its intensity, duration, and luminosity. The stimulus gives rise through an often complex computational process to a noisy internal signal called the decision variable. The stronger the stimulus, the greater the mean value of this noisy decision variable is. The subject responds when the decision variable exceeds a decision threshold. The stronger the stimulus is, the more likely the decision variable is to exceed the decision threshold; hence, the more likely the subject is to respond. The plot of the subject's response probability as a function of the strength of the stimulus (e.g., its intensity or duration or luminosity) is called the psychometric function.

In our analysis of conditioning, the conditioning protocol is the stimulus. The temporal intervals in the protocol—including the cumulative duration of the animal's exposure to the protocol—are the relevant parameters of the stimulus, as are the reinforcement magnitudes when they also vary. These stimulus parameters determine the value of a decision variable through a to-be-described computational process called rate estimation theory (RET). The decision variable is noisy because of both external and internal sources. The animal responds to the CS when the decision variable exceeds an acquisition threshold. The decision process is adapted to the characteristics of the noise.

The acquisition function in conditioning is equivalent to the psychometric function in a psychophysical task. Its rise (the increasing probability of a response as exposure to the protocol is prolonged) reflects the growing magnitude of the decision variable. The visual stimulus in the aforementioned example gets stronger as the duration of the flash is prolonged because the longer a light of a given intensity is continued, the more evidence there is of its presence (up to some limit). Similarly, a conditioning protocol gets stronger as the duration of the subject's exposure to it increases because the continued exposure to the protocol gives stronger and stronger objective evidence that the CS makes a difference in the rate of reinforcement (stronger and stronger evidence of CS–US contingency).

In modeling acquisition, we try to emulate psychophysical modeling by paying closer attention to quantitative results, rather than predicting only the directions of effects. However, our efforts to quantitatively test models of the simple acquisition process are hampered by a paucity of data on acquisition in individual subjects. Most published acquisition curves are group averages. These are likely to contain averaging artifacts. If individual subjects acquire a CR abruptly, but different subjects acquire it after different amounts of experience, the averaging across subjects will yield a smooth, gradual group acquisition curve, even though acquisition in each individual subject was abrupt. Thus, the form of the "psychometric function" (acquisition function) for individual subjects is not well established.

Quantitative facts about the effects of basic variables such as partial reinforcement, delay of reinforcement, and intertrial interval on the rate of acquisition and extinction also have not been as well established as one might suppose given the rich history of experimental research on conditioning and the long-recognized importance of these parameters.[2] In recent years, pigeon autoshaping has been the most extensively used appetitive-conditioning preparation. The most systematic data on rates of acquisition and extinction come from it. Data from other preparations, notably rabbit jaw-movement conditioning (another appetitive preparation), the rabbit nictitating-membrane preparation (aversive conditioning), and the conditioned suppression of appetitive responding (CER) preparation (also aversive), appear to be consistent with these data but do not permit as strong quantitative conclusions.

Pigeon autoshaping is a fully automated variant of Pavlov's classical-conditioning paradigm. The protocol for it is diagrammed in Figure 8A. The CS is the transillumination of a round button (key) on the wall of the experimental enclosure. The illumination of the key may or may not be followed after some delay by the brief presentation of a hopper full of food (reinforcement). Instead of salivating to the stimulus that predicts food, as Pavlov's dogs did, the pigeon pecks at it. The rate or probability of pecking the key is the measure of the strength of conditioning. As in Pavlov's original protocol, the CR (pecking) is the same or nearly the same as the unconditioned response elicited by the US. In this paradigm, as in Pavlov's paradigm, the food is delivered at the end of the CS whether or not the subject pecks the key. Thus, it is a classical-conditioning paradigm rather than an operant-conditioning paradigm. As an automated means for teaching pigeons to peck keys in operant-conditioning experiments, it has replaced experimenter-controlled shaping. It is now common practice to condition the pigeon to peck the key by reinforcing key illumination whether or not the pigeon pecks (a Pavlovian procedure) and only then intro-

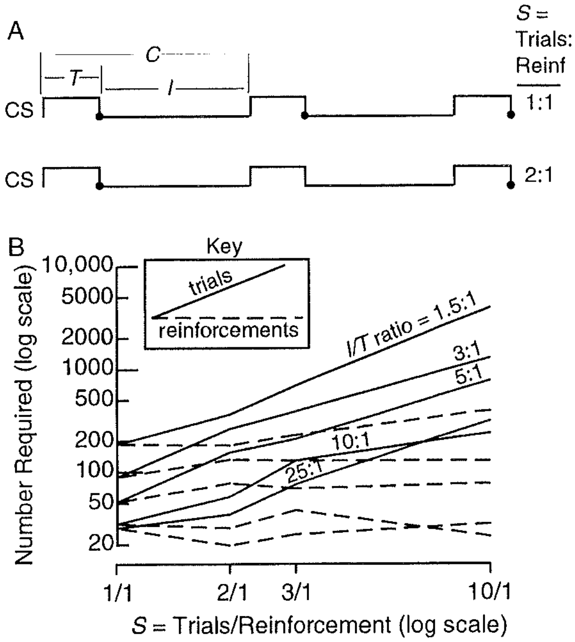

Figure 8. A: Time lines showing the variables that define a classical (Pavlovian) conditioning protocol—the duration of a conditioned-stimulus (CS) presentation (*T*), the duration of the intertrial interval (*I*), and the reinforcement schedule (*S* [trials/reinforcement]). The unconditioned stimulus (US [reinforcement]) is usually presented at the termination of the CS (black dots). For reasons shown in Figure 12, the US may be treated as a point event, an event whose duration can be ignored. The sum of *T* and *I* is *C*, the duration of the trial cycle. B: Trials to acquisition (solid lines) and reinforcements to acquisition (dashed lines) in pigeon autoshaping, as a function of the reinforcement schedule and the *I*/*T* ratio. Note that the solid and dashed lines come in pairs, with the members of a pair joined at the 1/1 value of *S*, because, with that schedule (continual reinforcement), the number of reinforcements and the number of trials are identical. The acquisition criterion was at least one peck on three out of four consecutive presentations of the CS. Reanalysis of data in Figure 1 of Gibbon, Farrell, Locurto, Duncan and Terrace (1980).

---

[2] This is due, in part, to the fact that meaningful data on acquisition could not be collected before the advent of fully automated conditioning paradigms. When experimenter judgment enters into the training in an on-line manner, as is the case when animals are "shaped," or when the experimenter handles the subjects on every trial (as in most maze paradigms), the skill and attentiveness of the experimenter is an important but unmeasured factor.

duce the operant contingency on responding. The discovery that pigeon key pecking—the prototype of the operant response—could be so readily conditioned by a classical (Pavlovian) rather than an operant protocol has cast doubt on the traditional assumption that classical and operant protocols tap fundamentally different association-forming processes (Brown & Jenkins, 1968).

Some well-established facts about the acquisition of a CR are as follows:

- The "strengthening" of the CR with extended experience: It takes a number of reinforced trials for an appetitive CR to emerge.
- No effect of partial reinforcement: Reinforcing only some of the CS presentations increases the number of trials required to reach an acquisition criterion in both Pavlovian paradigms (Figure 8B, solid lines) and operant discrimination paradigms (Williams, 1981). However, the increase is proportional to the thinning of the reinforcement schedule—the average number of trials per reinforcement (the thinning factor). Hence, the required number of reinforcements is unaffected by partial reinforcement (Figure 8B, dashed lines). Thus, the nonreinforcements that occur during partial reinforcement do not affect the *rate of acquisition*, defined as the reciprocal of reinforcements to acquisition.
- Effect of the intertrial interval: Increasing the average interval between trials increases the rate of acquisition; that is, it reduces the number of reinforcements required to reach an acquisition criterion (Figure 8B, dashed lines) and, hence, trials to acquisition (Figure 8B, solid lines). More quantitatively, reinforcements to acquisition are approximately inversely proportional to the *I*/*T* ratio (Figures 9 and 10), which is the ratio of the intertrial duration (*I*) to the duration of a CS presentation (*T*, for trial duration). If the CS is reinforced on termination (as in Figure 8A), then *T* is also the reinforcement latency or delay of reinforcement. This interval is also called the CS–US interval or the interstimulus interval. The effect of the *I*/*T* ratio on the rate of acquisition is independent of the reinforcement schedule, as can be seen from the fact that the solid lines are parallel in Figure 8B, as are, of course, the dashed lines.

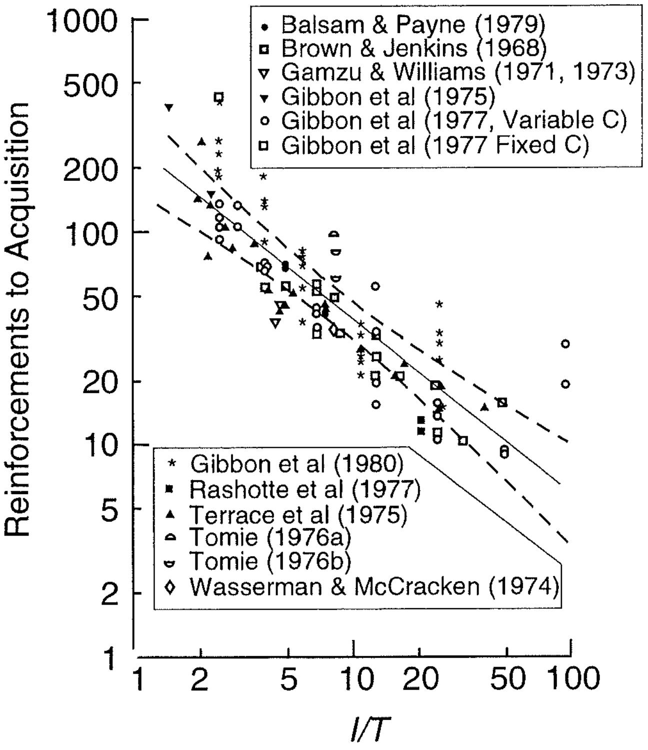

Figure 9. Reinforcements to acquisition as a function of the ratio of the duration of the intertrial interval (*I*) to the duration of the conditioned-stimulus presentation (*T*; double logarithmic coordinates). The data are from 12 experiments in several different laboratories.

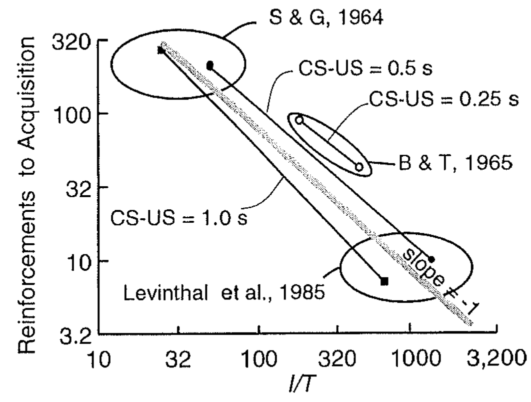

Figure 10. Selected data showing the effect of *I*/*T* ratio on the rate of eyeblink conditioning in rabbits, where *I* is the estimated amount of exposure to the experimental apparatus per conditioned-stimulus (CS) trial (the time when the subject was outside the apparatus was not counted) and *T* is the CS–US interval. We used 50% conditioned-response frequency as the acquisition criterion in deriving these data from published group acquisition curves. S & G, 1964 = Schneiderman and Gormezano (1964), 70 trials per session, session length approximately half an hour, *I* varied randomly with a mean of 25 s. B & T, 1965 = Brelsford and Theios (1965), single-session conditioning, *I*s were 45, 111, and 300 s, session lengths increased with *I* (1.25 and 2 hr for data shown). We do not show the 300-s data because those sessions lasted about 7 hr. Fatigue, sleep, growing restiveness, and so forth may have become important factors. Levinthal et al., 1985 = Levinthal, Tartell, Margolin, and Fishman (1985), one trial per 11-min (660-s) daily session. None of these studies were designed to study the effect of *I*/*T* ratio, so the plot should be treated with caution. Such studies are clearly desirable—in this and other standard conditioning paradigms. US = unconditioned stimulus.

- Delay of reinforcement: Increasing the delay of reinforcement, while holding the intertrial interval constant, retards acquisition—in proportion to the increase in the reinforcement latency (Figure 11, solid line). Because *I* is held constant while *T* is increased, delaying reinforcement in this manner reduces the *I*/*T* ratio. The effect of delaying reinforcement is entirely due to the reduction in the *I*/*T* ratio. Delay of reinforcement per se does not affect acquisition (Figure 11, dashed line).
- Timescale invariance: When the intertrial interval is increased in proportion to the delay of reinforcement, delay of reinforcement has no effect on reinforcements to acquisition (Figure 11, dashed line). Increasing the intertrial interval in proportion to the increase in CS duration means that all the temporal intervals in the conditioning protocol are increased by a common scaling factor. Therefore, we call this important result the "timescale invariance" of the acquisition process. The failure of partial reinforcement to affect rate of acquisition and the constant coefficient of variation in

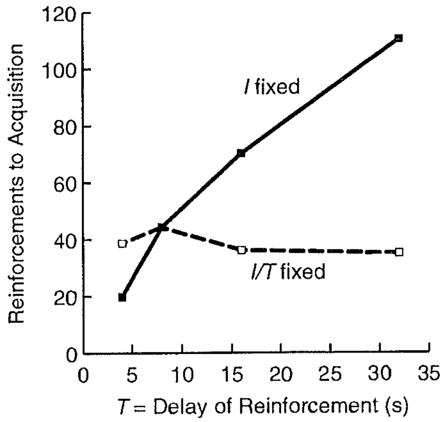

Figure 11. Reinforcements to acquisition as a function of delay of reinforcement (*T*), with the (average) intertrial interval (*I*) fixed (solid line) or varied (dashed line) in proportion to delay of reinforcement. For the solid line, *I* was fixed at 48 s. For the dashed line, the *I*/*T* ratio was fixed at 5. The data are replotted (by interpolation) from data originally reported in Gibbon et al. (1977).

reinforcements to acquisition (constant vertical scatter about the regression line in Figure 9) are other manifestations of timescale invariance, as we explain later.

- Irrelevance of reinforcement magnitude: Above some threshold level, the amount of reinforcement has little or no effect on the rate of acquisition. Increasing the amount of reinforcement by increasing the duration of food-cup presentation 15-fold does not reduce reinforcements to acquisition. In fact, the rate of acquisition can be dramatically increased by reducing reinforcement duration and adding the time thus saved to the intertrial interval (Figure 12). The intertrial interval, the interval when nothing happens, matters profoundly in acquisition; the duration or magnitude of the reinforcement does not.

- Acquisition requires contingency (the truly random control): When reinforcements are delivered during the intertrial interval at the same rate as they occur during the CS, conditioning does not occur (the truly random control, also known as the effect of background conditioning; Rescorla, 1968). The failure of conditioning under these conditions is not simply a performance block, because conditioned responding to the CS after random control training is not observable even with sensitive techniques (Gibbon & Balsam, 1981). The truly random control eliminates the contingency between CS and US while leaving the frequency of their temporal pairing unaltered. Its effect on conditioning implies that conditioning is driven by CS–US contingency, not by the temporal pairing of CS and US.

- Effect of signaling "background" reinforcers: In the truly random control procedure, acquisition to a target CS does occur if another CS precedes (and thereby signals) the "background" reinforcers (Durlach, 1983). These signaled reinforcers are no longer background reinforcers if, by a background reinforcer, one means a reinforcer that occurs in the presence of the background alone.

We have presented data from pigeon autoshaping to illustrate the basic facts of acquisition (Figures 8, 9, 11, and 12) because the most extensive and systematic quantitative data come from experiments using that paradigm. However, the same effects (and surprising lack of effects) seem to be apparent in other classical-conditioning paradigms. For example, partial reinforcement produces little or no increase in reinforcements to acquisition in a wide variety of paradigms (see citations in Table 2 of Gibbon, Farrell, Locurto, Duncan, & Terrace, 1980; see also Holmes & Gormezano, 1970; Prokasy & Gormezano, 1979), whereas lengthening the amount of exposure to the experimental apparatus per CS trial increases the rate of conditioning in the rabbit nictitating-membrane preparation by almost two orders of magnitude (Kehoe & Gormezano, 1974; Levinthal, Tartell, Margolin, & Fishman, 1985; Schneiderman & Gormezano, 1964; see Figure 10). Thus, it appears to be generally true that varying the *I*/*T* ratio has a much stronger effect on the rate of acquisition than does varying the

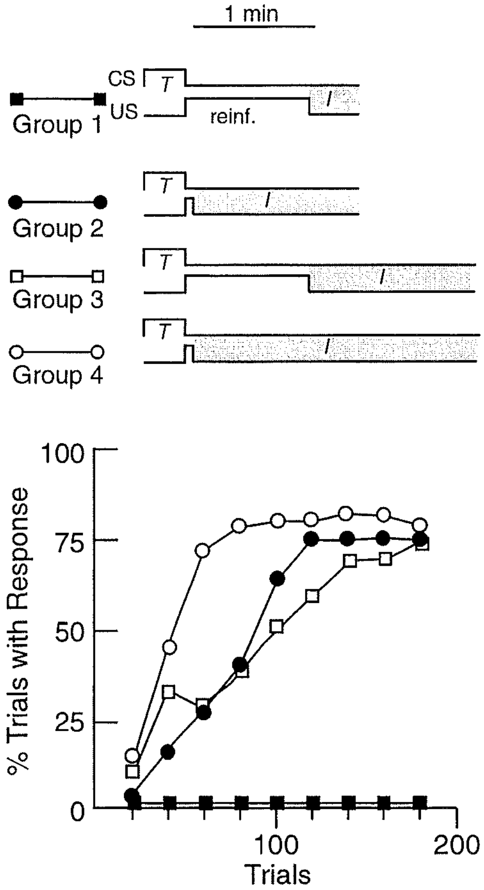

Figure 12. Effect on rate of acquisition of allocating time either to reinforcement (reinf.) or to the intertrial interval (*I*). Groups 1 and 2 had the same duration of the trial cycle ($T + I +$ reinforcement time), but Group 2 had its reinforcement duration reduced by a factor of 15 (from 60 to 4 s). The time thus saved was added to *I*. Group 2 acquired a conditioned response, whereas Group 1 did not. Groups 3 and 4 had longer (and equal) cycle durations. Again, a 56-s interval was used either for reinforcement (Group 3) or as part of *I* (Group 4). Group 4 acquired most rapidly. Group 3, which had the same *I*/*T* ratio as Group 2, acquired no faster than Group 2, despite getting 15 times more access to food per reinforcement. CS = conditioned stimulus; US = unconditioned stimulus. From "Intertrial Interval and Unconditioned Stimulus Durations in Autoshaping," by P. D. Balsam and D. Payne, 1979, *Animal Learning and Behavior, 7,* p. 478. Copyright 1979 by the Psychonomic Society. Adapted with permission.

degree of partial reinforcement, regardless of the conditioning paradigm used.

It also appears to be generally true that in both appetitive- and aversive-conditioning paradigms, varying the magnitude or intensity of reinforcement has little effect on the rate of acquisition. Increasing the magnitude of the water reinforcement in rabbit jaw-movement conditioning 20-fold has no effect on the rate of acquisition (Sheafor & Gormezano, 1972). Annau and Kamin (1961) examined the effect of shock intensity on the rate at which fear-induced suppression of appetitive responding is acquired. All of the groups receiving the three highest intensities (0.85, 1.55, and 2.91 mA) went from negligible levels of suppression to complete suppression on the 2nd day of training (between Trials 4 and 8). The group receiving the next lower shock intensity (0.49 mA) showed less than 50% suppression asymptotically. Kamin (1969a) later examined the effect of two levels of shock intensity on the rate at which CERs to a light CS and a noise CS were acquired. He used 1 mA, which is the usual level used in CER experiments, and 4 mA, which is a very intense shock. The 1-mA groups crossed the 50% median suppression criterion between Trials 4 and 5, whereas the 4-mA groups crossed this criterion between Trials 3 and 4. Thus, varying shock intensity from the minimum level that sustains a vigorous fear response up to very high levels has little effect on the rate of CER acquisition.

The lack of an effect of US magnitude or intensity on the number of reinforcements required for acquisition is counterintuitive and merits further investigation in a variety of paradigms. In such investigations, it will be important to show data from individual subjects to avoid averaging artifacts. For the same reason, it will be important not to bin the responses by session or number of trials, and so forth. What one wants is the real-time record of responding. Finally, it will be important to distinguish between the *asymptote* of the acquisition function and the *location* of its rise, defined as the number of reinforcements required to produce, for example, a half-maximal rate of responding. At least from a psychophysical perspective, only the latter measure is relevant to determining the rate of acquisition. In psychophysics, it has long been recognized that it is important to distinguish between the location of the psychometric function along the *x*-axis (in this case, reinforcements to acquisition), on the one hand, and the asymptote of the function, on the other hand. The location of the function indicates the underlying rate or sensitivity, whereas its asymptote reflects performance factors. The same distinction is used in pharmacology: The location (dose required) for the half-maximal response indicates affinity, whereas the asymptote indicates performance factors such as the number of receptors available for binding.

We do not claim that reinforcement magnitude is unimportant in conditioning. As we emphasize later on, it is a very important determinant of preference. It is also an important determinant of the asymptotic level of performance. And, if the magnitude of reinforcement varied depending on whether the reinforcement was delivered during the CS or during the background, we would expect magnitude to affect rate of acquisition as well. A lack of effect on rate of acquisition is observed (and, in our analysis, expected) only when there are no background reinforcements (the usual case in simple conditioning) or when the magnitude of background reinforcements is the same as the magnitude of CS reinforcements (the usual case when there is background conditioning).

Rate Estimation Theory

From a timing perspective, acquisition is a consequence of decisions that the animal makes about whether to respond to a CS. Our models for these decisions are adapted from Gallistel's (1990, 1992a, 1992b) earlier accounts, which we call RET. In our acquisition model, the decision to respond to the CS in the course of conditioning is based on the animal's growing certainty that the CS has a substantial effect on the rate of reinforcement. In simple conditioning, this certainty appears to be determined by the subject's estimate of the maximum possible value for the rate of background reinforcement given its experience of the background up to a given point in conditioning. Its estimate of the upper limit of what the rate of background reinforcement may be decreases steadily as conditioning progresses because the subject never experiences a background reinforcement (in simple conditioning). The subject's estimate of the rate of CS reinforcement, by contrast, remains stable because the subject gets reinforced after every so many seconds of exposure to the CS. The decision to respond is based on the ratio of these rate estimates, as shown in Figure 13. This ratio gets steadily larger as conditioning progresses because the upper limit on the background rate gets steadily lower. It should already be apparent why the amount of background exposure is so important in acquisition. It determines how rapidly the estimate for the background rate of reinforcement diminishes.

The ratio of two estimates for rates of reinforcement is equivalent to the ratio of two estimates of the expected interval between reinforcements (the interval–rate duality principle). Thus, any model couched in terms of rate ratios can also be couched in terms of the ratios of the expected intervals between events. When couched in terms of the expected intervals between reinforcements, the RET model of acquisition is as follows: Because the subject never experiences a background reinforcement in standard delay conditioning (after the hopper training), its estimate of the interval between background reinforcements gets longer in proportion to the duration of its unreinforced exposure to the background. By contrast, its estimate of the interval between reinforcements when the CS is on remains constant because it gets reinforced after every *T* seconds of CS exposure. Thus, the ratio of the two expected intervals gets steadily greater as conditioning progresses. When this ratio exceeds a decision threshold, the animal begins to respond to the CS.

The interval–rate duality principle means that the decision variables in SET and RET are the same kind of variables. Both decision variables are equivalent to the ratio of two estimated intervals. Rescaling time does not affect these ratios, which is why both models are timescale-invariant. This timescale invariance is, we believe, unique to timing-based models of conditioning with decision variables that are ratios of estimated intervals. It provides a simple way of discriminating experimentally between these models and associative models. There are, for example, many associative explanations for the trial-spacing effect (Barela, 1999), which is the strong effect that lengthening the intertrial interval has on the rate of acquisition (Figures 9 and 10). To our knowledge, none of them are timescale-invariant. That is, in none of them is it true that the magnitude of the trial-spacing effect is determined

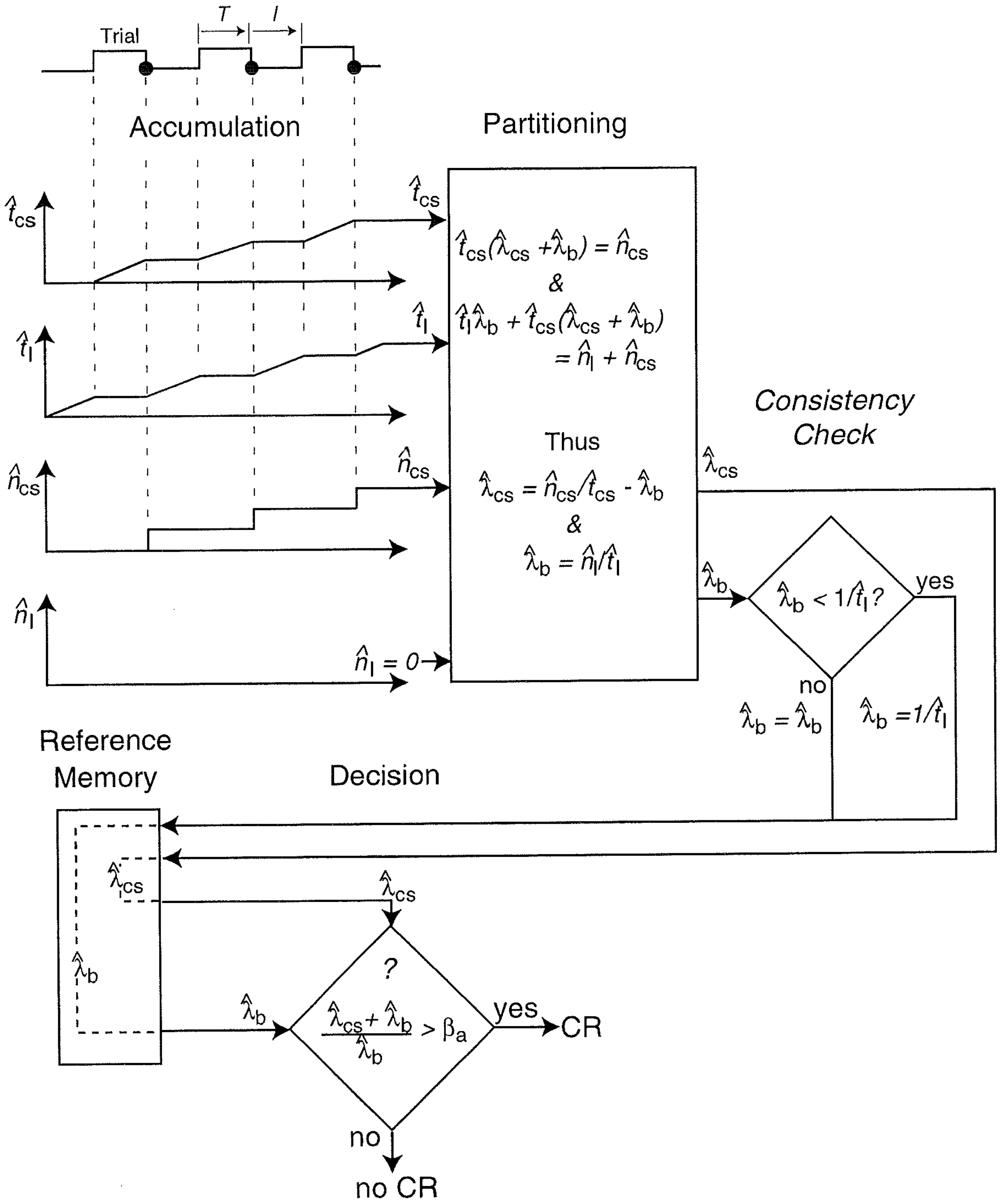

Figure 13. Functional structure (flow diagram) of the *whether* decision in acquisition. In simple conditioning, reinforcements (black dots) coincide with each conditioned-stimulus (CS) offset, and there are no background reinforcements (no dots during intertrial intervals). Subjective duration is cumulated separately for the CS ($\hat{t}_{CS}$) and for the background ($\hat{t}_I$), as are the subjective numbers of reinforcements ($\hat{n}_{CS}$ and $\hat{n}_I$) These values in working memory enter into the partition computation to obtain estimated rates of reinforcement for the CS ($\hat{\lambda}_{CS}$) and for the background ($\hat{\lambda}_b$). The rate estimates are continually updated and stored in reference memory. A rate estimate can never be less than the reciprocal of the cumulative interval of observation. When an estimate is lower than this (typically, an estimate of a rate of zero), it is replaced by the reciprocal of the total exposure to the background alone (consistency check). The decision that the CS predicts an increased rate of reinforcement occurs when the ratio of the rate of reinforcement expected when the CS is present ($\hat{\lambda}_{CS} + \hat{\lambda}_b$) to the estimated background rate of reinforcement ($\hat{\lambda}_b$) equals or exceeds a criterion, β_a. CR = conditioned response.

simply by the relative amounts of exposure to the CS and to the background alone in the protocol (Figure 11). The explanation of the trial-spacing effect given by Wagner's (1981) "sometimes opponent process" model, for example, depends on the rates at which stimulus traces decay from one state of activity to another. The size of the predicted effect of trial spacing will not be the same for protocols that have the same proportion of CS exposure to intertrial interval and differ only in their timescale, because longer timescales will lead to more decay. This timescale dependence is seen in the predictions of any model that assumes intrinsic rates of decay (of, e.g., stimulus traces, as in Sutton & Barto, 1990) or any model that assumes that experience is carved into trials (e.g., Rescorla & Wagner, 1972).

RET offers a model of acquisition that is distinct from, albeit similar in inspiration to, the model proposed by Gibbon and Balsam (1981). The idea underlying both models is that the decision whether to respond to a CS in the course of conditioning depends on a comparison of the estimated rate of CS reinforcement and the estimated rate of background reinforcement (cf. the comparator hypothesis in Cole, Barnet, & Miller, 1995a; Miller, Barnet, & Grahame, 1992). In our current proposal, RET incorporates scalar variability in the interval estimates, just as SET did in estimating the point within the CS at which responding should be seen. In RET, however, two new principles are introduced: First, the relevant time intervals are cumulated across successive occurrences of the CS and across successive intervals of background alone. The total cumulated time in the CS and the total cumulated exposure to the background are integrated throughout a session and even across sessions, provided no change in rates of reinforcement is detected.

Cumulations over separated occurrences of a signal have previously been shown to be relevant to performance when no reinforcers intervene at the end of successive CSs. These are the "gap" (Meck, Church, & Gibbon, 1985) and "split trials" (Gibbon & Balsam, 1981) experiments, which show that subjects do, indeed, cumulate successive times over successive occurrences of a signal. However, the cumulations proposed in RET extend over much greater intervals (and much greater gaps) than those used in the just-cited experiments. This raises the important question of how accumulation without (practical) limit may be realized in the brain. We conjecture that the answer to this question may be related to the question of the origin of the scalar variability in remembered magnitudes. Pocket calculators accumulate magnitudes (real numbers) without practical limit but not with a precision that is independent of magnitude. What is fixed is the number of significant digits, hence, the percent accuracy with which a magnitude (real number) may be specified. The scalar noise in remembered magnitudes gives them the same property: A remembered magnitude is only specified to within plus or minus a certain percentage of its "true" value, and the decision process is adapted to take account of this. Scalar uncertainty about the value of an accumulated magnitude may be inherent in any scheme that permits accumulation without practical limit, for example, through a binary cascade of accumulators as suggested by Gibbon, Malapani, et al. (1997) and developed quantitatively by Killeen and Taylor (in press). Our point is that scalar uncertainty about the value of a quantity may be inherent in a scale-invariant computational device, a device capable of working with magnitudes of any scale.

The second important way in which the RET model of acquisition differs from the earlier SET model is that it incorporates a partitioning process into the estimation of rates. Partitioning is fundamental to RET because RET starts from the observation that when only a few reinforcements have occurred in the presence of a CS, it is inherently ambiguous whether they should be credited entirely to the CS, entirely to the background, or some to each. Thus, any process that is going to make decisions based on separate rate estimates for the CS and the background needs a mechanism that partitions the observed rates of reinforcement among the possible predictors of those rates. The partitioning process in RET leads in some cases (e.g., in the case of "signaled" background reinforcers; see Durlach, 1983) to estimates for the background rate of reinforcement that are not the same as the observed estimates assumed by Gibbon and Balsam's (1981) model.

We postpone discussion of the partitioning process until we come to consider the phenomena of cue competition because cue-competition experiments highlight the need for a rate partitioning process in any timescale-invariant model of acquisition. The only thing that one needs to know about the partitioning process at this point is that when there have been no reinforcements of the background alone, it attributes a zero rate of reinforcement to the background. This is equivalent to estimating the interval between background reinforcements to be infinite, but the estimate of an infinite interval between events can never be justified by a finite period of observation. A fundamental idea in our theory of acquisition is that a failure to observe any background reinforcements during the initial exposure to a conditioning protocol should not and does not justify an estimate of zero for the rate of background reinforcement. It only justifies the conclusion that the background rate is no higher than the reciprocal of the total exposure to the background so far. Thus, RET assumes that the estimated rate of background reinforcement when no reinforcement has yet been observed during any intertrial interval is $1/\hat{t}_I$, where $\hat{t}_I$ is the subjective measure of the cumulative intertrial interval (the cumulative exposure to the background alone; see consistency check in Figure 13). (See Table 2 for definitions of the symbols used in the exposition of RET.)

Correcting the background rate estimate delivered by the partitioning process in the case in which there has been no background USs adapts the decision process to the objective uncertainty inherent in a finite period of observation without an observed event. (In other words, it recognizes that absence of evidence is not evidence of absence.) Note that this correction is consistent with partitioning in later examples in which reinforcements are delivered in the intertrial interval. In those cases, the estimated rate of background reinforcement, $\hat{\lambda}_b$, is always $\hat{n}_I/\hat{t}_I$, the cumulative number of background reinforcements divided by the cumulative exposure to the background alone.

As conditioning proceeds with no reinforcers in the intertrial intervals, $\hat{t}_I$ gets longer and longer, so $1/\hat{t}_I$ gets smaller and smaller. When the ratio of the rate expected during the CS and the background rate exceeds a threshold, conditioned responding appears. Thus, conditioned responding makes its appearance when

$$\frac{\hat{\lambda}_{CS} + \hat{\lambda}_b}{\hat{\lambda}_b} > \beta,$$

Table 2
Symbols and Expressions in Rate Estimation Theory of Acquisition

Symbol or expression	Meaning
T	Duration of a conditioned stimulus (CS) presentation, which is equal to the reinforcement latency in delay conditioning
I	Intertrial interval
I/T	Ratio of the intertrial interval to the trial duration
$\hat{t}_{\mathrm{CS}}$	Cumulative exposure to the CS
$\hat{t}_{\mathrm{I}}$	Cumulative intertrial interval
$\hat{n}_{\mathrm{CS}}$	Cumulative number of reinforcements while CS was present (CS reinforcements)
$\hat{n}_{\mathrm{I}}$	Cumulative number of intertrial reinforcements
$\hat{\lambda}_{\mathrm{CS}}$	Rate of reinforcement attributed to a CS
$\hat{\lambda}_{\mathrm{b}}$	Estimated rate of background reinforcement
$\dfrac{\hat{\lambda}_{\mathrm{CS}} + \hat{\lambda}_{\mathrm{b}}}{\hat{\lambda}_{\mathrm{b}}}$	Decision variable in acquisition, ratio of rate of reinforcement when CS is present to rate of background reinforcement
N	Number of CS reinforcements required for acquisition

Note. A hat on a variable indicates that it is a subjective estimate. A symbol without a hat refers to a physically measurable variable.

where β is the threshold or decision criterion. Assuming that the animal's estimates of numbers and durations are proportional to the true numbers and durations (i.e., that subjective number and subjective duration, represented by the symbols with hats, are proportional to objective number and objective duration, represented by the same symbols without hats), we have

$$\hat{\lambda}_{\mathrm{CS}} + \hat{\lambda}_{\mathrm{b}} = n_{\mathrm{CS}}/t_{\mathrm{CS}} \text{ and } \hat{\lambda}_{\mathrm{b}} = n_{\mathrm{I}}/t_{\mathrm{I}},$$

so that (by substitution) conditioning requires that

$$\frac{n_{\mathrm{CS}}/t_{\mathrm{CS}}}{n_{\mathrm{I}}/t_{\mathrm{I}}} > \beta.$$

Equivalently (by rearrangement), the ratio of CS reinforcers to background reinforcers, $n_{\mathrm{CS}}/n_{\mathrm{I}}$, must exceed the ratio of the cumulated trial time to the cumulated intertrial (background alone) time by some multiplicative factor,

$$\frac{n_{\mathrm{CS}}}{n_{\mathrm{I}}} > \beta\,\frac{t_{\mathrm{CS}}}{t_{\mathrm{I}}}. \tag{1}$$

It follows that N, the number of CS reinforcements required for conditioning to occur in simple delay conditioning, must be inversely proportional to the I/T ratio. The left-hand side of Equation 1 is equal to N because, by the definition of N, the CR is not observed until $n_{\mathrm{CS}} = N$, and n_{I} is implicitly taken to be 1 when the estimated rate of reinforcement is taken to be $1/t_{\mathrm{I}}$. On the right-hand side of Equation 1, the ratio of cumulated intertrial interval time (cumulative exposure to the background alone $= t_{\mathrm{I}}$) and the cumulated CS time (t_{CS}) is, on average, the I/T ratio. Thus, conditioned responding to the CS should begin when

$$n_{\mathrm{CS}} > \beta(I/T)^{-1}. \tag{2}$$

Equation 2 means that, on average, the number of trials to acquisition should be the same in different protocols with different durations for I and T but the same I/T ratio. It also implies that reinforcements to acquisition should be inversely proportional to the I/T ratio.

In Figure 9, which is replotted from Gibbon and Balsam (1981), data from a variety of studies show that this inverse proportionality between reinforcements to acquisition and the I/T ratio is only approximately what is in fact observed. The slope of the best fitting line through the data in Figure 9 is $-.72 \pm .04$, which is significantly less than the value of -1 (99% confidence limit $= -.83$), which means that there is a linear rather than strictly proportional relation. The fact that the slope is close to 1 indicates, however, that the relation can be regarded as approximately proportional.

The derivation of a linear (rather than proportional) relation between $\log N$ and $\log(I/T)$ and of the scalar variability in reinforcements to acquisition (the constant vertical scatter about the regression line in Figure 9) is given in Appendix A. Intuitively, it rests on the following idea: N is the CS presentation (trial) at which subjects first reach the acquisition criterion. This means that for the previous $N - 1$ trials, this criterion was not exceeded. Because there is noise in the decision variable, for any given average value of the decision variable that is somewhat less than the decision criterion, there is some probability that the actually sampled value on a given trial will be greater than the criterion. Thus, there is some probability that noise in the decision variable will lead to the satisfaction of the acquisition criterion during the period when the average value of the variable remains below criterion. The more trials there are during the period when the average value of the decision variable is close to but still below the decision criterion, the greater the likelihood of this happening. In probabilistic terms, conditioning requires N to be such that $N - 1$ failures to cross threshold precede it, and this occurs with probability $\prod_{k=1}^{k=N-1} P_k$, where P_k is the probability of failure on the kth trial. As N increases, the chance of $N - 1$ failures before the first success becomes smaller; hence, the chance of prematurely exceeding the criterion increases. It is this feature that, in Figure 9, reduces the slope of the N versus I/T function below -1, which is the value predicted by Equation 2.

The important conclusion to be drawn from Figure 9 is that the speed of conditioning is constant at constant I/T ratios, as RET predicts, and that the rate of acquisition varies approximately in proportion to the I/T ratio. This accounts for most of the previously listed quantitative findings about acquisition.

1. Effect of trial spacing: Increasing I without increasing T results in a higher I/T ratio, hence more rapid conditioning. RET correctly predicts the form and magnitude of this effect.

2. Effect of delay of reinforcement: Increasing T without increasing I results in a lower I/T ratio, hence slower conditioning. Again, RET correctly predicts the form and magnitude of this effect.

3. Timescale invariance: Increasing I and T by the same factor does not change the rate of conditioning. The points in Figure 9 with the same I/T ratio show approximately equal rates of conditioning, even though the absolute values of I and T differ substantially among points at the same ratio (at the same point along the abscissa; see also Figure 11).

4. No effect of partial reinforcement: When reinforcers are given only on some fraction of the trials, cumulative exposure to the CS

per CS reinforcement increases by the inverse of that fraction, but so does cumulative exposure to the background per CS reinforcement. For example, reinforcing only 1/2 of the trials increases the amount of exposure to the CS per reinforcement by 2 (from T to $2T$). But each T seconds of exposure to the CS is accompanied by I seconds of exposure to the background alone. Doubling the amount of CS exposure per reinforcement doubles the amount of background-alone exposure per CS reinforcement as well. Therefore, the ratio of these two cumulative exposures (t_{CS} and t_I) after any given number of reinforcements remains unchanged. No decrement in rate of acquisition should be seen, and none is, indeed, found. In RET, this very important experimental result is another manifestation of the timescale invariance of conditioning because partial reinforcement does not change the relative amounts of CS exposure and background exposure per reinforcement.

5. No effect of reinforcement magnitude: When reinforcement magnitude is increased, it increases the estimated rate of reinforcement[3] in both the signal and in the background by the same factor; hence, these changes in reinforcement magnitudes cancel, leaving I/T unchanged. Again, no improvement in rate of acquisition is expected, and none is found. If there were a contrast between the magnitude of reinforcements given during the intertrial intervals and the magnitude given during the CS, then RET predicts that the ratio of these contrasting reinforcement magnitudes would strongly affect rate of acquisition. However, when there are no reinforcements during the intertrial intervals (the usual case), RET predicts that varying magnitudes of reinforcement will have no effect because the "consistency check" stage in the computation of the decision variable implicitly assumes that the yet-to-occur first background reinforcement will have the same magnitude as the reinforcements so far experienced.

6. Acquisition variability: The data points in Figure 9 show an approximately constant range of vertical scatter about the regression line in log–log coordinates. In the model of acquisition just presented, this scalar variability in reinforcements to acquisition results from the increasing variability in the estimate of t_I, the total accumulated intertrial time, in comparison with the relatively stable variability in the estimate of the average interval of CS exposure between reinforcements, $t_{\mathrm{CS}}/n_{\mathrm{CS}}$. Intuitively, the estimated interreinforcement interval in the presence of the CS ($1/[\hat{\lambda}_{\mathrm{CS}} = \hat{\lambda}_b]$) becomes increasingly stable as n_{CS} increases, whereas the sampling noise in the estimate of the background interreinforcement interval gets greater in proportion as that estimate gets larger (scalar variability). Because of the scalar property, the variability in the estimate of N in Equation 2 is proportional to its size, hence constant on the log scale. The basic threshold prediction and its expected variance are detailed in Appendix A.

Summary of Acquisition

Most of the presently known quantitative facts about the rate of acquisition follow directly from the assumption that the animal begins to respond to the CS when the ratio of two rate estimates exceeds a criterion: The numerator of the ratio is the subject's estimate of the rate of reinforcement in the presence of the CS. The denominator is the estimate of the background rate of reinforcement. The ratio may be thought of as the subject's measure of how similar the rate of CS reinforcement is to the rate of background reinforcement. In simple conditioning, when the background alone is never reinforced, the denominator is the reciprocal of the cumulative duration of the interval between trials, whereas the numerator is the rate of reinforcement when the CS is present. If the decision ratio is taken to be a ratio of expected interreinforcement intervals, then the predictions follow from the assumption that conditioned responding begins when the expected interval between background reinforcements exceeds the expected interval between CS reinforcements by a threshold factor. These are equivalent formulations (the interval–rate duality principle).

Acquisition of a Timed Response

There is no CR to the CS until the *whether* criterion has been met. The timing of the responses that are then observed is known to depend, at least eventually, on the distribution of reinforcement latencies that the animal observes. It is this dependence that is modeled by SET, which models the process leading to a CR under well-trained conditions, in which the animal has decided (earlier in its training) that the CS merits a response (the *whether* decision) what the appropriate comparison interval for that particular response is and what the appropriate threshold value is. A model for the acquisition of an appropriately timed CR is needed to describe the process by which these latter decisions are made during the course of training, because SET presupposes that these decisions have already been made. It models only mature responding, the responding observed once comparison intervals and thresholds have been decided.

It is tempting to assume that no such decisions are necessary, that the animal simply samples from the distribution of remembered intervals to obtain the particular remembered interval that constitutes the denominator of the decision ratios in SET on any one trial. This would predict exponentially distributed response latencies in experiments in which the observed CS–US intervals are exponential, and normally distributed response latencies in cases in which there is a single, fixed CS–US interval. We are inclined to doubt that this assumption would survive detailed scrutiny of the distributions actually observed and their evolution over the course of training, but we are not aware of published data of this kind. Consider an experiment in which a rat has come to fear a shock that occurs at some random but low rate when a CS is present (e.g., as in the background conditions of Rescorla, 1968). The shock delays after CS onset are exponentially distributed, and this distribution is so shallow that it is common for shocks not to occur for many minutes. It seems unlikely that onset of the rat's fear response is ever delayed by many minutes after the onset of the CS under *these* conditions, in which the shock is equally likely at any moment! But this is what one has to predict if it is assumed that the rat simply samples from the distribution of remembered latencies. Also, casual observation of training data from the peak procedure suggests that the termination of conditioned responding to the CS when the expected reinforcement latency has passed develops later in training than does the delay of anticipatory

---

[3] Rate is now used to mean the *amount* of reinforcement per unit of time, which is the product of reinforcement magnitude and number of reinforcements per unit of time. Later, when it becomes important to distinguish between the *number* of reinforcements per unit of time and the magnitudes of those reinforcements, we call this "income" rather than rate. It is the same quantity as expectancy of reinforcement, H, in Gibbon (1977).

responding (cf. Rescorla, 1967). This implies that it takes longer (more training experience) to decide on an appropriate stop threshold than to decide on an appropriate start threshold.

The need to posit timing-acquisition processes by which the animal decides in the course of training on appropriate comparison intervals (and perhaps also on appropriate decision thresholds) becomes even clearer when one considers more complex paradigms such as the time-left paradigm with one very short and one very long standard interval. In this paradigm, the decision to switch from the standard side to the time-left side uses the harmonic mean of the two standard intervals as the comparison value (the denominator in the decision variable). However, on those trials in which the subject does not switch to the time-left side before the moment of commitment, and thereby ends up committed to the standard delays, one observes the effects of three more timing decisions. After the moment when the program has committed the subject to the standard side, and hence one of the two standard delays, the likelihood of responding rises to a peak at the time of the first standard delay (first start decision); if food is not delivered then, it subsides (first stop decision), to rise to a second peak at the time of the second latency (second start decision). Thus, in this experiment, three different reference intervals (expectations) are derived from one and the same experienced distribution (the distribution of delays on the standard side)—one expectation for the changeover decision, one for the decision that causes the early peak in responding on the standard side, and one for the decision that causes the late peak. Clearly, an account is needed of how, in the course of training, the animal decides on these three different reference intervals and appropriate thresholds. There is no such account at present. Its development must await data on the emergence of timed responding (i.e., appropriate acquisition data).

A related issue concerns the acquisition of the CR in trace-conditioning paradigms. In these paradigms, the US does not occur during the CS but rather some while after the termination of the CS. Thus, the onset of the CS does not predict an increase in the rate of US occurrence. Rather, the offset of the CS predicts that a US will occur after a fixed latency. For acquisition of a response to the CS to occur under these conditions, the animal must decide that the latency from CS onset to the US is appreciably much shorter than the US–US latency. As in the acquisition of a timed response, this would seem to require a decision process that examines the distribution of USs relative to a time marker.

Extinction

Associative models of conditioning are event-driven; changes in associative strengths occur in response to events. Extinction is the consequence of nonreinforcements, which are problematic "events," because a nonreinforcement is the failure of a reinforcement to occur. If there is no defined time when a reinforcement ought to occur, then it is not clear how to determine when a nonreinforcement has occurred. In RET, this problem does not arise because extinction is assumed to occur when a decision variable involving an elapsing interval exceeds a decision criterion. The decision variable is the ratio of the currently elapsing interval without a reinforcement to the expected interreinforcement interval. Before elaborating, we list some of the salient empirical facts about extinction, against which different models of the process may be measured:

Extinction Findings

Weakening of the conditioned response with extended experience of nonreinforcement. It takes a number of unreinforced trials before the CR ceases. How abruptly it ceases in individual subjects has not been established. That is, the form of the psychometric extinction function in individual subjects is not known.

Partial-reinforcement extinction effect. Partial reinforcement during the original conditioning increases trials to extinction, the number of unreinforced trials required before the animal stops responding to the CS. However, the increase is proportional to the thinning of the reinforcement schedule (Figure 14B, solid lines); hence, it does not affect the number of reinforcements that must be omitted to produce a given level of extinction (Figure 14B, dashed lines). Thus, both delivered reinforcements to acquisition and omitted reinforcements to extinction are little affected by partial reinforcement.

No effect of I/T ratio on rate of extinction. The *I/T* ratio has no effect on the number of reinforcements that must be omitted to reach a given level of extinction (Figure 14B, dashed lines) and, hence, no effect on trials to extinction (Figure 14B, solid lines). This lack of effect on the rate of extinction contrasts strikingly with the strong effect of the same variable on the rate of acquisition (Figure 14A). As in the case of acquisition, this result is best established in the case of pigeon autoshaping, but it appears to be generally true that partial reinforcement during acquisition has little effect on the number of reinforcements that must be omitted to produce extinction (for an extensive tabulation of such results, see Gibbon et al., 1980).

Rates of extinction may be equal to or faster than rates of acquisition. After extensive training in an autoshaping paradigm, the number of reinforcements that must be omitted to reach a

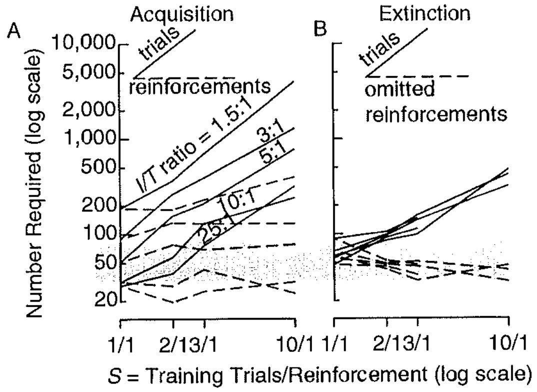

Figure 14. Effect of the *I/T* (intertrial interval/trial duration) ratio and the reinforcement schedule during training on acquisition and extinction of autoshaped pecking in pigeons. A: Reproduced from Figure 8. B: Partial reinforcement during training increases trials to extinction in proportion to the thinning factor (*S*); hence, it has no effect on omitted reinforcements to extinction. The *I/T* ratio, which has a strong effect on reinforcements to acquisition, has no effect on omitted reinforcements to extinction. This figure is based on data in Gibbon, Farrell, Locurto, Duncan, and Terrace (1980) and Gibbon, Baldock, Locurto, Gold, and Terrace (1977).

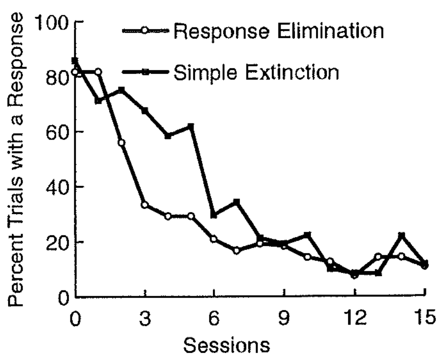

Figure 15. Rate of response elimination under random reinforcement compared with rate of simple extinction. Two groups ($N = 6$) received 15 sessions of autoshaping training with a rate of conditioned-stimulus (CS) reinforcement of one unconditioned stimulus (US) per 30 s, a CS duration of 12 s, and a mean intertrial interval of 132 s (thus, the ratio of I/T [intertrial interval/trial duration] = 10). Data shown are from subsequent sessions of extinction (no USs at all) or random control (USs delivered at an average rate of one per 30 s during both intertrial intervals and CS presentations). The data are from Aronson, Balsam, and Gibbon (1991).

modest extinction criterion (a decline to 50% of the preextinction rate of responding) is roughly the same as the number of reinforcements required to reach a modest acquisition criterion (one peck in three out of four CS presentations), when the I/T ratio during acquisition is in the most commonly used range (see the gray band across Figures 14A and 14B). If the I/T ratio is smaller than usual, then acquisition takes many more reinforcements than must be omitted to produce extinction. Under these conditions, the rate of extinction is faster than the rate of acquisition.

Rates of extinction are comparable to rates of response elimination. After extensive training in the autoshaping paradigm, the decline of responding in ordinary extinction proceeds at about the same rate as the decline of responding when response elimination is produced not by withholding reinforcement of the CS but rather by introducing the truly random control, that is, by shifting from a zero rate of background reinforcement to a rate of background reinforcement that equals the rate observed when the CS is present (Durlach, 1986)—see Figure 15.

Model

The conceptual framework that we are elaborating attributes changes in conditioned behavior to decisions based on decision variables that grow with the duration of relevant experience. Central to this perspective is the assumption that different decisions are based on different decision variables. Thus, the acquisition of a response to a CS that predicts a higher rate of reinforcement is based on a different comparison than is the timing of a response to a CS that predicts a reinforcement after some fixed latency. Similarly, the disappearance of the CR to a CS in the course of extinction is assumed to be based on yet another comparison, a comparison between the cumulative amount of unreinforced CS exposure since the last CS reinforcement and the expected amount of CS exposure per reinforcement. In the version of RET presented here, extinction occurs when

$$\frac{\hat{I}_{\mathrm{CS\ no\ R}}}{I\hat{R}I_{\mathrm{CS}}} = \frac{\hat{\lambda}_{\mathrm{CS}}}{\hat{\lambda}_{\mathrm{CS\ no\ R}}} > \beta,$$

that is, when the ratio of the estimated (subjective) interval that has elapsed since the last reinforcement ($\hat{I}_{\mathrm{CS\ no\ R}}$) to the expected interval between CS reinforcements ($I\hat{R}I_{\mathrm{CS}}$) is greater than a threshold factor, β. By the principle of interval–rate duality, the decision variable may also be thought of as the ratio between the estimated rate of CS reinforcement up to the last reinforcement ($\hat{\lambda}_{\mathrm{CS}}$) and the estimated rate since then ($\hat{\lambda}_{\mathrm{CS\ no\ R}}$). The model is developed more formally and with a treatment of variance issues in Appendix A. (See Table 3 for definitions of the symbols used in the exposition of the extinction model.)

Weakening of the conditioned response. The decision variable grows in proportion to the duration of extinction, because the longer extinction lasts, the more unreinforced exposure to the CS accumulates. Equivalently, the subject's estimate of the rate of CS reinforcement since the last such reinforcement becomes an ever smaller fraction of its previous estimate. This explains why it takes some number of omitted reinforcements to produce extinction. From this viewpoint, it is a *relatively* prolonged interval without reinforcement that leads to extinction, rather than the "occurrence" of a nonreinforcement. *Occurrence* is in quotation marks because a nonreinforcement is an event only in the mind of the experimenter or the theoretician (and, possibly, in the mind of the subject); it is not an event in the ordinary physical sense. *Relatively* is italicized because it is the expected interval between reinforcements that determines what constitutes a long interval without reinforcement. Here, as everywhere in our models, the only intervals that matter are relative intervals (the principle of timescale invariance).

The decision process mediating extinction operates in parallel with the decision process mediating acquisition because it is a general process that animals use to detect changes in the rates of significant events. Before the experimenter implements extinction, the subject has been checking all along for a change in the rate of CS reinforcement, but this check comes up negative (in favor of the no-change hypothesis) until some time into the extinction phase of the experiment.

Table 3
Symbols and Expressions in Rate Estimation Theory of Extinction

Symbol or expression	Meaning
$\hat{I}_{\mathrm{CS\ no\ R}}$	Amount of conditioned-stimulus (CS) exposure since the last reinforcement credited to the CS
$I\hat{R}I_{\mathrm{CS}}$	Expected interval between CS reinforcements equals $1/\hat{\lambda}_{\mathrm{CS}}$
$\hat{\lambda}_{\mathrm{CS\ no\ R}}$	Estimated rate of CS reinforcement since the last reinforcement credited to the CS equals $1/\hat{I}_{\mathrm{CS\ no\ R}}$
$\frac{\hat{I}_{\mathrm{CS\ no\ R}}}{I\hat{R}I_{\mathrm{CS}}} = \frac{\hat{\lambda}_{\mathrm{CS}}}{\hat{\lambda}_{\mathrm{CS\ no\ R}}}$	Decision variable in extinction

Note. A hat on a variable indicates that it is a subjective estimate, a quantity in the head representing a physically measurable external variable. If the hat is omitted, then the symbol refers to the physically measurable external variable.

The partial-reinforcement extinction effect. The effect of partial reinforcement on trials to extinction follows directly from the assumption that the decision to stop responding is based on the ratio of the currently accumulating interval of unreinforced CS exposure to the expected amount of CS exposure per reinforcement. Thinning the reinforcement schedule during training by a given factor increases the expected amount of CS exposure per reinforcement by that factor. For example, when a 10:1 schedule of reinforcement is used during acquisition (an average of 10 presentations of the CS per reinforcement), the expected amount of CS exposure per reinforcement is $10T$ rather than simply T, where T is trial duration. In our model of extinction, the expected interval between reinforcement is the denominator of the decision variable, the quantity against which the animal compares the currently elapsing period of unreinforced CS exposure. For the decision variable to reach a criterion (fixed-threshold) value during extinction, the numerator must increase by the same factor as the denominator. Hence, the number of trials required to reach the extinction decision threshold must increase in proportion to the thinning of the partial-reinforcement schedule. This increase in trials to extinction is the partial-reinforcement extinction effect.

Because the required amount of unreinforced CS exposure increases in proportion to the thinning factor, the number of reinforcements that must be omitted to reach a given level of extinction remains constant. In other words, there is a partial-reinforcement extinction effect only if one looks at trials to extinction (or required amount of unreinforced CS exposure). From a timing perspective, this is not an appropriate unit. The appropriate unit is omitted reinforcements to extinction (or the relative amount of CS exposure). If one takes the number of omitted reinforcements to extinction as the appropriate unit of analysis, there is no partial-reinforcement extinction effect (Figure 14). Notice that the prediction that trials to extinction must increase in proportion to the thinning of the reinforcement schedule does not depend on parametric assumptions. The proportionality between trials to extinction and the preextinction schedule of reinforcement is another manifestation of timescale invariance.

If one assumes that reinforcement strengthens and nonreinforcement weakens associative connections, then it is paradoxical that partial reinforcement has no effect on reinforcements to acquisition even though the rate of extinction (the amount of weakening per nonreinforcement) may be as great as or greater than the rate of acquisition (the amount of strengthening per reinforcement). The paradox vanishes if one thinks in terms of ratio comparison of expected intervals between reinforcements. From that perspective, partial reinforcement has no effect on reinforcements to acquisition because it proportionately extends both exposure to the CS and exposure to the background, leaving their ratio unchanged. And partial reinforcement has no effect on omitted reinforcements to extinction because the same number of nonreinforcements produces any given ratio between the expected interreinforcement interval and the interval without reinforcement. In both cases, the lack of an effect is a manifestation of the underlying timescale invariance of the conditioning process. This is obscured by the associative analysis because in that framework, the focus is on trials (CS presentations) rather than the ratios of accumulated intervals (or, equivalently, the ratios of estimated rates of reinforcement).

The I/T ratio and the rate of extinction. The I/T ratio has no effect on the rate of extinction because the currently cumulating interval of unreinforced CS exposure is not compared with the expected interval between background reinforcements; it is compared with the expected interval between CS reinforcements. The striking difference between the effect of the I/T ratio on acquisition and its lack of effect on extinction follows from a fundamental aspect of timing theory—different decisions rest on different comparisons (cf. Miller's comparator hypothesis; Cole et al., 1995a; Miller et al., 1992). Different decisions rest on different comparisons because different comparisons are appropriate to detecting different properties of the animal's experience. The background rate of reinforcement is the appropriate comparison when the effect of a CS on the rate of reinforcement is in question; the previous rate of CS reinforcement is the appropriate comparison when the question is whether the rate of CS reinforcement has changed. This is a general feature of modeling within a decision-theoretic framework, not just of our models. It is taken for granted, for example, in psychophysical modeling that different judgments about the same stimulus are based on different decision variables.

Rates of extinction and rates of acquisition. Rates of extinction may be as fast as or faster than rates of acquisition because rate of acquisition depends on the I/T ratio, whereas rate of extinction does not. When an unusually low I/T ratio is used during acquisition, the number of reinforcements that must be delivered is high, whereas the number of reinforcements that must then be omitted to produce extinction is unaffected. This makes it possible for the rate of extinction to be faster than the rate of acquisition, even though partial reinforcement has no effect. These two findings are irreconcilable in most associative models because the only way to minimize (not eliminate!) the effect of partial reinforcement on acquisition is to make the weakening effect of nonreinforcement (hence, the rate of extinction) much less than the strengthening effect of reinforcement (hence, the rate of acquisition).

Rate of response elimination and rate of extinction. The rate of response elimination is comparable to the rate of extinction because when the background rate of reinforcement is raised to match the rate of CS reinforcement, the partitioning scheme (presented below) immediately stops crediting to the CS reinforcements that occur during the CS. Thus, the input to the rate-change decision is the same in response elimination as in simple extinction: In both cases, there is an abrupt cessation of reinforcements credited to the CS. (See Figure 21, which is presented later in the Cue Competition section.)

Conjoining Decisions

We have argued for three distinct decision processes mediating three distinct conditioning phenomena—the decision when to respond to a CS that predicts reinforcement at a fixed latency, the decision whether to respond to a CS that predicts a change in the rate of reinforcement, and the decision to stop responding to a CS when it no longer predicts a change in the rate of reinforcement. Moreover, it seems necessary to assume a fourth decision process, which we have not modeled—a process that decides whether there is one or more (relatively) fixed latencies of reinforcement, as opposed to a random distribution of reinforcement latencies. This fourth process mediates the acquisition of a timed response.

We assume that the computations underlying these decision processes operate in parallel. In other words, in common with most perceptual modeling, we assume that the nervous system is simultaneously processing the stimulus (the conditioning protocol) in several different ways. However, the manifestation of these different types of stimulus processing in conditioned behavior often depends on a conjunction of decisions. For example, in the peak procedure, in which the subject responds to the CS but only in a period surrounding the expected reinforcement time, the response to the CS is not seen unless the subject has decided that (a) the rate of reinforcement during the CS is substantially higher than in its absence, (b) there is a relatively fixed latency between CS onset and reinforcement, and (c) the time that has so far elapsed on this particular trial is approximately equal to the expected latency of reinforcement.

Cue Competition

The modern era in the study of conditioning began with the discovery that conditioning to one CS does not proceed independently of the conditioning that occurs to other CSs in the protocol. The more or less simultaneous discovery of the effects of background conditioning (Rescorla, 1968), blocking and overshadowing (Kamin, 1967, 1969a), and relative validity (Wagner, Logan, Haberlandt, & Price, 1968) made the point that what an animal learns about one CS strongly affects what it learns about other CSs. Alternatively, the behavioral expression of what it has learned about one CS is affected by what it has learned about other CSs (cf. Miller's comparator hypothesis; Cole, Barnet, & Miller, 1995b; Miller et al., 1992). The discovery of cue competition led to the Rescorla–Wagner (Rescorla & Wagner, 1972) model and to other contemporary associative models that explain how experience with one stimulus affects the observed conditioning to another stimulus (e.g., Mackintosh, 1975; Miller & Matzel, 1989; Pearce, 1994; Pearce & Hall, 1980; Wagner, 1981).

Some well-established facts concerning cue interactions during acquisition are as follows:

- Blocking and the effect of background conditioning: Blocking is said to occur when one CS is presented alone at least some of the time and together with a second CS some of the time. If the rate of reinforcement during presentations of the first CS is unaffected by the presence or absence of the second CS, then the second CS does not get conditioned no matter how often it is reinforced (i.e., paired with the US).
- Overshadowing: If two CSs are always presented and reinforced together, a CR generally develops much more strongly to one CS than to the other (Kamin, 1967, 1969a, 1969b; Mackintosh, 1971, 1976; Reynolds, 1961)—see Figure 16.
- Relative validity: When one CS (called the "common cue") occurs in combination with a second CS on some trials but in combination with a third CS on other trials, the CS that gets conditioned is the one that can, by itself, predict the observed pattern of US occurrences. In other words, the relatively more valid cue gets conditioned (Wagner et al., 1968)—see Figure 17.
- One-trial overshadowing: The competitive exclusion of one CS by another CS in the overshadowing protocol is manifested after a single conditioning trial (James & Wagner, 1980; Mackintosh & Reese, 1970).
- Retroactive reversal of overshadowing and blocking: Subjects do not respond to an overshadowed CS if they are tested before the overshadowing CS is extinguished, but they do respond to it if they are tested after the overshadowing CS is extinguished (Kaufman & Bolles, 1981; Matzel, Schachtman, & Miller, 1985; see Baker & Mercier, 1989, for a review). Thus, extinction of the overshadowing CS retroactively removes the overshadowing. Sometimes, the complementary effect, retroactive blocking, is also obtained. Subsequent reinforcement of an overshadowed CS retroactively blocks the overshadowing CS (Cole et al., 1995a). Retroactive blocking is reliably obtained only in sensory preconditioning protocols, in which the stimulus paired with a CS is another CS rather than a conventional US. It is not generally obtained when conventional CS–US overshadowing protocols are used (Grahame, Barnet, & Miller, 1992).

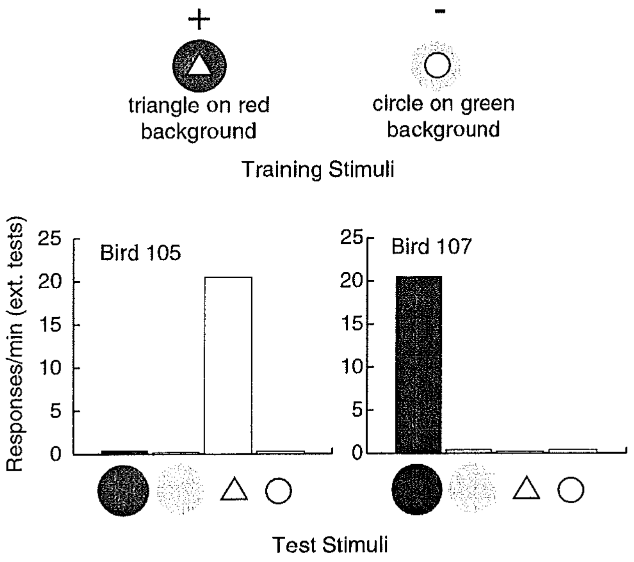

Figure 16. Overshadowing in operant appetitive conditioning. Two pigeons were trained to peck the key with a triangle on a red background and not the key with a circle on a green background. When tested with the four stimulus elements (red, green, triangle, circle) separately, 1 bird responded entirely to the red, whereas the other bird responded entirely to the triangle, although both of these elements had, of course, been paired with reinforcement throughout training. The data are from Reynolds (1961). Overshadowing is often not as complete as in this example. ext. = extinction (no reinforcement).

- Inhibitory conditioning: When a CS predicts the omission of reinforcement, a CR develops that is more or less the antithesis of the "excitatory" response (e.g., avoidance of the CS rather than approach to it). Such a CS is said to be an inhibitory CS (or CS−). When a CS− is paired with an excitatory CS (CS+), the CR elicited by the CS+ is reduced or abolished. This is called the "summation test" for conditioned inhibition.

Partitioning and Predictor Minimization

In RET, cue competition is explained by properties of the partitioning process, the process that credits rates of reinforcement to the different CSs presented during conditioning. Among the CSs that may be credited with some rate of reinforcement is the background (the experimental chamber), which is simply treated as

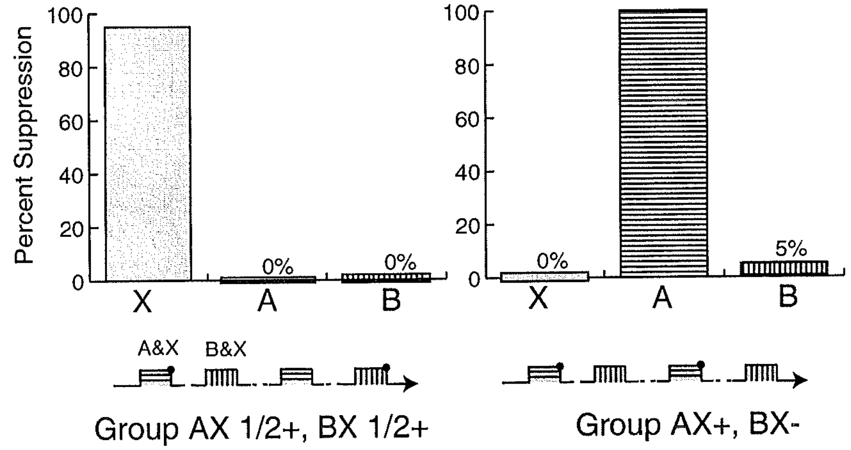

Figure 17. The effect of relative validity on conditioned emotional response conditioning in rats. Although footshock (dots) was paired with both X and A on 50% of all presentations, X was the more valid predictor in Group AX 1/2+, BX 1/2+, whereas A was the more valid predictor in Group AX+, BX−. When responses to the individual stimuli were tested, only the more valid conditioned stimulus elicited a conditioned response (fear-induced suppression of appetitive responding). These are the Stage I data from Wagner, Logan, Haberlandt, and Price (1968, Table 3, p. 175). Not all of the results in that article showed the all-or-nothing property that one sees in these data.

another CS, with no special status. The properties of the partitioning process are largely derived from a simple principle—rate additivity—which is implicit in the structure of the mechanism that computes the rates of reinforcement credited to each of the experimentally manipulated stimuli (Gallistel, 1990). Some properties of the partitioning process depend on a second principle—the principle of parsimony (or predictor minimization). This second principle comes into play only in those cases in which the principle of rate additivity does not determine a unique solution to the rate estimation problem. In such cases, the principle of parsimony (Occam's razor) minimizes the number of predictors.

The structure of the partitioning process in RET is entirely determined by these two principles because all of the explanations in RET that depend on partitioning are mathematically derivable consequences of them. (Figure 18 shows a flow diagram for the computational structure implied by these two principles.) Mathematical derivations of the predictions of these principles are given in Gallistel (1990, 1992a), and a spreadsheet implementation of the model is given in Gallistel (1992b). However, for one to understand predictions of the partitioning model in the case of most of the classic cue-competition protocols, no mathematics are really necessary. To make these explanations as intuitively accessible as possible, we avoid mathematical derivation in what follows, referring the mathematically inclined reader to the aforementioned citations and to Appendix B.

Blocking

We begin our discussion of the partitioning mechanism with its application to blocking. In a blocking protocol, an independently conditioned CS is combined on some trials with the target CS, without a change in the expected interreinforcement interval. Because the rate of reinforcement when the target CS is present does not differ from the rate when the blocking CS is presented alone, the additive combination of expected rates requires that the reinforcement rate attributed to the target stimulus be zero (see the partitioning box in Figure 19). Hence, the target stimulus does not acquire the power to elicit a CR no matter how often it is paired with the reinforcer.

RET predicts that if the US magnitude is changed at the same time that the second CS is introduced, then a rate—or, rather, an income—will be attributed to that CS (Dickinson, Hall, & Mackintosh, 1976). An *income* is a rate of reinforcement (number of reinforcements per unit of time) multiplied by a reinforcement magnitude. If the reinforcement magnitude goes up when the second CS is introduced, then income goes up, and the newly introduced CS is credited with that increase. That is, it is credited with the amount of income not predicted by the first CS. If the reinforcement magnitude goes down when the new CS is introduced, then it is credited with a negative effect on income. We deal with incomes at much greater length in the Operant Choice section. So far, we have been ignoring reinforcement magnitude, because in most conditioning experiments it does not vary. However, RET generalizes in a straightforward way to the case in which it is income (rate multiplied by magnitude) that must be predicted rather than simply rate. No new assumptions are necessary. Income simply replaces simple rate in the calculations (systems of equations).

There is a close analogy between the explanation of blocking in terms of rate partitioning and its explanation in the Rescorla–Wagner (Rescorla & Wagner, 1972) model, which has been the most influential associative explanation of cue competition. In the Rescorla–Wagner model, associative strengths combine additively, and at asymptote, the sum of the associations to a given US cannot exceed the upper limit on possible net associative strength for that US. In the timing model, estimated rates of reinforcement (or estimated incomes) combine additively, and their sums must equal the observed rates of reinforcement (or incomes). However, in RET, unlike in the Rescorla–Wagner model, the constraint holds

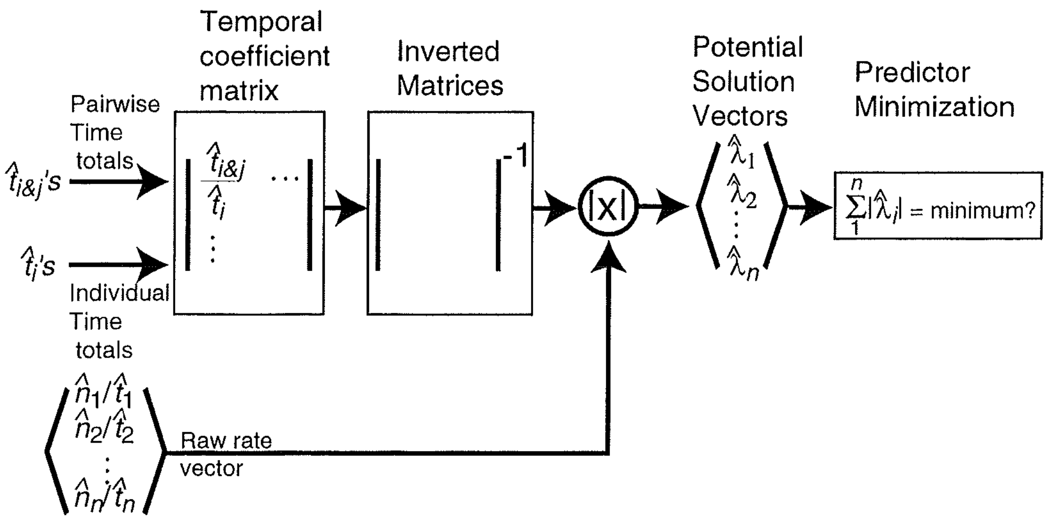

Figure 18. The functional structure of the rate-estimating process. The principle of rate additivity reduces the rate estimation problem to a well-understood problem in linear algebra, the problem of finding the solution to a system of simultaneous linear equations. The coefficients of the matrix that defines the system of equations are the ratios of individual time totals and the pairwise time totals (cumulative amounts of exposure to each conditioned stimulus [CS] in the denominators and cumulative exposure to each pairwise combination of CSs in the numerators). The raw rate vector consists of the rate estimates made by ignoring other CSs and simply dividing the cumulative exposure to each stimulus by the number of reinforcements obtained in the presence of that stimulus. Inverting the temporal coefficient matrix and multiplying it by the raw rate vector gives the corrected rate estimates in all cases in which there exists a unique additive solution. The determinant of the matrix is zero when the system of simultaneous equations implied by the assumption of additivity does not have a solution. This happens whenever there are redundant CSs, in which case there are not as many independent equations as there are rates to be estimated. In such cases, potential solution vectors come from lower order matrices (systems of equations that ignore one or more CSs). The principle of predictor minimization determines which of the lower order solutions is taken as the "correct" solution. It selects the solution that minimizes the sum of the absolute values of the predicted rates. For more details, see Appendix B. $\hat{n}_i$ = subjective number of reinforcements in the presence of the *i*th CS; $\hat{t}_i$ = subjective cumulative exposure to the *i*th CS; $\hat{t}_{i\&j}$ = subjective cumulative exposure to the combination of the *i*th CS and the *j*th CS (regardless of what other CSs may also have been present); $\hat{\lambda}_i$ = subjective rate of reinforcement for the *i*th CS.

at every point in conditioning (not only at asymptote), and it is an external constraint (the observed rates) rather than an internal free parameter (maximum possible associative strength). Indeed, there are no free parameters in the rate estimation process: The theoretically posited rate estimates in the subject's head are completely determined by the observed rates of reinforcement.

Background Conditioning and the Contingency Problem

In the truly random control protocol, the rate of reinforcement when the background alone is present is the same as the rate that is observed when a transient CS, such as a tone or a light, is also present (Figure 20). The principle of rate additivity requires that the background be credited with a rate that explains the rate that is observed when it alone is present. It also requires that the sum of this rate and the rate credited to the CS equals the rate observed when the CS is present. The unique solution to this double constraint ascribes a zero rate of reinforcement to the CS. This explains the profoundly important discovery that conditioning depends on a CS–US contingency rather than on the temporal pairing of the CS and the US (Rescorla, 1968). In RET, the ratio of the rate of reinforcement when the CS is present to the rate of reinforcement when it is absent (the background rate) is the measure of contingency. When this ratio equals one, there is no contingency.

Response Elimination

Recall that responding to a CS may be abolished equally rapidly either by ordinary extinction, in which the CS is no longer reinforced, or by the so-called response-elimination procedure, in which the CS continues to be reinforced but the rate of reinforcement during the intertrial interval (background reinforcement), which has been zero, is now made equal to the rate of CS reinforcement. The partitioning process in RET has the property that the attribution of further reinforcements to the CS ceases as soon as the rate of background reinforcement is raised to equal the rate of CS reinforcement. This is because the rate estimates at every point in conditioning depend only on the totals accumulated up to that point. Thus, it does not matter whether the reinforcements in the background come before or after the CS reinforcements. In Figure 20, the background reinforcements come before the CS reinforcement. In Figure 21, they come afterward, as they would when response elimination begins, but the result is the same: The three reinforcements during the intertrial interval have

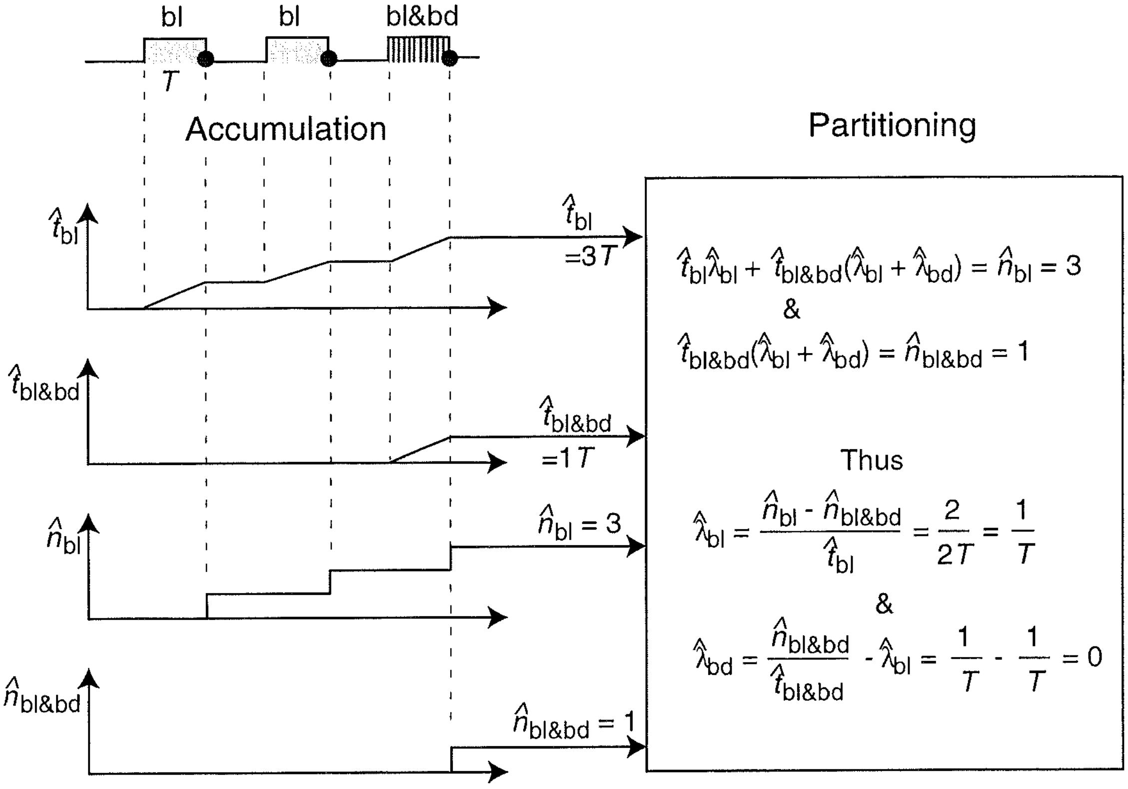

Figure 19. The principle of rate additivity predicts blocking. Rate additivity implies the top two equations in the partitioning box. Solving them gives $1/T$ (trial duration) for $\hat{\lambda}_{bl}$, which is the subjective rate of reinforcement of the blocking (bl) stimulus. Note that any value other than this would fail to account for the rate of reinforcement observed on trials when the blocking stimulus is presented alone. By additivity, the subjective rate of reinforcement attributed to the blocked (bd) stimulus ($\hat{\lambda}_{bd}$) must be the raw rate for that stimulus, which is $1/T$, minus the rate estimate for the blocking stimulus, which is also $1/T$. Hence, the rate of reinforcement attributed to the blocked stimulus must be zero. $\hat{t}_{bl}$ = subjective cumulative exposure to the blocking stimulus; $\hat{t}_{bl\&bd}$ = subjective cumulative simultaneous exposure to the blocking and blocked stimulus; $\hat{n}_{bl}$ = subjective number of reinforcements in the presence of the blocking stimulus; $\hat{n}_{bl\&bd}$ = subjective number of reinforcements in the presence of the combined blocking and blocked stimulus.

the effect of forcing the reinforcement that occurs during the CS to be credited to the background rather than to the CS. Because the commencement of background reinforcement immediately robs all further CSs of credit for the reinforcements that occur in their presence, the onset of background reinforcement marks the beginning of an interval in which no further reinforcements are attributed to the CS. Thus, ordinary extinction and response elimination are two different ways of reducing to zero the apparent rate of CS reinforcement. From the standpoint of the extinction-producing decision process, which compares the currently cumulating amount of apparently unreinforced CS exposure with the expected interval between reinforcements credited to the CS, they are identical. Thus, the decision to stop responding to the CS occurs after the same number of "omitted" reinforcements in both cases, whether those "omissions" are real or apparent.

Signaling Background Reinforcements

Rate additivity also has the consequence that so-called signaled background reinforcers do not affect the estimate of the rate of reinforcement for the chamber alone; hence, signaling background reinforcers eliminates the blocking effect of such reinforcers. A signaled background reinforcer is simply a reinforcer that occurs in the presence of another CS (the signaling CS). Because there are never reinforcements when only the background is present, the background rate of reinforcement must be zero. The zero estimate for the background rate of reinforcement forces the reinforcers that occur during the signaling CS to be credited to it (proof is given in Gallistel, 1990), which is why signaled background reinforcements do not prevent the acquisition of a CR to the target CS (Durlach, 1983; Goddard & Jenkins, 1987). Note that this explanation requires a substantial amount of unreinforced exposure to the background in the absence of any other CSs, which has been shown to be necessary (Cooper, Aronson, Balsam, & Gibbon, 1990). The advantage that RET has over the alternative timing account of blocking and background conditioning offered by Gibbon and Balsam (1981) is that it explains why signaling background reinforcers eliminates the blocking effect of background conditioning.

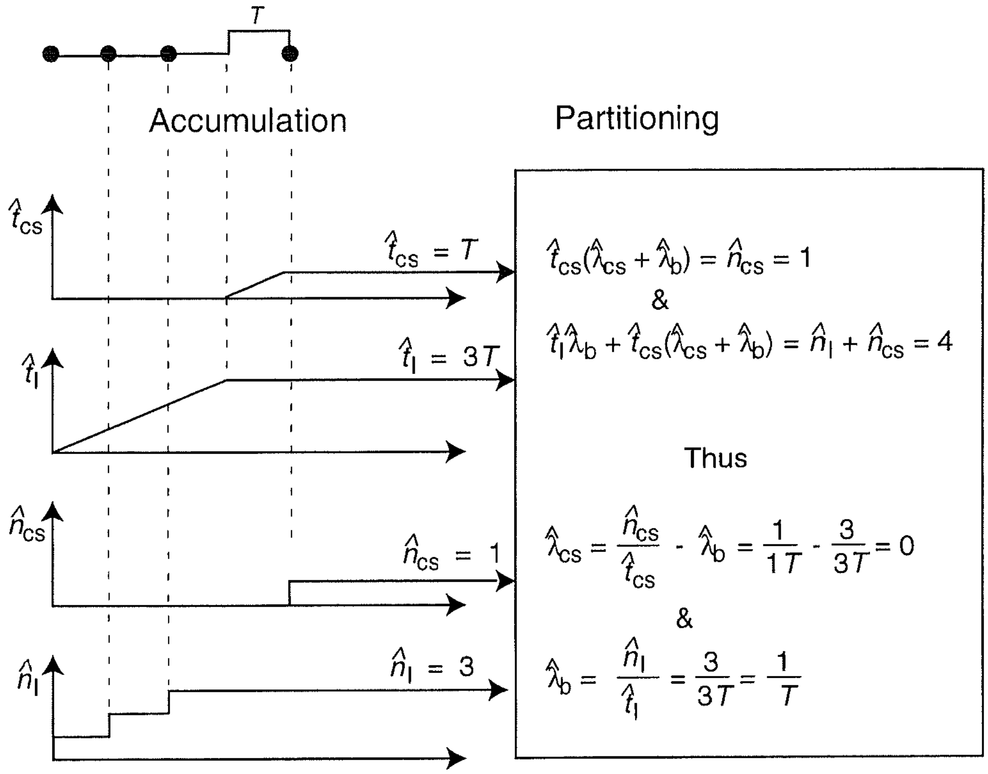

Figure 20. The principle of rate additivity predicts the effect of background conditioning. When the rate of reinforcement during the intertrial interval equals the rate during a conditioned stimulus (CS), the rate estimate for the background equals the raw rate estimate for the CS. When the rate estimate for the background is subtracted from that raw rate estimate for the CS, the resulting estimate for the CS alone is zero. Thus, what matters is not whether the unconditioned stimulus (US) is paired with the CS but whether the rate of US occurrence changes when the CS comes on (CS–US contingency). T = trial duration; $\hat{t}_{CS}$ = subjective cumulative exposure to the CS; $\hat{t}_I$ = subjective cumulative intertrial interval (i.e., cumulative exposure to the background alone); $\hat{n}_{CS}$ = subjective number of CS reinforcements; $\hat{n}_I$ = subjective number of intertrial (background alone) reinforcements; $\hat{\lambda}_{CS}$ = subjective rate of CS reinforcement; $\hat{\lambda}_b$ = subjective rate of background reinforcement.

Overshadowing

The principle of additivity does not always determine a unique solution to the rate estimation problem. In overshadowing protocols, two CSs are always presented together, but the CR develops to one CS and not the other (Figure 16), or, at least, more strongly to one than to the other. When two CSs have always occurred together, any pair of rate estimates that sums to the observed rate of reinforcement is consistent with the additivity constraint. Suppose, for example, that a tone and a light are always presented together for 5 s, after which reinforcement is always given. Reinforcement is never given when these two CSs are not present, so the estimate of the background rate of reinforcement must be 0. One solution to the rate estimation problem credits a rate of 1 reinforcement/5 s to the tone and a rate of 0 reinforcement to the light. This solution is consistent with the principle of rate additivity. But so is the solution that credits 1 reinforcement/5 s to the light and 0 reinforcement to the tone. And so is the solution that credits 0.5 reinforcement/5 s to the light and 0.5 reinforcement/5 s to the tone, and so is every combination of rates that sums to 1 reinforcement/5 s. Thus, the principle of rate additivity does not determine a unique solution to the rate estimation problem in cases in which there are redundant CSs.

The principle of predictor minimization eliminates redundant CSs in such a way as to minimize the number of predictors (CSs) credited with any predictive power, that is, with a nonzero rate of reinforcement. The requirement that a solution to the rate estimation problem must minimize the number of predictors eliminates all those solutions that impute part of the observed rate of reinforcement to one of the two redundant CSs and part to the other. Thus, this principle pares the infinite population of possible additive solutions down to only all-or-none solutions. In the case shown in Figure 16, there does not appear to be, in principle, any nonarbitrary way of deciding which of the two all-or-none solutions should be preferred, so it is not surprising to find that 1 subject credited the triangle whereas the other credited the red background. The very arbitrariness of this outcome suggests that the principle underlying overshadowing is the elimination of redundant predictors. The situation is, we believe, analogous to the situation with ambiguous figures, stimuli that support two or more mutually exclusive percepts. The perceptual system resolves the

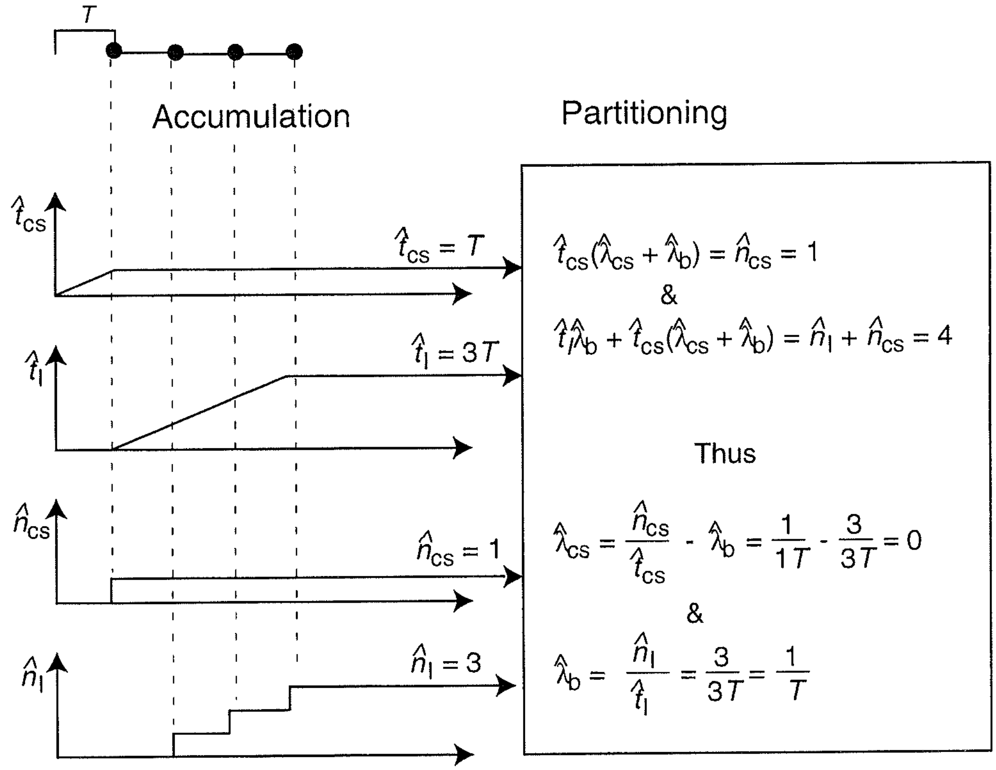

Figure 21. Explanation of response elimination: When intertrial reinforcers appear, they force a corresponding proportion of previous conditioned-stimulus (CS) reinforcements to be credited to the background rather than to the CS. Thus, the cumulative number of reinforcements credited to the CS stops growing as soon as the background begins to be reinforced at a rate equal to the rate of CS reinforcement. Note that partitioning depends only on the total cumulations, which are the same in this figure and Figure 20. This is a form of path independence. T = trial duration; $\hat{t}_{CS}$ = subjective cumulative exposure to the CS; $\hat{t}_I$ = subjective cumulative intertrial interval (i.e., cumulative exposure to the background alone); $\hat{n}_{CS}$ = subjective number of CS reinforcements; $\hat{n}_I$ = subjective number of intertrial (background alone) reinforcements; $\hat{\lambda}_{CS}$ = subjective rate of CS reinforcement; $\hat{\lambda}_b$ = subjective rate of background reinforcement.

conflict in favor of one percept or the other, even if the resolution is arbitrary.

In other cases, there may be an auxiliary principle that favors one solution over another. There may be a priori biases (Foree & LoLordo, 1973; LoLordo, Jacobs, & Foree, 1982), or a CS that has a greater observed range of variation (hence, higher contrast between its presence and its absence) might be preferred over a CS with a smaller range of variation. This auxiliary principle would explain why manipulating the relative intensities of the two CSs can determine which CS is overshadowed (Kamin & Gaioni, 1974; Mackintosh, 1976).

The principle of predictor minimization requires that one of two redundant CSs be credited with no rate of reinforcement. Although this is the result shown in Figures 16 and 17, it is not always the case that the subject fails to respond at all to the overshadowed CS. There are many cases in the literature in which there was some response to both CSs (e.g., Kamin, 1969a; Mackintosh, 1971; Wagner et al., 1968). Blocking is also sometimes incomplete (Ganesan & Pearce, 1988), although this finding is rarer. The explanation of overshadowing and blocking in many associative models depends on the assumed values of salience parameters, so these models can explain intermediate degrees of overshadowing (at least qualitatively) by an appropriate choice of values for these free parameters. The explanation of overshadowing and blocking in RET, by contrast, does not involve any free parameters: The principle of predictor minimization dictates that one or the other CS be given credit, but not both. Thus, the only way that intermediate degrees of overshadowing or blocking can be reconciled with RET is to assume that the subjects respond to the overshadowed CS not because that CS is itself credited with a rate of reinforcement, but because the overshadowed CS predicts the presence of the overshadowing CS, and the overshadowing CS is credited with a rate of reinforcement. Many readers will recognize this as an appeal to something akin to what are called "within-compound" associations (Dickinson & Burke, 1996; Rescorla & Cunningham, 1978), which are a well-established phenomenon, often incorporated into associative models as well. This interpretation raises the question of how one can determine whether the response to a redundant CS is due to the primary conditioning of that CS (a rate of reinforcement credited to that CS) or to second-order conditioning of that CS to the other CS. Techniques for doing this have been developed by Holland (1990), so this explanation can be tested. Of course, it would also be desirable to understand why one sees this phenomenon in some cases but not others.

One-Trial Overshadowing

The selective imputation of the observed rate of reinforcement to only one of two redundant CSs may occur after the first trial on which the two CSs are reinforced, because predictor minimization applies whenever ambiguity (multiple additive solutions) arises, and ambiguity arises the first time that there is a reinforcement of the jointly presented CSs. In the Rescorla–Wagner model, the overshadowing of one CS by another can develop only over repeated trials. Thus, experiments demonstrating overshadowing after a single trial have been thought to favor models in which an attentional process of some kind excludes one CS from access to the associative process (Mackintosh, 1975; Pearce & Hall, 1980). RET explains one-trial overshadowing without recourse to selective attention. Like ordinary overshadowing, it is an immediate consequence of the predictor minimization principle. That is, it is a consequence of the process for determining which stimuli get credited with which rates of reinforcement.

Relative Validity

The predictor minimization principle also predicts the relative validity effect first demonstrated by Wagner et al. (1968). Their two protocols had three CSs, A, B, and X. In both protocols, the X stimulus, that is, the common cue, was reinforced on half of the trials. In both protocols, it occurred together with the A stimulus on half of the trials and together with the B stimulus on the other half. In one protocol, however, only the AX trials were reinforced, whereas in the other, half of the AX trials and half of the BX trials were reinforced. Subsequently, unreinforced test trials with each stimulus presented in isolation showed that subjects exposed to the first protocol (only AX trials reinforced) developed a CR to the A stimulus but not to the X or the B stimulus. Subjects exposed to the second protocol (half of each kind of trial reinforced) developed a CR to the X stimulus but not to the A or the B stimulus, despite the fact that both were reinforced just as frequently as the X stimulus (see Figure 17).

Both protocols in Wagner et al.'s (1968) experiment give rise to two possible solutions, one involving only one CS and the other involving two CSs. The predictor minimization principle dictates the one-CS solution (the most valid CS). In the protocol in which both AX and BX trials are reinforced half the time, the one-CS solution credits all of the reinforcements to X. The alternative to this solution credits both to A and to B the rate of reinforcement credited to X by the one-CS solution. The predictor-minimizing machinery selects the one-CS solution. In the other protocol, in which only AX trials are reinforced, the one-CS solution credits all reinforcements to A. Alternatively, reinforcements on AX trials may be credited to X, but this predicts an equal rate of reinforcement on BX trials. To explain the absence of reinforcements on BX trials, this solution attributes an equal and opposite rate of reinforcement to B. The predictor-minimizing machinery rejects the two-CSs solution in favor of the one-CS solution. Thus, predictor minimization explains both overshadowing and the effect of the relative validity of a cue. Predictor minimization is a mechanical or algorithmic implementation of Occam's razor.

Retroactive Reversals

RET also predicts that the overshadowing of one CS by another will be reversed by subsequent extinction of the overshadowing CS. This result is a joint consequence of the additivity and predictor minimization principles. Predictor minimization credits the observed rate of reinforcement to only one CS, which leads to overshadowing in the first phase of conditioning. When the overshadowing CS is shown by subsequent experience not to predict the rate of US occurrence previously imputed to it, the additive partitioning of rates forces all of the USs previously credited to the overshadowing CS to be credited retroactively to the overshadowed CS. That is, the additional experience radically alters the solution to the rate estimation problem.

The retroactive effects of later reinforcements have already been seen in our explanation of response elimination. They are a consequence of the path independence of the partitioning process, its indifference to the order in which various CSs and CS combinations have been experienced. In RET, rate estimates depend only on the accumulated time and number totals. It does not matter how they were accumulated.[4]

The reversal of overshadowing by subsequent training with the overshadowing CS alone is difficult to explain for associative models that explain overshadowing in terms of the effect of cue competition on the strength of the association to the overshadowed

---

[4] In our analysis, the only alternative to the retroactive revision of previously computed rates of reinforcement is to conclude that the rate of reinforcement predicted by the overshadowing CS has changed. In the original formulation of RET (Gallistel, 1990), the conclusion that the rate had changed was prevented by a third constraining principle, the "rate–inertia" principle. The rate–inertia principle is that the rate attributed to a CS is not assumed to have changed unless the analysis for the nonstationarity of rate (the analysis that leads to the extinction decision) decides that it has. In the case of retroactive unblocking, there is no evidence of nonstationarity *if* the rates originally imputed to the overshadowing and overshadowed stimuli are reversed. After the reversal, there is no evidence that the rates of reinforcement imputed to each CS have changed since conditioning began.

Reversing rate estimates made under initially ambiguous conditions is not the same as deciding there has been a change in the rates. The system concludes, in effect, that its earlier estimates of the rates were erroneous, as opposed to concluding that its earlier estimates were correct and that the rates themselves have now changed. The latter conclusion is prevented by the rate–inertia constraint. This constraint imposes an arbitrary resolution on an inherently ambiguous state of affairs. The second scenario—that the rates themselves have changed—is just as consistent with the animal's experience as is the conclusion that previous rate accreditations were erroneous. It may be that, under some circumstances, the rate–inertia constraint does not operate, in which case earlier estimates will not be revised in the light of later evidence. This would explain why retrograde blocking is sometimes observed and sometimes not (Grahame et al., 1992). In retrograde blocking, subsequent reinforced presentations of the overshadowed CS retroactively block conditioning to the overshadowing CS. Generally speaking, in perceptual theory, ambiguity-resolving constraints are context-specific. They apply in some circumstances but not all. A principled analysis of which contexts will and will not support retroactive blocking is clearly necessary for this to be a satisfactory explanation.

CS (Baker & Mercier, 1989; Barnet, Grahame, & Miller, 1993a; Hallam, Matzel, Sloat, & Miller, 1990; Miller et al., 1992; Miller & Grahame, 1991; Miller & Matzel, 1989). It requires that the strength of the association to the overshadowed CS increase in the absence of any further experience with that CS. The strength of the association to the overshadowed CS must increase not simply as a consequence of the passage of time but rather as a consequence of the animal's subsequent experience with the overshadowing CS. This violates a fundamental principle of most associative models, which is that the only CSs whose associations with the US are modified on a given trial are the CSs present on that trial (but see Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994). These considerations have led to the suggestion that overshadowing is not due to an effect on associative strength but rather to a decision process that translates associative strengths into observed responses (Baker & Mercier, 1989; Cole et al., 1995a; Matzel et al., 1985). This suggestion moves associative theory in the direction of the kind of theory that we are describing, particularly when it is coupled with the temporal coding hypothesis (Barnet, Arnold, & Miller, 1991; Barnet et al., 1993a; Barnet, Grahame, & Miller, 1993b; Barnet & Miller, 1996; Cole et al., 1995b; Matzel, Held, & Miller, 1988; Miller & Barnet, 1993), which is that the animal learns the CS–US interval in addition to the CS–US association.

Retroactive effects may also be shown using the relative validity protocol (Cole et al., 1995a). Here, too, a subsequent disambiguating experience reverses an animal's initial conclusions about which CS predicts an observed rate of reinforcement. Again, the additivity and predictor minimization constraints together force this retroactive reversal of the rate estimates attributed to two CSs. The kind of analysis that we propose brings the study of learning close to the study of psychophysics and perception, not only in that it has explicitly formulated decision processes but also in that it has principles that constrain the interpretation of inherently ambiguous data, thereby resolving the ambiguities.

The timing-model explanations of the effects of background conditioning, overshadowing, and relative validity do not depend on assumptions about the values of free parameters. They are direct results of the two principles that determine the structure of the rate estimation machinery. By contrast, associative explanations of these phenomena depend on parametric assumptions. The explanations of the effects of background conditioning and relative validity were central features of the influential article by Rescorla and Wagner (1972), which is the most frequently cited contemporary theory of associative conditioning. However, it is unclear whether the Rescorla–Wagner model can explain both effects using a single set of parametric assumptions (Gallistel, 1990, pp. 412–417). Attention to whether and how explanations depend on parametric assumptions should be part of the evaluation of competing theories of conditioning.

Inhibitory Conditioning

The explicitly unpaired and feature-negative protocols. The additive partitioning of observed rates generates the phenomena of conditioned inhibition, without any further assumptions. Two protocols that produce what is called "conditioned inhibition" are the explicitly unpaired protocol and the feature-negative protocol. In the first, reinforcements occur at some random rate except for when a transient CS is present. Thus, when the CS comes on, the expected rate of reinforcement decreases to zero. If the reinforcers are positive reinforcers, the CR that develops is avoidance of the CS (Wasserman, Franklin, & Hearst, 1974). In the feature-negative paradigm, one CS (the CS+) is reinforced when presented alone but not when presented together with the other CS (the CS−). Thus, on those trials in which the CS− comes on along with the CS+, the rate of reinforcement decreases to zero from the rate predicted by the CS+. In time, the CR is seen to the CS+ but not to the combination of CS+ and CS−. When the CS− is then tested in combination with another separately conditioned CS+, the CR elicited by this other CS+ is reduced or eliminated (for reviews of the conditioned-inhibition literature, see LoLordo & Fairless, 1985; Rescorla, 1969).

The additive partitioning of observed rates of reinforcement dictates that the subjective rate of reinforcement for the CS− in the above described inhibitory-conditioning paradigms be negative because the rate attributed to the sometimes co-occurring CS (the background or the CS+), which is positive, and the rate attributed to the CS− must sum to the rate observed when the CS− is also present. Objective rates cannot be negative, anymore than amounts of money can be negative, but subjective rates can be negative as easily as bank balances can. Subjective rates are quantities or signals in the brain, just as bank balances are numbers in a book or bit patterns in a computer.

Subjective rates are used in the process of arriving at estimates of expected interreinforcement intervals, just as debts (negative assets) are used in arriving at estimates of net worth. Adding a negative rate estimate to a positive rate estimate reduces the estimated rate of reinforcement, which lengthens the estimated interval between reinforcements, thereby weakening or eliminating the CR. The lengthening of the expected interreward interval when the CS− comes on is what elicits conditioned avoidance of the CS−.

In this analysis, the conditioned effects of the CS− have nothing to do with inhibition in the neurophysiological sense. From a timing perspective, these phenomena are misnamed in that it seems odd to call the reduction in an expected rate of reinforcement or, equivalently, the lengthening of an estimated interval between reinforcements an instance of inhibition. Moreover, when we come to consider recent results from backward second-order conditioning experiments, the reader will see that calling these CSs "inhibitors" leads to confusion and perplexity.

The overprediction protocol. Although conditioned inhibition is normally produced by omitting reinforcement when the target CS is present, it can also be produced by protocols in which the target CS is reinforced every time it is presented (Kremer, 1978; Lattal & Nakajima, 1998). In the first phase of this protocol, two CSs are separately presented and reinforced. In the second phase, they are presented together and accompanied by a third CS. Each presentation of the three-CS compound is reinforced. Because rate estimates combine additively in RET, the rate of reinforcement to be expected when two CSs are presented together is twice the rate that is expected when each is presented separately. But only one reinforcement is in fact given on the three-CS trials. There is a unique additive solution to the resulting discrepancy between the predicted and observed rates, namely, that the rate of reinforcement ascribed to the third CS is equal in magnitude to the rates ascribed to each of the first two CSs but opposite in sign. Asymptotically, $\hat{\lambda}_1$, $\hat{\lambda}_2 = 1/T$, and $\hat{\lambda}_3 = -1/T$. This protocol does indeed produce inhibitory conditioning of the third CS, despite the fact that the CS is paired with reinforcement on every occasion on which it is presented. Note that the predictor minimization principle does not come into play here because there is a unique additive solution. Predictor minimization operates only when there is more than one additive solution.

A related protocol weakens the CR to two independently conditioned CSs by pairing them and giving only one reinforcement per paired presentation (Kamin & Gaioni, 1974). Here, there is no third CS to whose influence the missing reinforcements can be ascribed. Hence, the rates of reinforcement ascribed to each of the two CSs must be reduced when they are presented together without doubling the amount of reinforcement per presentation. It is readily shown that when the number of joint presentations has grown to equal the number of (initial) individual presentations, $\hat{\lambda}_1 = \hat{\lambda}_2 = 2/3(N/T)$. At that point, the paired presentations have reduced the original (prepairing) rates (N/T) by one third.

Backward, Second-Order, and Trace Conditioning

Three dichotomies that figure prominently in textbook presentations of basic conditioning are the dichotomy between forward and backward conditioning, the dichotomy between primary and second-order conditioning, and the dichotomy between delay and trace conditioning. In each case, the latter type of conditioning is usually observed to be less effective than the former. A recent series of brilliantly conceived and executed experiments in the laboratory of Ralph Miller have shown that temporal coding is fundamental to an understanding of these dichotomies (for a review, see Savastano & Miller, 1998). Backward conditioning (the US precedes the CS) is less effective than forward conditioning, not because it produces weaker associations but because the subject learns that the US precedes the CS; hence, the CS does not ordinarily enable the subject to anticipate the US. Second-order conditioning is less effective than primary conditioning because the subject learns that the second-order signal will be followed at a predictable latency by the primary signal, which, being closer in time to the US, is a more precise predictor of it. Trace conditioning is less effective than delay conditioning for the same reason: The subject learns that CS onset will be followed at a predictable latency by CS offset, which, being closer in time to the US, is a more precise predictor of it. By exploiting these insights, Miller and his collaborators have been able to arrange conditions that reverse each of these conventional observations, making backward conditioning appear stronger than forward conditioning, second-order conditioning appear as strong as primary conditioning, and trace conditioning appear stronger than delay conditioning.

The experiments to be reviewed from Miller's lab use a CER paradigm in which a CS signals impending shock, the fear of which suppresses drinking (the licking of a water tube). On test trials, the shock is omitted so that the CR to the CS is uncontaminated by the response to the shock. The latency to resume licking after the CS comes on is the measure of the strength of the CR.

If the shock coincides with CS offset, the procedure is delay conditioning. If the shock occurs some while after the offset of the CS, it is trace conditioning. If the shock onset precedes CS onset, the procedure is backward conditioning. If the CS is paired with shock (by delay, trace, or backward conditioning), the procedure is a primary-conditioning procedure. If the CS is not conditioned directly to shock but rather has been paired with another CS that has already been directly conditioned, the procedure is secondary conditioning. Secondary conditioning is one of two forms of second-order conditioning. In the other form, called sensory preconditioning, two neutral CSs are paired in Phase 1, and then one of them is paired with a US in Phase 2. Thus, secondary conditioning and sensory preconditioning differ only in whether the CS–US phase of training comes before or after the CS–CS phase.

One experiment (Cole et al., 1995b) looked at trace conditioning versus delay conditioning, backward conditioning, and second-order conditioning. In the secondary-conditioning version of the experiment, the subjects were first given primary conditioning, using either a delay protocol or a trace protocol. In the delay protocol, the CS lasted 5 s, and the shock coincided with its offset (see the top protocol in Figure 22A). In the trace protocol, the CS also lasted 5 s, but the shock did not occur until 5 s after its offset (see the middle protocol in Figure 22A).

Some subjects were immediately tested for their degree of suppression to the CS (Cole et al., 1995b). These tests revealed the expected difference between the results of delay versus trace conditioning: The trace-conditioned subjects showed less fear of the CS than the delay-conditioned subjects. This outcome is doubly predicted by associative theory because, first, trace conditioning is weaker than delay conditioning, even when the CS–US latency is the same. This is traditionally taken to indicate that the associations formed to stimulus traces are weaker than the associations formed when the stimulus is present at the moment of reinforcement. Second, in the present case, the trace group had a greater delay of reinforcement (10 s vs. 5 s). Prolonging the interval from CS onset to reinforcement is traditionally thought to weaken the strength of the association by reducing the degree to which the stimulus onsets are temporally paired.

The subjects that were not immediately tested for the strength of primary conditioning next experienced a brief phase of backward secondary conditioning (see the bottom protocol in Figure 22A). During this training, the CS that they had already learned to fear was followed by another CS, also lasting 5 s. As usual, the phase of secondary conditioning was kept brief because the primarily conditioned CS was not reinforced during secondary conditioning. From the standpoint of the primary conditioning, the secondary-conditioning phase is a brief extinction phase; it must be kept brief to prevent extinction of the primary conditioning.

From the standpoint of associative theories, the group that received primary trace conditioning followed by backward secondary conditioning should show no conditioned responding to the second-order CS, for two reasons: First, the primary conditioning itself was weaker in this group because it was trace conditioning rather than delay conditioning. Second, the secondary conditioning was backward conditioning. Backward pairing is commonly assumed to produce no (or very little) conditioning.

From a timing perspective, however, this group should show a strong CR to the secondary CS, as one can see by looking at the diagram of the three protocols in Figure 22A. In fact, their response to the secondary CS should be stronger than the response of the group that received primary delay conditioning followed by backward secondary conditioning. In the latter group, the expected interval to shock when the secondary CS (the clicker) comes on is zero. (This expectation can be seen in Figure 22A by following the dashed vertical line from the shock in the primary-conditioning

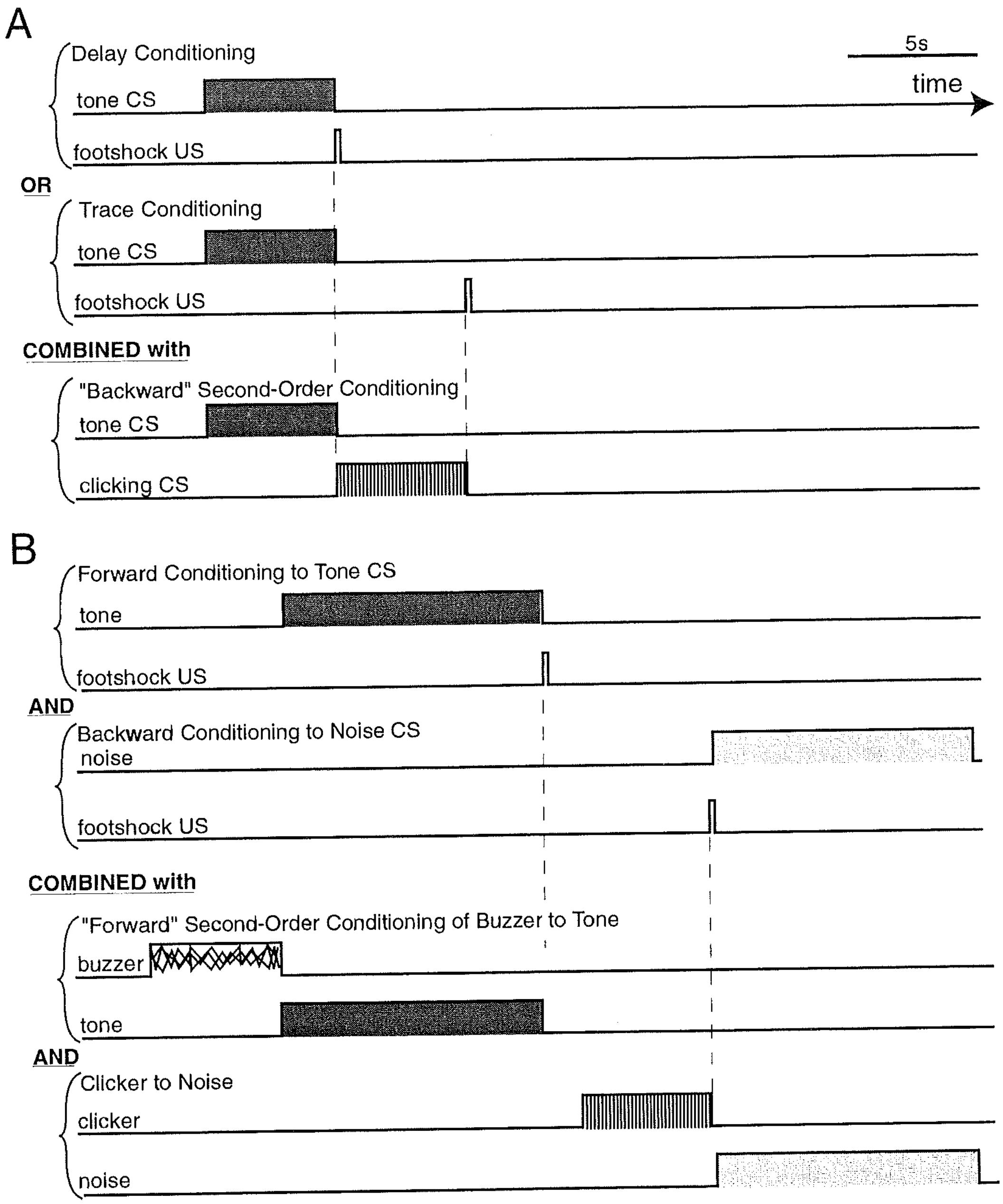

Figure 22. A: Diagram of experiments by Cole, Barnet, and Miller (1995b). The tone (black box) was the primary conditioned stimulus (CS); the clicker (vertical striped box) was the second-order CS. B: Diagram of experiment by Barnet, Cole, and Miller (1997). The tone and the buzzer (black and gray boxes, respectively) were the primary CSs. The dashed vertical lines are aids to perceiving the expected temporal relation between the second-order CSs and the shock (the unconditioned stimulus [US]) when the remembered CS–US intervals from the different phases of conditioning are summed to yield the expected interval between the onset of the second-order CS and the US.

phase down to the onset of the secondary CS.) They should show little CR for the same reason that animals that are given simultaneous primary conditioning fail to respond to the CS. By the time they realize that the CS has come on, they have nothing to fear. In fact, this group showed very little fear (very little lick suppression). But for the group that got trace conditioning, the expected interval to shock was 5 s when the clicker came on. They should have been afraid of the clicker, and they were, despite the fact that the clicker itself was never paired with shock.

The reaction to the clicker in the group that got primary trace conditioning followed by backward second-order conditioning was not significantly weaker than the reaction to the primary CS (the tone) in the control subjects that got simple delay conditioning. This comparison is not entirely valid because the clicker (the secondary CS) is generally found to be more potent in this kind of conditioning than the tone used for primary conditioning. Nonetheless, the comparison emphasizes that the fear elicited by combining what the animal has learned about the temporal relations between the three stimuli (the tone, the clicker, and the shock) during different phases of conditioning can be a strong fear. These are not marginal effects.

Cole et al. (1995b) also did a sensory-preconditioning version of the experiment. Recall that in secondary conditioning the primary protocol is given first and the second order is given second, whereas in sensory preconditioning this order is reversed. This variation in training order is important from an associative perspective. When the second-order conditioning occurs before the primary conditioning, conditioning cannot be explained in terms of secondary reinforcement (the idea that a CS associated with a US takes on the reinforcing properties of the US). From a timing perspective, this variation in procedure is of little theoretical importance because the learning of temporal intervals is not dependent on reinforcement. The only effect of reversing the training order is that the primary-conditioning phase becomes an extinction phase for the second-order conditioning rather than vice versa. This is a serious consideration only if the second phase of conditioning (whether primary or second order) is prolonged. But it is not, precisely in order to avoid this problem. From a timing perspective, one should get basically the same result in the sensory-preconditioning version of this experiment as in the secondary-conditioning version. That was in fact the case.

Again, the strength of the fear reaction to the clicker in the trace group was as strong as the reaction to the tone in the delay group, despite the fact that the tone was directly paired with shock whereas the clicker was not. (Again, however, this comparison is somewhat vitiated by the fact that the CSs used in different roles were deliberately not counterbalanced.) The strong fear reaction to the clicker was observed only in the group that received the supposedly weaker form of primary conditioning—trace conditioning. In the group that received the supposedly stronger form of primary conditioning, the fear reaction to the clicker was much weaker.

In these experiments by Cole et al. (1995b), the relative strengths of the reactions to the second-order CSs were the reverse of the relative strengths of the reactions to the primary CSs. From an associative perspective, these reversals are paradoxical. The reaction to the second-order CS is assumed to be mediated by the second-order and primary associations conducting in series. The second-order training was always the same in these experiments; only the primary training differed. So, the strengths of the reactions to the second-order CSs should have been determined by the strength of the primary associations, which they clearly were not.

A second experiment (Barnet, Cole, & Miller, 1997) made the same point. There were two 10-s-long CSs in the phase of primary conditioning, one forwardly conditioned and the other backwardly conditioned (top two protocols in Figure 22B). There was no fixed temporal relation between these two CSs, but there was a fixed temporal relation between each CS and shock. Shock always occurred at the offset of the one CS but immediately preceding the onset of the other CS.

In the second-order phase, two different secondary CSs, each 5 s in duration, were forwardly paired with the two primary CSs (bottom two protocols in Figure 22B). Thus, the onset of one second-order CS preceded by 5 s the onset of the forwardly conditioned primary CS, whereas the onset of the other second-order CS preceded by 5 s the onset of the backwardly conditioned primary CS. (As in the previously discussed experiment, there were secondary-conditioning and sensory-preconditioning versions of this experiment.)

The strengths of the fear reactions to the primary CSs showed the usual difference between forward and backward conditioning; that is, reactions to the backwardly conditioned primary CS were weaker. However, the strengths of the reactions to the second-order CSs were reversed. The second-order CS that predicted the backwardly conditioned primary CS elicited more fear than did the second-order CS that predicted the forwardly conditioned primary CS.

The diagram of the temporal relations between CSs and USs (Figure 22B) makes the timing explanation of these results more or less self-evident. What one needs to consider is the expected interval to shock at the onset of a CS—the comparison quantity in SET. In the tests of the primary CSs, when the forwardly conditioned CS (the tone) comes on, the expected interval to shock is 10 s. When the backwardly conditioned CS (the noise) comes on, however, it is 0 s. As previously explained, this should result in little fear, given the decision rule specified by SET. In contrast, when the secondarily conditioned buzzer comes on, which predicts the onset of the tone, the expected interval to shock is 15 s. Moreover, the onset of a still better warning stimulus—one that more precisely predicts the shock—is expected in 5 s. When the secondarily conditioned clicker comes on, however, the expected interval to shock is only 5 s, and probably more importantly, no further warning stimulus is expected because the shock is expected before the primarily conditioned noise. Thus, the shock is coming very soon, and there will be no further warning.

The diagrams in Figure 22 make it clear why one might wish to speak of subjects' forming a temporal map during conditioning (Honig, 1981). The ability to remember temporal intervals and to add, subtract, and divide them—the central capabilities assumed in our timing models—gives the animal such a map.

A third experiment (Barnet & Miller, 1996) shows the usefulness of the timing perspective in understanding both backward conditioning and conditioned inhibition. This intricately designed experiment requires for its interpretation both RET and SET. It brings into play most of the theoretical analyses we have so far developed.

Recall that in RET a CS is a conditioned inhibitor if its estimated rate of reinforcement for that CS is negative. When a CS to which a negative subjective rate of reinforcement is attributed is presented together with a previously conditioned excitor, the negative rate of reinforcement sums with the positive rate attributed to the excitor. The expected rate of reinforcement is thereby reduced. This test—presenting the putative conditioned inhibitor together with a conditioned excitor—is called the summation test for conditioned inhibition. Recall also that in SET, when there is a fixed interval between CS onset and reinforcement, the expected interval to reinforcement is used to time a CR that anticipates the reinforcement. All of these principles come into play in what follows.

The experiment (Barnet & Miller, 1996), which is diagrammed in Figure 23, involved three training phases. The first phase forwardly conditioned an excitor to be used in the summation test. That is, the subjects were taught to fear a stimulus that predicted shock. The second phase backwardly conditioned an inhibitor to be used in the summation test. The third phase paired a third CS with the so-called inhibitor, without further shocks (second-order conditioning of an inhibitor).

In the first phase, each presentation of the excitatory CS lasted 30 s, and there was a 1-s footshock coincident with its offset. Thus, the expected rate of reinforcement in the presence of this excitor was two shocks per minute. A prolonged phase of backward inhibitory conditioning followed, carried out in the same chamber as the excitatory conditioning but with a different CS. Each of the twelve 20-min sessions in this phase contained eight 30-s presentations of the inhibitory CS, with each presentation of this CS immediately preceded by a 1-s shock. The background against which the shock was presented was a flashing houselight that came on 30 s before the shock. In the analysis of this experiment from the perspective of RET, it is assumed that this flashing houselight constituted the background and that the backwardly conditioned CS, which came on when the shock and the background were terminated, can be treated as having suppressed the background and, hence, the rate of reinforcement predicted by it, which was two shocks per minute.

Finally, there was a phase of secondary conditioning in which the subjects learned that yet another CS predicted the onset of the "inhibitory" CS. The reason that quotation marks now appear around *inhibitory* will soon be evident. The secondary CS lasted only 5 s, and the 30-s "inhibitory" CS came on at its offset.

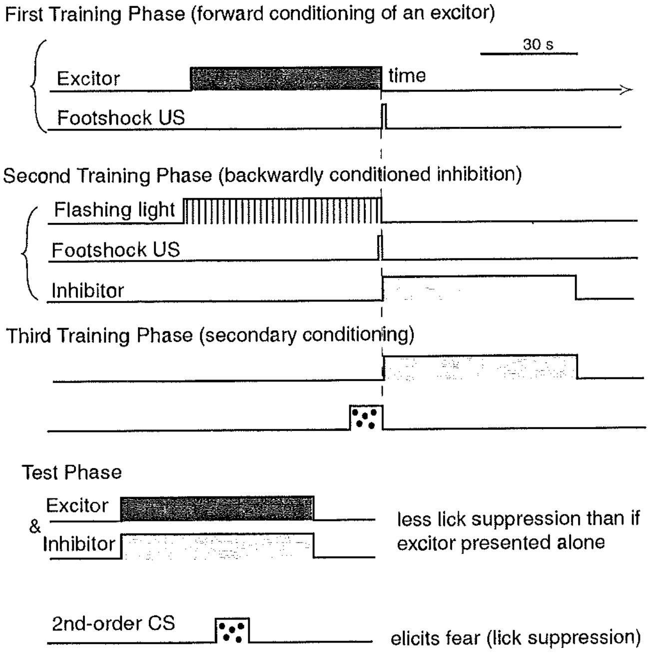

Figure 23. Schematic diagram of an experiment by Barnet and Miller (1996) in which the inhibitor conditioned stimulus (CS) reduced (inhibited fear), but a second-order CS, whose only connection to shock was that it predicted the inhibitor CS, nonetheless elicited (excited) fear. A consideration of the temporal relations diagrammed here makes this seeming paradox intelligible. US = unconditioned stimulus.

As always, the phase of secondary conditioning was kept brief to avoid extinction of the primary conditioning. We note in passing that the account of extinction offered by RET makes it clear why this works. In associative theories, every nonreinforced trial weakens the net excitatory strength of the CS–US associations, and the initial trials in a series of nonreinforced trials have the greatest weakening effect. Therefore, a moderately numerous sequence of unreinforced trials should have a big impact on the (net excitatory) strength of a CS–US association, but, in fact, such strings have no detectable impact (Prokasy & Gormezano, 1979). In the RET model of extinction, by contrast, the first few nonreinforced segments of CS exposure have no effect. Extinction does not occur until the ratio between the interval without CS reinforcement and the expected interval between CS reinforcements gets large. Until the subject decides that the rate of reinforcement has changed, there is no extinction.

The secondary conditioning was conducted in a different chamber from the primary conditioning, as were the summation test and the test for the subjects' CR to the secondary CS. Thus, all of the conditions in which shock never occurred were run in a context that gave the animal no reason to expect shock.

The summation test showed that the CS that suppressed the flashing became a conditioned inhibitor. Presenting this backwardly conditioned stimulus together with the excitor CS in the test phase of the experiment diminished the subject's fear of the excitor. This is the standard test for conditioned inhibition. The backwardly conditioned experience taught the animal that the CS suppressed an excitor. When the backwardly conditioned stimulus was combined with another independently conditioned excitor, it did not eliminate the fear caused by the excitor, but it did reduce it. This is to be expected if the inhibitory power of the backwardly conditioned stimulus arose from its suppression of an excitatory "background" (the flashing light) rather than from its direct suppression of shocks themselves.

That backward conditioning should establish an effective inhibitor is itself interesting. It shows once again that forward temporal pairing is not essential to conditioning. However, the theoretically most interesting part of this experiment comes from the results of the test for the effect of the secondary CS—the CS whose onset predicted that the conditioned "inhibitor" would come on in 5 s. The secondary CS excited fear (manifested in a strong suppression of licking), even though its only connection to the shock was through a primary CS that inhibited fear. It is this result that has led us to place quotation marks around "inhibitor" and its cognates. If a stimulus has an inhibitory effect on the CR, then another stimulus, whose only connection to the reinforcer is that it predicts the inhibitory stimulus, ought itself to inhibit the CR. But it does not inhibit the CR; on the contrary, it strongly excites it.

SET predicts this result, for reasons diagrammed in Figure 23. In the backward-conditioning phase, the animal learned that the interval from the onset of the "inhibitory" CS to the onset of shock was -1 s. In the secondary-conditioning phase, it learned that the interval from the onset of the secondary CS to the onset of the "inhibitory" CS was 5 s. The expected interval to shock when the secondary CS came on should have been the sum of these two intervals, which was 4 s. The subjects should have feared an impending shock, and they did.

RET and SET are not in conflict here. RET deals only with the rate of reinforcement that the animal expects when the inhibitory CS is present. That expectation is zero shocks because the backwardly conditioned stimulus suppresses the excitor. The rate estimation mechanism takes account of the analytic impossibility of directly observing a negative rate. SET deals with a different aspect of the animal's decision making, the timing of a CR controlled by the expected interval to the reinforcement (the *when* decision, in contrast to the *whether* decision). SET gives (in part) a quantitative formalization of Honig's (1981) temporal map and Miller's (Miller & Barnet, 1993) temporal coding hypothesis. Intuitively, any animal with a temporal map—any animal that has learned the temporal relations between the various stimuli—will show what Barnet and Miller (1996) found, as the diagram in Figure 23 makes clear. Their findings are not surprising if it is assumed that knowledge of the temporal intervals in conditioning protocols is the foundation of conditioned behavior.

In a second version of this experiment, Barnet and Miller (1996) added an extinction phase after the backward-conditioning phase. In this version, the animal was first taught that a shock immediately preceded the onset of the "inhibitory" CS (the backward-conditioning phase) and then that this was no longer true (the extinction phase). In the extinction phase, the "inhibitory" CS was repeatedly presented in the absence of shock—and (importantly) also in the absence of the flashing light that presaged shock (the effective background stimulus for the shock experience during the backward conditioning). From a timing perspective, the extinction phase should persuade the animal that, although it once was true that the inhibitory CS was preceded by a shock, that is no longer true. However, this extinction experience should not affect the animal's estimate of the power of the backwardly conditioned stimulus to suppress an excitatory background stimulus (the flashing light), because this stimulus was absent during extinction. The extinction experience, if anything, confirmed that when the "inhibitory" CS was present, the threatening flashing light was not present.

This second version of the experiment (the version with an interpolated extinction phase) emphasizes the importance of the distinction between the *whether* decision and the *when* decision. The extinction phase removes the expectation that is used in the excitatory *when* decision. This *when* decision leads to a fear response to the secondary CS (and the consequent suppression of licking). But the extinction does not remove the expectation that makes the primary CS an effective conditioned inhibitor, the expectation that the primary CS can suppress CSs that predict shock, thereby lowering the expected rate of shock reinforcement.

The results from this second version (with interpolated extinction) were as expected from the joint application of SET and RET. The extinction phase did not affect the results of the summation test. Despite the intervening extinction phase, presenting the "inhibitory" CS together with the excitor diminished the subjects' fear of the excitor, thereby moderately reducing the amount of lick suppression, just as in the version without an extinction phase. The extinction phase, however, greatly reduced the animals' fear of the secondary CS. Thus, the extinction phase removed the basis for the anticipatory reaction to the secondary CS (the reaction that occurs before the primary CS comes on), without removing the basis for the reduction in fear caused by the actual presence of the "inhibitory" CS.

Further versions of this basic experiment replaced the backward conditioning of the inhibition with a phase of conventional, feature-negative inhibitory conditioning. In this more conventional protocol, the "inhibitory" effect of one CS was created by pairing that CS from time to time with an otherwise reinforced CS (hence, an excitor) and omitting the reinforcement on those trials. Historically, this is the procedure, first used by Pavlov himself, that gave rise to the concept of a conditioned inhibitor.

The use of the more conventional procedure for establishing conditioned "inhibition" eliminates the expectation that figured in the previously described prediction because there no longer is a fixed temporal interval between the onset of the "inhibitory" CS and the onset of the reinforcer. SET operates only when there is such an expectation. But RET still operates because the additive partitioning of observed rates forces the attribution of a negative rate of reinforcement to the influence of the "inhibitory" CS. Moreover, according to SET, the animal learns the duration of the secondary CS during the phase of secondary conditioning. Thus, when a stimulus that has been secondarily conditioned to the "inhibitory" CS comes on, the animal expects that—after an interval equal to the duration of the secondary CS (SET)—the rate of shock will go down (RET). The two models together predict that, in this version of the experiment, the secondary CS will alleviate fear, just as does the conditioned inhibitor itself. And this was the result that was in fact obtained. In other words, timing theory explains why and how the conditioned "inhibition" created by the standard protocol (the one Pavlov himself used) differs from the conditioned inhibition created by the backward-conditioning protocol. In the process, it shows why these phenomena are better thought of not in terms of inhibition and excitation but rather in terms of the content of conditioning—the interval and rate expectations established by the conditioning. The importance of the extensive series of experiments from Miller's laboratory in recent years (Savastano & Miller, 1998)—only very partially reviewed here—is that they demonstrated beyond reasonable question that the fact that subjects learn the intervals in the conditioning protocol is the key to understanding second-order conditioning and trace conditioning.

Operant Choice

The previously presented models for the timing, acquisition, and extinction of the CR share with modern psychophysical models a decision-theoretic conceptual framework. In this framework, a decision mechanism determines a response depending on whether the value of a task-specific decision variable is greater than or less than some criterion. The idea that different aspects of conditioned behavior depend on different decision variables is a fundamental feature of the conceptual framework that we propose (as it is in sensory psychophysics). We now extend this idea to the analysis of operant choice. We describe two different decision mechanisms in operant choice, which operate in different contexts.

Opting Versus Allocating

Most of the theoretical literature on operant choice attempts to specify the effect of a given pattern of reinforcement on the subjective value (attractiveness) of an option (a key or a lever that the subject may peck or press, respectively). Implicit in most such analyses is the assumption that there is a fixed monotonic mapping between the relative subjective values of two options and observed preference (the relative frequency with which they are chosen or the relative amounts of time or numbers of responses devoted to them). However, this will not be the case if different decision mechanisms operate in different choice contexts. Suppose, for example, that in one context, choosing serves to select the best option, whereas in another context, choosing serves to allocate time or effort among the options so as to optimize net gain. In the first case, the choice mechanism is optimal to the extent that it always selects the best option. In the second case, the choice mechanism is optimal to the extent that it allocates time and effort in the best way. These are different goals. We assume that they cannot be and are not mediated by one and the same choice mechanism. Rather, the animal invokes different choice mechanisms, depending on what it has learned about the context. We call choice behavior mediated by the first mechanism *opting behavior* and choice behavior mediated by the second mechanism *allocating behavior.*

The Opting Decision

The desirability of distinguishing between opting and allocating became apparent in the course of a "failed" experiment by Brunner, Gallistel, Fairhurst, and Gibbon (2000). Their experiment was intended to be a variant of the well-known matching paradigm in which subjects choose between two manipulanda reinforced by concurrent variable interval (VI) schedules. These schedules make reinforcement available at a VI following the collection of a preceding reinforcement. Usually, the distribution of intervals is approximately exponential. Schedules are said to run concurrently when two response options are presented continuously and simultaneously, each reinforced on its own schedule. The paradigm models the situation that a foraging animal faces when it must allocate its foraging time among patches that differ in food density. Herrnstein (1961) discovered that, under these circumstances, the ratio of the dwell times and responses allocated to the options during a session tended to match the ratio of the reinforcements obtained from them. The study and analysis of this "matching" phenomenon have formed a substantial part of the experimental and theoretical literature on operant choice in the past four decades (Commons, Herrnstein, & Rachlin, 1982; Davison & McCarthy, 1988; Herrnstein, 1991; Herrnstein & Prelec, 1991; Nevin, 1982).

In the failed version of this experiment, pigeons were first trained with keys presented one-by-one rather than together. Six pigeons received 25 sessions of initial training in which each of two differently illuminated keys (red or green) was presented 15 times per session. Each time a key was presented, it remained on until the pigeon had collected a reinforcement from it. The amount of time that elapsed on any one trial between the illumination of the key, which signaled its availability as an option, and the arming of the reinforcement mechanism varied from trial to trial in accord with the intervals programmed into a standard VI scheduling tape. The average scheduled interval depended on which key was illuminated; it was 20 s when the red key was illuminated and 60 s when the green key was illuminated. When these subjects were then shifted to concurrent training (both keys were illuminated at the same time), they did not show a degree of preference for the richer key that matched its relative richness (i.e., a 3:1 preference). All 6 pigeons preferred the richer key exclusively; they almost never chose the leaner key. This is called overmatching. Overmatching is seldom observed when concurrent VI schedules constitute the options, and this degree of overmatching—98% to 100% choice of the richer schedule in 6 out of 6 subjects—is without published precedence. The effect of the initial training, with each key presented separately, on subsequent preference under concurrent conditions was robust and enduring. Three of the 6 subjects showed exclusive preference for the richer key throughout 25 sessions of concurrent training.

Related experiments in this series showed that breaking up the usually continuous concurrent-schedules paradigm into discrete trials, with each trial terminating in a reinforcement on one or the other of the two simultaneously operative keys, did not produce overmatching. On the contrary, it reliably produced some degree of undermatching, provided that the schedule on the key not being pecked elapsed in the discrete trial procedure exactly as in the continuous procedure, that is, provided that the concurrent intervals did not begin anew with each trial. Thus, it was presumably the presentation of the options individually during initial training, not the breaking up of the experience with each key into discrete one-reinforcement trials, that resulted in an exclusive preference for the richer key. A similar result was obtained by Edmon, Lucki, and Grisham (1980), who trained pigeons on multiple VI schedules—a protocol in which only one schedule is available at any one time, with the different schedules signaled by different key lights. When these subjects were tested with keys presented concurrently, they showed almost exclusive preferences for the keys that signaled the richer of the two choices. These findings call into question the assumption that matching behavior is the result of the relative strengths of the responses.

We suggest that in the course of the initial training, the subjects in these experiments learned that this was not a context that permitted the allocation of behavior among concurrent options. In our model, the opting mechanism mediates choice between mutually exclusive options, options that cannot be concurrently exploited. Our model for the decision process that operates in such contexts is the signal-detection model for forced choices based on noisy decision variables. The expected delays of reinforcement for the two options are represented internally by (memory) signals with scalar sampling variability (Figure 24). The remembered

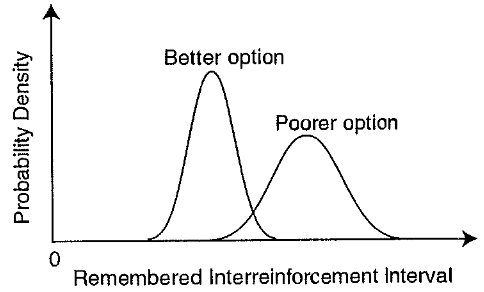

Figure 24. The signal-detection conception applied to choice behavior in an opting context. The signals representing the parameters of the outcome (in this case, the interval between reinforcements) have scalar variability (noise). To the extent that the probability density functions for the distributions of these signals overlap, there is some likelihood that, on a given choice trial, the subject will be mistaken about which is the better option. Such mistakes are due to sampling error.

interval for a given option—the signal generated for that option when memory is consulted—varies from one consultation of memory to the next. The decision mechanism reads (samples) memory to get the value associated with each option, and it always chooses the option with the lower sampled value. The choice of the option with the longer interreinforcement interval occurs only as a result of sampling error, that is, only on trials in which the sampled signal for the leaner option happens to be shorter than the sampled signal for the richer option. Thus, when the distributions from which the choice mechanism is sampling are well separated, the leaner option is almost never chosen. That is why pigeons that are trained with only one VI option available at any one time almost invariably choose the richer of the two options when they are subsequently offered as concurrently exploitable options. The subjects' choice behavior is no longer mediated by the decision process used to allocate behavior among concurrently exploitable options because the subjects have learned that these options are not concurrently exploitable.

In the framework that we propose, the sources of nonnormative decision making are to be sought in the idiosyncrasies of the processes that compute the decision variables. The decision variables in opting behavior are expected (subjective) incomes. The subjective income from a single reinforcement is the subjective magnitude of that reinforcement divided by its subjective delay. If the reinforcers are appetitive (positive reinforcement), then the option with the greatest average subjective income is preferred; if the reinforcers are aversive (negative reinforcement), then the option with the smallest average subjective income is preferred. However, the computation of subjective income has repeatedly been shown to depart from the normative computation in two ways: (a) Only the intervals when a reinforcement is pending are taken into account. We call this the principle of subjectively sunk time. (b) When the amount and delay of reinforcement vary, the average subjective income is the expectation of the ratios rather than the ratio of the expectations (Bateson & Kacelnik, 1996). The ratios in question are the ratios between the magnitudes and the delays of reinforcement. Each such ratio is an income datum. It is a well-established experimental fact that laboratory animals average income data in computing the expected subjective income for an option. This is commonly termed *hyperbolic discounting* or *harmonic averaging*. On the one hand, this finding supports our contention that the rate of reinforcement (the reciprocal of the delay) is the fundamental variable in conditioning (rather than probability of reinforcement), because when reinforcement magnitude does not vary, averaging income data is equivalent to averaging rate data. On the other hand, this way of computing the average subjective income can cause the resulting subjective expectation to depart arbitrarily far from the true expectation, so it is decidedly nonnormative. This departure from what normative considerations would lead one to expect is perplexing.

Subjectively sunk time. The intertrial intervals when no option is present, the latencies to choose, and the intervals devoted to consuming (or otherwise dealing with) the reinforcements are not taken into account in computing the expected income (Bateson & Kacelnik, 1996; Mazur, 1991). In our model, the cumulative interreinforcement interval timer does not run during these intervals; it runs only when a reinforcement is pending. We call these subjectively sunk times by analogy to the microeconomic concept of a sunk cost. A sunk cost is a cost that appears to be relevant to the computation of the utility of an alternative but, in fact, is not and ought to be ignored in rational (normative) economic decision making. A subjectively sunk interval is an interval that is not included in the computation of subjective income.

Subjectively sunk intervals are not necessarily objectively sunk intervals, and this discrepancy can lead to strongly nonnormative preferences. Suppose, for example, that an animal is repeatedly allowed to choose between Food Rewards A and B, where the magnitude of A is always 1.5 times the magnitude of B, but the delay of B—the latency between choice and receipt, that is, the time when the reward is pending—is 0.5 times the delay of A. If the animal chooses B, it gets its reward 50% sooner, but what it gets is only two thirds of what it would have gotten had it chosen A. Suppose, as is commonly the case, that session time is fixed or that the intervals during which reinforcements are pending account for only a small proportion of session time. Under these conditions, reductions in the average amount of time during which a reward is pending are offset by increases in subjectively sunk time (intertrial intervals) during which no reward is pending, so shortening the average delay between choice and reinforcement (the interval when a reinforcement is pending) has a negligible effect on the objective income, that is, on the amount of reinforcement per unit of session time. Under these circumstances, the rational decision maker should strongly prefer the bigger reward as opposed to the one that is delivered more quickly. The rational decision maker should not treat the intertrial intervals and the decision times as sunk intervals. But because the delay-of-reinforcement timer does not run during these intervals, subjects strongly prefer the quicker reward. This irrational preference for the quicker but smaller reward has been called a lack of self-control brought on by the prospect of immediate reinforcement (Rachlin, 1974, 1995). From a timing perspective, it is a consequence of the principle of subjectively sunk times and hyperbolic discounting.

Hyperbolic discounting and harmonic averaging. The principle that the decision variables on which opting is based are subjective incomes—subjective reinforcement magnitudes divided by interreinforcement intervals—is equivalent to the hyperbolic discounting principle, which has often been used to describe the effect of delaying reinforcement on the subjective value of the reinforcement (Fantino, 1969; Killeen, 1982; Mazur, 1984, 1986; McDiarmid & Rilling, 1965; Rilling & McDiarmid, 1965). In Mazur's (1984) formulation,

$$V = \frac{A}{1 + kD}, \tag{3}$$

where V is the value of a reward (positive reinforcement), A is the amount (magnitude), and D is the delay. This may be rewritten as

$$V = \frac{A}{D + k'}, \text{ with } k' = \frac{1}{k},$$

which reduces to *magnitude* (amount) divided by interreward interval (delay) when k' is zero (or negligibly small relative to D).

Equation 3 is a necessary modification of the income formula when it is applied to very short delays. Without the 1 in the denominator, the income produced by a given amount of food would go to infinity as the delay went to 0. Clearly, as the delay becomes negligible, the income ought to become asymptotically equal to the subjective magnitude of the food. This is accomplished by adding the 1 to the denominator. Psychologically, this may be thought of as taking into account that below some value, differences in the objective delay are psychologically negligible. Put another way, delays are scaled relative to the delay that is perceived as immediate, that is, as lasting no longer than a single unit of subjective time. The irreducible unit of time may also be thought of as the latency to begin timing a delay (see Gibbon et al., 1984). If the reinforcement is delivered in less than this latency, then it has negligible delay.

When an option sometimes produces one delay and sometimes another, the expected delay of reinforcement that governs opting decisions is the harmonic average (Bateson & Kacelnik, 1995, 1996; Brunner & Gibbon, 1995; Brunner, Gibbon, & Fairhurst, 1994; Gibbon, Church, Fairhurst, & Kacelnik, 1988; Killeen, 1968; Mazur, 1984; McDiarmid & Rilling, 1965; Schull, Mellon, & Sharp, 1990). The harmonic average of a series of interreinforcement intervals is the reciprocal of the average of the reciprocals. The reciprocal of an interreinforcement interval is a rate (the rate–interval duality principle). Thus, this well-established and surprising way of computing the expectations that govern opting behavior is equivalent to computing the rate of reinforcement associated with each delivery (the reciprocal of the interval by which delivery was delayed), averaging these rates, and then taking the reciprocal of that average to obtain the expected delay of reinforcement.

It does not matter whether one is pitting a variable delay against a fixed delay of reinforcement or a probabilistic reinforcement against a certain reinforcement (delivered at a fixed delay). With probabilistic reinforcement, each occurrence of a given option, that is, each occurrence of a given discriminative stimulus, is reinforced with some probability less than one. From a timing perspective, this is another way of producing a VI schedule of reinforcement. Timing theory assumes that the brain processes probabilistic reinforcement schedules as if the relevant variable were the rate of reinforcement, not its probability. Thus, the hyperbolic equation that predicts the results of experiments titrating a fixed-delay option against a VI option should also predict the results of experiments that titrate a fixed-delay option against a probabilistically reinforced option, and this is indeed the case (Mazur, 1989, 1997; Rachlin, Logue, Gibbon, & Frankel, 1986; Rachlin, Raineri, & Cross, 1991).

The harmonic average of different delays is always shorter than the arithmetic average. Thus, harmonic averaging yields the experimentally well-documented preference for a variable delay of reward (positive reinforcement) over a fixed delay with the same (objective) expectation (Autor, 1969; Davison, 1969; Fantino, 1969; Herrnstein, 1964; Mazur, 1984; Pubols, 1962). As this example illustrates, the averaging of income data leads to nonnormative opting behavior. The normative (correct) way to compute expected income is to divide average magnitude (the expected magnitude) by average interreinforcement interval (the expected delay of reinforcement). This quantity is the ratio of the expectations. What subjects do instead is average the income data; that is, they compute the expectation of the ratios (magnitude/delay). Why they do so is a mystery (cf. Bateson & Kacelnik, 1996).

Matching and the Allocating Decision Mechanism

Many discussions of matching behavior—and, in particular, the controversy about matching versus maximizing—rest, we believe, in part on the implicit assumption that rational choice should always take the form of an opting decision; that is, the mechanism should always choose the subjectively better alternative, within the limits established by the noise in the decision variables. Part of the fascination of matching behavior for students of choice and rational decision making is that it appears to violate this principle; the animal frequently leaves the better alternative for the poorer alternative. It does so even when the alternatives differ greatly in richness, so that discrimination failure is unlikely to be a significant factor. We have just seen that when birds are first given the VI schedules as mutually exclusive options, they subsequently almost never confuse the poorer option with the better option, even for options for which the expectations differ by a factor of only three. Matching, by contrast, is obtained with much greater differences (Figure 29, presented later).

It is not obvious, however, that either the subject or the theorist should attend to the average rates of reinforcement. In a concurrent VI schedule, when the scheduling mechanism times out, the scheduled reinforcement is held for the subject to collect whenever it next responds on the manipulandum. The longer the subject has been away from a manipulandum, the more certain it is that a reinforcer has been scheduled on it. Thus, the longer a subject has stuck with one choice, the more certain it is that, at that moment, the other choice is likely to pay off. This consideration is the foundation of the *local maximizing* account of matching behavior (Shimp, 1969). In this account, the subject learns the relation between time away from a manipulandum and probability of immediate payoff. It leaves the richer option for the leaner option when it estimates that the momentary probability of a payoff from the leaner option is higher than the momentary probability of a payoff from the richer option. This account adheres to the principle that the decision mechanism always chooses what appears to be the better option at the moment of choice.

Local maximizing accounts of matching must be distinguished from *global maximizing* accounts (Baum, 1981; Rachlin, 1978; Staddon & Motheral, 1979). The latter emphasize the important point that matching (on concurrent VI schedules and in many other situations) is the rational way to behave in that it maximizes overall return. Because immediate reward is reliably available at the leaner alternative if it has long gone unvisited, subjects should leave the richer alternative for the leaner from time to time. Adaptive advantage is very likely the ultimate (evolutionary) cause of matching behavior. The question is, What is the proximate cause of the decision to leave the richer alternative for the leaner? Global maximization is a theory about ultimate causation, whereas local maximization is a theory of proximate (or "molecular") causation, as is the model that we propose.

Our model of the decision process in matching rests on three well-established experimental findings: the stochastic nature of visit termination, the effect of the relative rate of reinforcement on the likelihood of visit termination, and the effect of the overall rate of reinforcement on the likelihood of visit termination. It builds on earlier theorizing that treated the likelihood of visit termination—hence, the rate of switching between the options—as the fundamental dependent variable (Myerson & Miezin, 1980; Pliskoff,

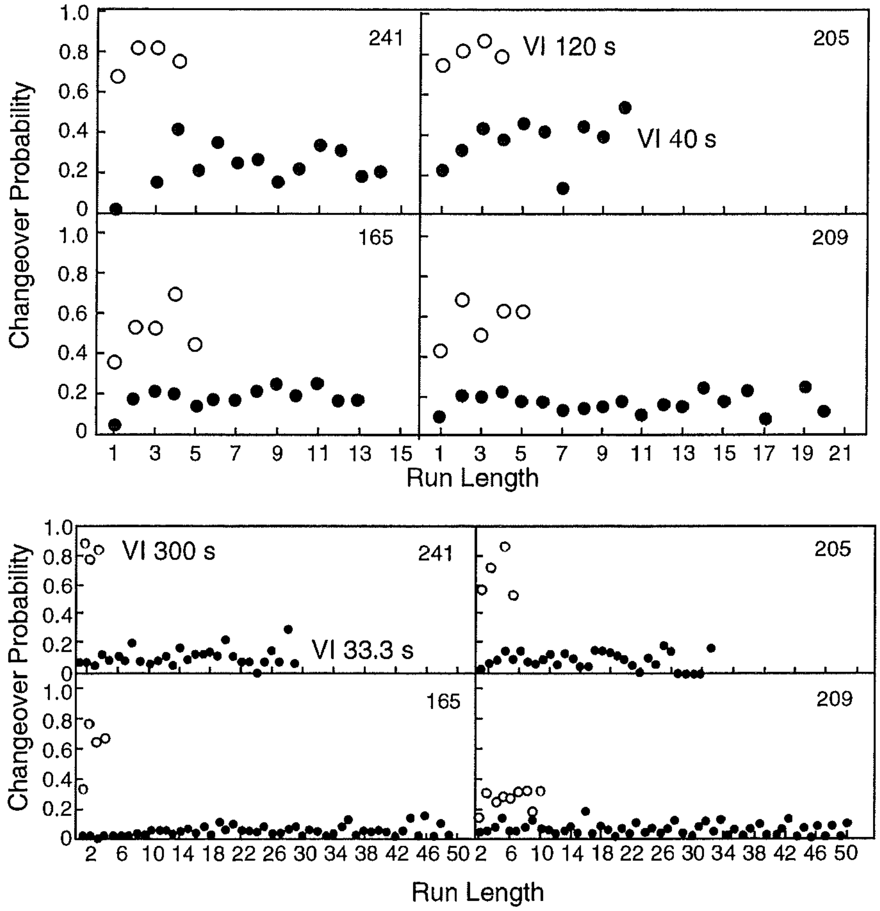

Figure 25. The probability of a changeover key response (a switch to the other option) as a function of the number of responses since the last changeover. Note that probability increases little, if at all, with the length of the run, despite the fact that the longer the run, the more certain it is that there is a reinforcement waiting to be collected from the other option. Note also that the lower the probability of switching, the longer runs last. Expected run length is the reciprocal of the switching probability (leaving rate). VI = variable interval. From "A Markov Model Description of Changeover Probabilities on Concurrent Variable-Interval Schedules," by G. M. Heyman, 1979, *Journal of the Experimental Analysis of Behavior, 31,* pp. 45–46. Copyright 1979 by the Society for the Experimental Analysis of Behavior. Reprinted with permission.

1971). Ours is a model of the decision process that terminates a visit.

Leaving is stochastic. The local maximizing account predicts that the likelihood that the subject will switch from one option to the other increases with the duration of its stay (the interval during which it remains with one option, also called the dwell time). However, Heyman (1979) showed that the momentary probability of a changeover from one option to the other was constant; it did not vary as a function of the number of responses that the pigeon had made on the key it was currently pecking (Figure 25). This means that the rate of switching (the momentary likelihood of visit termination) does not change as stay duration increases, which, in turn, implies that the frequency of visits terminating at a given duration declines by a fixed percentage for each increment in duration. In other words, the frequency of observed departures should be an exponential function of dwell time, which it is (Figure 26).

The discovery that the likelihood of leaving an option at any one moment depends only on the average rates of reinforcement, and not on the momentary likelihood that the switch will be reinforced, led Heyman (1979, 1982) to suggest that the decision process in matching was an elicited (unconditioned) Markov process, in which average rates of reinforcement directly determine the momentary likelihoods that the subject will leave one option for the other. (These momentary likelihoods are hereafter called the "leaving rates.") Heyman argued that matching is not conditioned in the Skinnerian or operant sense; that is, it is not shaped by the prior consequences of such behavior in such a situation (the law of effect); rather, it is simply what the animal is built to do given a certain input. The suggestion that matching is "unconditioned" behavior is very much in the spirit of our models. In our models, an innate (unconditioned) decision mechanism determines what the animal will do given the interreinforcement intervals in its memory.

Effects of relative and overall rates of reward on leaving rate. Heyman (1979) also noted an important property of the leaving rates, namely, that they summed to a constant. Myerson and

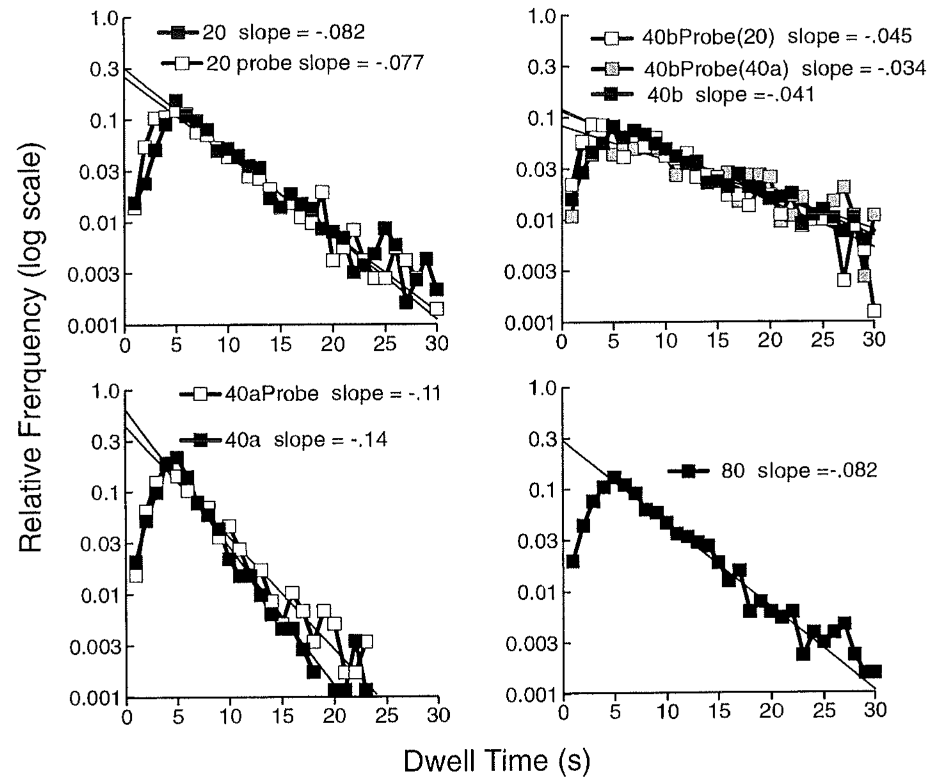

Figure 26. Dwell-time distributions (frequency of stays of a given duration plotted against duration) averaged over 6 subjects. The ordinate is logarithmic, so the linear decline seen after the changeover duration implies an exponential distribution of stay durations. The slope of the regression line through the linear decline is proportional to the leaving rate (momentary likelihood of a departure). The leaving rate depends on the rate of reinforcement available in the alternative schedule; hence, it is approximately 4 times greater when a variable interval (VI) 40-s schedule has been paired with a VI 20-s schedule (lower left panel) than when the same schedule has been paired with a VI 80-s schedule (upper right panel). Notice, also, that the leaving rate is unchanged on the probe trials, when a different alternative is offered. What matters is the remembered alternative rate of reinforcement, not the alternative offered on the probe trial. This figure is redrawn, with recomputed regression lines, from data in Gibbon (1995, Figure 4, p. 213).

Miezin (1980) showed that the same was true in the data of Baum and Rachlin (1969). They further conjectured that the value of the constant to which the leaving rates summed would be an increasing function of the overall reinforcement rate (the sum of the two reinforcement rates). As they noted, this assumption was incorporated into Killeen's (1975) model of choice in the form of an arousal factor, which depends on the overall rate of reinforcement. These arousal assumptions make intuitive sense because the greater the sum of the leaving rates, the more rapidly the subject cycles through the options. To collect reinforcements from both options at close to the rate at which they become available, the duration of a visit cycle (the time consumed in sampling both options) must be no longer than the average overall interval between reinforcements. This interval—the expected interval between reinforcements without regard to source—is the reciprocal of the sum of the rates of reinforcement at the different sources. However, it is a waste of effort to cycle between the options many times during one expected interreinforcement interval. Thus, the time taken to cycle through the options should depend on the expected interval between reinforcements, and it does (Figure 27).

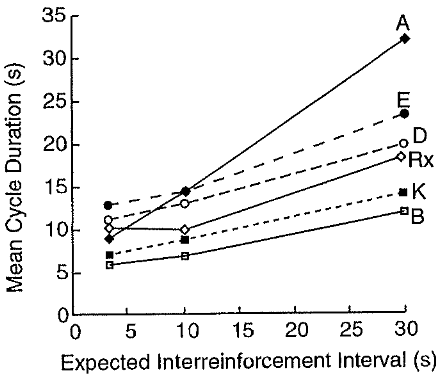

Figure 27. Cycle duration as a function of the expected interval between programmed reinforcement in 6 self-stimulating rats, with concurrent and equal variable interval (VI) schedules of brain-stimulation reinforcement. The overall rate of reinforcement changed, both between sessions and once in each 2-hr session. The expected interreinforcement interval is the reciprocal of the expected overall rate of reinforcement, which is the sum of the reciprocals of the two VIs. The cycle duration includes the changeover (travel time). The data are from Mark (1997).

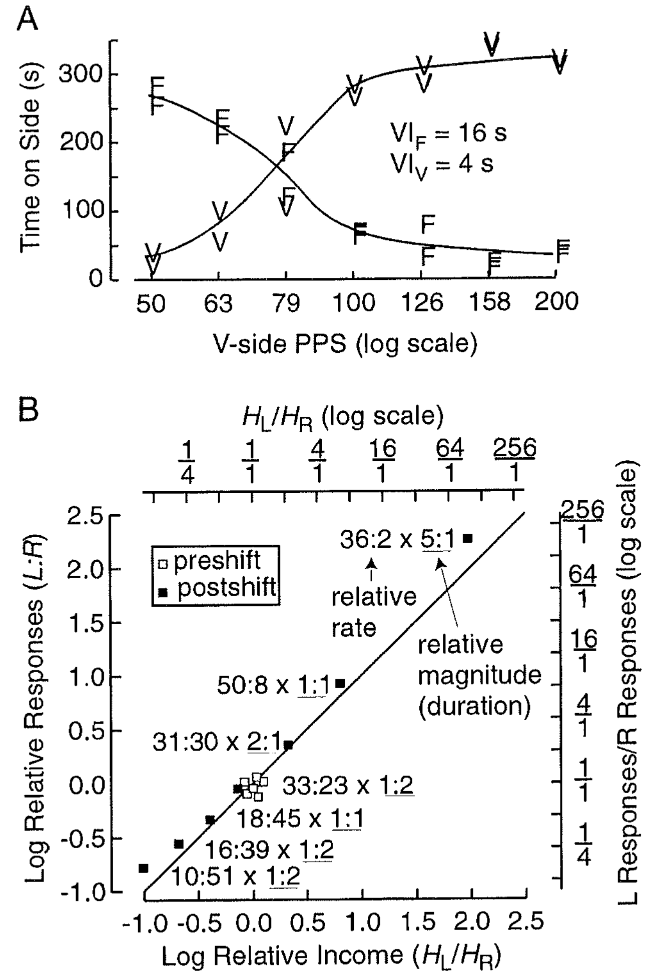

Figure 28. A: Times allocated to competing levers delivering brain-stimulation rewards on concurrent variable interval (VI) schedules to a rat moving back and forth between them so as to exploit them both. The magnitude of the rewards delivered on the F side was fixed. The magnitude of the rewards delivered on the V side was varied from one 10-min trial to the next by varying the pulse frequency in the 0.5-s trains of rewarding pulses. Preference varied from a strong preference for the F side to a strong preference for the V side, depending on the magnitude of the V reward. The strength of the preference is the ratio of the two time allocations. Based on data in Leon and Gallistel (1998, Figure 2, p. 270). PPS = pulses per second. B: Preference for the left (L) key over the right (R) key in a pigeon responding on concurrent VI schedules, as a function of the relative incomes. Relative income (H_L/H_R) is the product of the ratio of the reward magnitudes (M_L/M_R; underlined) and the ratio of the experienced rates of reward (λ_L/λ_R): $H_L/H_R = (\lambda_L/\lambda_R)(M_L/M_R)$. The preshift data are from initial control sessions in which both the scheduled rates of reinforcement and the reward magnitudes were equal. The data are from Keller and Gollub (1977).

Effect of relative reward magnitude on leaving rate. Although reinforcement magnitude has no effect on rate of acquisition, it has a dramatic effect on preference (Figure 28). Indeed, it has a scalar effect; that is, relative stay durations are strictly proportional to relative reward magnitudes. Catania (1963) reported matching when the relative magnitude of the reinforcements was varied rather than their relative rate. When the obtained rates of reinforcement were approximately equal but the duration of food access per reinforcement was, for example, twice as long on one side, pigeons spent approximately twice as much time on the side that produced double the income. The *income* from an option is the amount of reward obtained from that option per unit of time in the apparatus, not per unit of time or number of responses invested in that option. The latter quantity, amount of reward per unit of time invested in an option (or per response), is the *return*. The distinction between income and return is crucial in what follows. Note that it is possible to have a large income from a low return investment and a small income from a high return investment.

The average income is the product of the average rate at which rewards are obtained and the average magnitude of a reward. Put another way, the income from an option is proportional to the rate of reinforcement of that option, with the reinforcement magnitude as the constant of proportionality (scaling factor), and with rate of reinforcement specified in units of session time (not time devoted to the option). Expected income is the same quantity that Gibbon (1977) called "expectancy of reinforcement" and symbolized as *H*, which is the symbol that we use here.

Catania's (1963) result implies that it is relative income that matters in matching behavior, not rate or probability of reinforcement. Moreover, if the same result—time-allocation ratios that match income ratios—is obtained over a wide range of relative magnitudes of reinforcement, that is, if the effect of the relative magnitude of reinforcement is independent of the effect of relative rate of reinforcement, then strong conclusions follow. First, subjective rate and subjective magnitude are simply proportional to objective rate and objective magnitude. Second, subjective rates and subjective magnitudes combine multiplicatively to determine subjective incomes, which are the decision variables that determine how time is allocated among alternatives. Third, the observed time-allocation ratio is equal to the ratio of the subjective incomes; that is, the ratio of the expected stay durations equals the ratio of the expectancies. These are strong conclusions because they assert maximally simple relations between observed objective quantities (rates and magnitudes of reinforcement and time-allocation ratios) and inferred quantities in the head (subjective rates, subjective magnitudes, and the ratios of subjective incomes).

Keller and Gollub (1977) varied both relative reinforcement magnitude (duration of the bird's access to the grain hopper) and relative rate of reinforcement. They found that, for some training conditions at least, the time-allocation ratio was approximately equal to the income ratio (Figure 28B; see also Harper, 1982, for a similar result under group conditions). This finding has not been consistently replicated (e.g., Logue & Chavarro, 1987; see Davison, 1988, for a review), although none of the failures to replicate varied both relative reinforcement magnitude and relative reinforcement rate over a large range, so it is difficult to know how seriously to take the observed departures from matching.

Recently, Leon and Gallistel (1998) tested the extent to which the effect of varying the relative rate of reinforcement was independent of the relative magnitude of the reinforcers. They used brain-stimulation reward and varied reinforcement magnitude, *M*, by varying the pulse frequency in the 0.5-s trains of pulses that constituted individual reinforcements. Thus, the relative magnitude of the reinforcement was varied without varying reinforcement duration. On one side (the F side), the reward magnitude was fixed; on the other side (the V side), the reward magnitude varied from one 10-min trial to the next. It varied from a level near the threshold for performance to the saturation level, which is the largest possible reward producible in a given subject by stimulation through a given electrode.

The programmed relative rate of reward varied from 4:1 in favor of the F side to 1:4 in favor of the V side. The actually obtained rate ratio, λ_V/λ_F, varied by about two orders of magnitude, from about 10:1 in favor of the F side to about 1:10 in favor of the V side. Thus, the experiment tested the extent to which the ratio of times allocated to the F and V sides remained proportional to the relative obtained rate of reward in the face of large differences in the relative magnitude of reward. These large differences in the magnitudes of the rewards produced large differences in the extent to which the animal preferred one side or the other at a given relative rate of reward.

Leon and Gallistel (1998) found that changing the relative rate of obtained reward by a given factor changed the time-allocation ratio by about the same factor, regardless of the relative magnitudes of the reward (except when one of the reward magnitudes was close to the threshold for performance). Over about three orders of magnitude (from time-allocation ratios of 30:1 to 1:30), the time-allocation ratio equaled the ratio of subjective incomes (Figure 29).

The when-to-leave decision. The findings just reviewed imply that allocating behavior is mediated by a stochastic decision mechanism in which the leaving rate for one option is determined by the subjective incomes from it and from the other options. Mark and Gallistel (1994) suggested a model for the when-to-leave decision that directly incorporates these principles. Their model may be seen as a further development of a model by Myerson and Miezin (1980). The decision to leave is generated by a Poisson process (a process formally equivalent to the process of emitting a particle in the course of radioactive decay). The leaving rate (emission rate) is assumed to be determined by the incomes from the two options, in accord with two constraints. One constraint is that the ratio of the leaving rates equals the inverse ratio of the incomes. We formulate this constraint in terms of expected stay durations rather than leaving rates because the symbol for the leaving rate on side i would naturally be λ_i, but we have already used that symbol to represent the reinforcement rate on that side. Also, it is the stay durations that are directly measured. Recall that the expected stay duration, $E(d)$, is the reciprocal of the leaving rate. Therefore, the first constraint is

$$E(d_1)/E(d_2) = H_1/H_2. \tag{4}$$

The other constraint is that the sum of the leaving rates (i.e., the sum of the reciprocals of the expected stay durations) be proportional to the sum of the subjective incomes:

$$\frac{1}{E(d_1)} + \frac{1}{E(d_2)} = c(H_1 + H_2). \qquad (5)^{5}$$

In practice, the sum of the leaving rates (departures per minute of time spent at one or another lever) appears to be a linearly increasing function of the sum of the incomes (Figure 30).

Whereas Mark and Gallistel (1994) built the empirically derived constraints directly into the structure of the mechanism that makes the when-to-leave decision, Gibbon (1995) suggested a psychological mechanism from which these principles may be derived. He proposed that the decision to leave is based on repetitive sampling from the population of remembered incomes. (Recall that each individual reward gives rise to a remembered income or reinforcement expectancy, which is equal to the subjective magnitude of the reward divided by the subjective duration of the interval required to obtain it.) The subject leaves the option that it is currently exploiting whenever the income sample for the other option is greater than the sample from the population for the currently exploited option. When it is not greater, the subject stays where it is (so-called hidden transitions). Gibbon further assumed that the rate at which the subject samples from the population of remembered incomes (called its level of "arousal") is proportional to the overall income (the sum of the average incomes from the two options). Gibbon showed that the likelihood that a sample from one exponentially distributed population of remembered incomes would be greater than a sample from another such population was equal to the ratio of the (scaled) rate parameters for the two distributions. This explains why the ratio of the leaving rates equals the inverse of the income ratio. Also, the higher the rate of sampling from the two populations, the sooner a sample will satisfy the leaving condition. This explains why the duration of a visit cycle (the sum of the expected stay durations, plus the "travel" or switching time) goes down as the overall rate of reinforcement goes up.

These models of the allocating decision process explain the surprising results from a series of experiments designed initially by Williams and Royalty (1989) to test a key assumption underlying previous models of matching, namely, that matching behavior depends on either the probability that a response will be reinforced or the return from an option, that is, the amount of reward obtained per unit of time (or per response) invested in that option (Davis, Staddon, Machado, & Palmer, 1993; Herrnstein & Vaughan, 1980). The assumption that the strength of an operant is a monotonic function of the probability that it will be reinforced is the foundation of most operant theorizing. Thus, these experiments are fundamentally important to our understanding of the foundations of operant behavior.

In these experiments (Belke, 1992; Gibbon, 1995), the subject is trained with two different pairs of concurrent VI options. One pair might be a red key reinforced on a 20-s VI schedule and a green key reinforced on a 40-s VI schedule. During training, the presentation of this pair alternates with the presentation of another pair consisting of a white key reinforced on a 40-s VI schedule and a blue key reinforced on an 80-s VI schedule. Once matching behavior is observed on both pairs, brief unreinforced probe trials are intermingled with the training trials. On a probe trial, the pigeon confronts one key from each pair. Its preference between these two keys is measured by its time- or response-allocation ratio.

The results from these experiments are difficult to reconcile with the assumption that the relative strength of two operants depends on the relative frequency with which they have been reinforced. When the key reinforced on a 40-s VI schedule is the leaner of two options, the subject makes many fewer responses on it than when it is the richer of the two options because it spends

---

[5] The formulation of this assumption in Mark and Gallistel (1994) contained a mathematical error, which is corrected here.

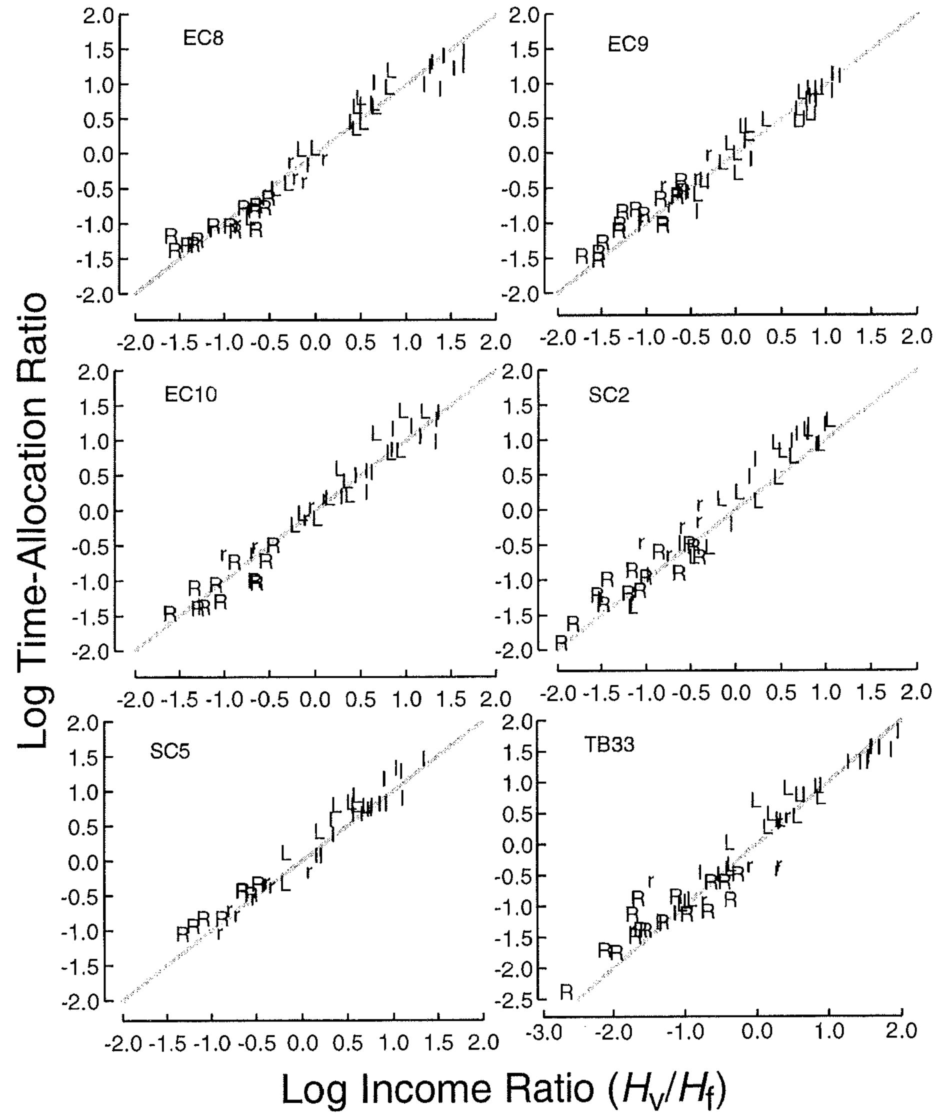

Figure 29. The log of the time-allocation ratio plotted against the log of relative income (H_v/H_f) for 6 rats responding for brain-stimulation reward on concurrent variable interval schedules, with relative rate of reward (λ_v/λ_f) and relative subjective reward magnitude (M_v/M_f) varying over about two orders of magnitude (1:10 to 10:1). The different symbols identify different subconditions. The gray lines represent the identity function, which has a slope of 1 and an intercept at the origin. The letter and number combination in the upper left-hand corner of each panel refers to each of the 6 subjects. From "Self-Stimulating Rats Combine Subjective Reward Magnitude and Subjective Reward Rate Multiplicatively," by M. I. Leon and C. R. Gallistel, 1998, *Journal of Experimental Psychology: Animal Behavior Processes, 24,* p. 272. Copyright 1998 by the American Psychological Association. Adapted with permission.

much more of its time on that alternative in the first case than in the second. Because subjects rapidly cycle between the options, the amount of time (hence, the number of responses) that they invest in each option has little effect on the incomes realized. Because the income remains almost constant while the investment varies (Heyman, 1982), the return realized from an option increases as the investment in it decreases, and so does the probability that a response on that option will be reinforced.

Theories of matching, which assume that relative probability of reinforcement (or relative return) is what determines choice, rely on the fact that the probability of reinforcement goes up as the probability of responding goes down. They assume that the animal adjusts its investments (response probabilities) until the probabilities of reinforcement are equal. Thus, when the leaner key from the richer pair is paired with the richer key from the leaner pair, the subject should prefer the leaner key from the richer pair because the probability of reinforcement it has experienced on that key is much higher than the probability of reinforcement it has experienced on the richer key from the leaner pair (despite the fact that both keys have produced about the same income). In fact, how-

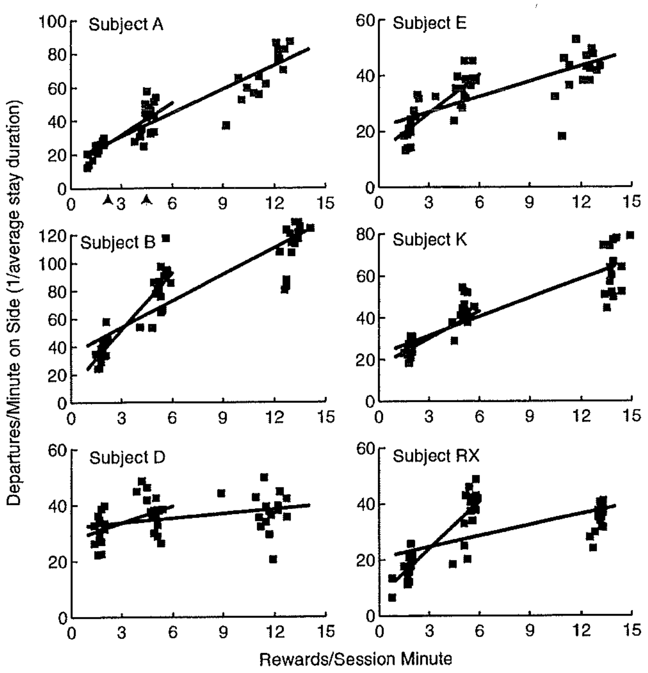

Figure 30. The sum of the leaving rates (total departures/total time on side) as a function of the overall reward rate for 6 rats responding for brain-stimulation reward on concurrent variable interval (VI) schedules of reinforcement. The schedules were the same on both sides and alternated at approximately 1-hr intervals between VI 60, VI 20, and VI 6.7 s. The arrows beneath the abscissa in the upper left panel indicate the programmed overall rates of reinforcement in the probe-trial experiments of Belke (1992) and Gibbon (1995) with pigeon subjects. Because the effect of the overall rate of reinforcement seems to saturate somewhat at the highest reinforcement rates, regression lines have been fitted both to the entire data set and to the data from the two leaner rates of overall reinforcement. The data are from Mark (1997).

ever, the subject prefers the richer key from the leaner pair by a substantial margin (Belke, 1992; Gibbon, 1995). The empirical results are strongly in the opposite direction from the result predicted by the conventional assumption that the relative strength of operants depends on their relative frequency of reinforcement.

It might be thought that the preference for the richer key from the leaner pair is a consequence of a contrast effect, that is, the value of that key has been enhanced by the contrast with a poorer alternative during training. But this would not explain the surprising strength of the preference. On the probe trials, the 40-s VI key from the leaner pair is preferred 4:1 over the 40-s VI key from the richer pair (Belke, 1992; Gibbon, 1995). By contrast, it is preferred only 2:1 over the 80-s VI key with which it is paired during training. Moreover, when the richer key from the richer pair is pitted against the richer key from the leaner pair, the richer key from the leaner pair is preferred 2:1 (Gibbon, 1995), a result that is exceedingly counterintuitive. This last result seems to rule out explanations in terms of contrast.

These startling results are predicted by Mark and Gallistel's (1994) and Gibbon's (1995) models of the when-to-leave decision. Simultaneously solving Equations 4 and 5 for the leaving rates (the reciprocals of the expected stay durations) yields $1/E(d_1) = cH_2$ and $1/E(d_2) = cH_1$. In words, the leaving rate for a key should be proportional to the income from the other key. Or, in terms of expected stay durations, the richer the other option or options, the shorter the expected stay duration. It follows that the expected stay duration for the 40-s VI key for which an 80-s VI key has been the other option is four times longer than the expected stay duration for a 40-s VI key for which a 20-s VI key has been the other option, and twice as long as the expected stay duration for a 20-s VI key for which a 40-s VI key has been the other option. (Compare the slopes of the regression lines in Figure 26, which are proportional to the empirically determined leaving rates.)

On our analysis, the results depend critically on the fact that the stay on each key gets shorter as the overall rate of reward goes up and the cycle duration gets correspondingly shorter. As can be seen in Figure 30, the slope and intercept of the functions relating cycle duration to the overall rate of reward vary appreciably between subjects. Thus, a further test of our explanation would be to determine whether the magnitude of the preference observed on the probe trials in individual subjects can be more accurately predicted if one separately determines for each subject the function relating the sum of the leaving rates to the sum of the reinforcement rates.

Our explanation of the probe results assumes that the decision to leave a key during a probe trial depends not on the income attributed to the key actually present as an alternative on that trial but rather on the income that has generally been available elsewhere when the key that the bird is currently pecking was one of the options. The parameter of the stochastic leaving decision depends on what the subject remembers having obtained elsewhere in the past, not on the income "promised" by the alternative presented on the probe trial. Recall that if a subject is initially trained with concurrent VI keys presented separately—not as options that can be concurrently exploited—then it shows an exclusive preference for the richer key when the keys are presented simultaneously. That result and the results of these probe experiments imply that the decision mechanism that mediates matching is invoked only when past experience indicates that two or more options may be exploited concurrently. Only the remembered incomes from the options that have actually been exploited concurrently affect the leaving rate for a given option. Options that have not been concurrently exploitable, but that are now present, do not enter into the decision to leave an option to sample another.

How the allocating mechanism may produce exclusive preference. An option that is never tried produces no income. Thus, the subjective income from an option must depend on the extent to which the subject samples that option, at least in the limit. As we have already noted, the sampling rate for concurrent VI options is such that large variations in the pattern and amount of sampling have little effect on the incomes realized from the options (Heyman, 1982). This is not true, however, with many other scheduling algorithms. For example, when reinforcements are scheduled by concurrent ratio schedules, which deliver reinforcement after a number of responses that is approximately exponentially distributed about some mean value, then any response-allocation ratio that is less than an exclusive preference for the richer schedule results in an income ratio that is less than the response-allocation ratio. This is an analytic fact, not an empirical fact. It was first pointed out by Herrnstein and Loveland (1975), who went on to note that, with variable ratio schedules, the response-allocation ratio can equal the income ratio only in the limit when all of the responses are allotted to the richer option and, hence, all of the income is obtained from that option. Thus, the exclusive preference for the richer schedule, which is observed with concurrent variable ratio schedules of reinforcement, does not imply that this behavior is controlled by the opting mechanism rather than the allocating mechanism. Exclusive preference for one option can result from destabilizing positive feedback between the subject's time-allocation behavior and the interreinforcement intervals that it experiences. Longer interreinforcement intervals cause diminished sampling of an option, which causes the animal to experience still longer interreinforcement intervals from that option, and so on, until sampling of an option eventually goes to zero (cf. Baum, 1981, 1992).

Summary

A coherent theory of operant choice can be erected on the same foundations as the theory of Pavlovian conditioning. In both cases, probability of reinforcement (whether response reinforcement or stimulus reinforcement) is irrelevant. The behaviorally important variable is the interval between reinforcements in the presence of the CS (in Pavlovian terminology; the discriminative stimulus or secondary reinforcer in operant terminology). Equivalently, the behaviorally important variable is the rate of reinforcement. The behaviorally important effect of the relevant variable is not the strengthening of associations (or of response tendencies). Rather, experienced intervals are stored in memory and read from memory when they are needed in the computations that yield decision variables. Distinct decision mechanisms mediate distinct kinds of decisions. In Pavlovian conditioning, there are distinct mechanisms for deciding whether to respond, when to respond, and whether to quit responding. In operant choice, there are distinct mechanisms for deciding which of two mutually exclusive options to choose (opting behavior) and for allocating behavior among concurrently exploitable options.

The decision variables in operant choice are subjective incomes (reinforcement expectancies), which are subjective reinforcement magnitudes divided by the subjective interval between reinforcements. Under concurrent conditions, in which both options are continuously present, the time used in computing the income from an option is session time (more precisely, the time when both options are present), not the time specifically devoted to (invested in) an option. In discrete-choice paradigms, in which the options are only intermittently available and the choice of one option more or less immediately precludes the choice of the other, the intervals used to compute subjective incomes are the delays between a choice and the delivery of the reinforcement (the intervals when a reinforcement has been chosen and is pending). Other intervals are ignored in the computation of incomes; they are subjectively sunk times.

Because reinforcement magnitude divided by delay of reinforcement goes to infinity as the subjective interval goes to zero, there must be a lower bound on subjective intervals. This lower bound—the subjective instant—establishes a scaling constant for subjective time. This leads to the experimentally well-documented hyperbolic relation between the amount by which a reward is delayed and its subjective value.

In discrete-choice paradigms in which the subjects must compute the value of an option that involves more than one delay of reinforcement, they average the income value of each reinforcement. The reinforcement expectancy in these cases is the expected ratio of individual reward magnitudes and their delays rather than the ratio of the expectancies. The rationale for this kind of averaging is unclear.

Contrasting Conceptual Frameworks

We have elaborated a different conceptual framework for the understanding of conditioning. In this section, we highlight the fundamental differences between this framework and the associative framework that has so far been the basis for all influential models of conditioning.

Different Answers to Basic Questions

That the two conceptual frameworks are fundamentally different is apparent from a consideration of the contrasting answers they offer to the basic questions addressed by the material taught in an introductory course on learning:

1. *Why does the CR appear during conditioning?*

Standard answer: Because the associative connection gets stronger.

Timing answer: Because the decision ratio for the whether-to-respond decision grows until it exceeds a decision threshold.

2. *Why does the CR disappear during extinction?*

Standard answer: Because there is a loss of net excitatory associative strength. This loss occurs either because the excitatory association itself has been weakened or because a countervailing inhibitory association has been strengthened.

Timing answer: Because the decision ratio for the whether-to-stop decision grows until it exceeds the decision threshold.

3. *What is the effect of reinforcement?*

Standard answer: It strengthens excitatory associations.

Timing answer: It marks the beginning or the termination of one or more intervals—an interreinforcement interval, a CS–US interval, or both.

4. *What is the effect of delay of reinforcement?*

Standard answer: It reduces the increment in associative strength produced by a reinforcement.

Timing answer: It lengthens the remembered interreinforcement interval, the remembered CS–US interval, or both.

5. *What is the effect of nonreinforcement?*

Standard answer: The nonreinforcement (the no US) weakens the excitatory association, or it strengthens an inhibitory association.

Timing answer: The timer for the most recent interreinforcement interval continues to accumulate.

6. *What happens when nothing happens (during the intertrial interval)?*

Standard answer: Nothing.

Timing answer: The timer for the background continues to accumulate.

7. *What is the effect of CS onset?*

Standard answer: It opens the associative window in the mechanism that responds to the temporal pairing of two signals. That is, it begins a trial during which the updating of associative strengths will occur.

Timing answer: It starts a timer (to time the duration of this presentation), and it causes the cumulative exposure timers to resume cumulating.

8. *What is the effect of varying the magnitude of reinforcement*?

Standard answer: It varies the size of the increment in the excitatory association.

Timing answer: It varies the remembered magnitude of reinforcement.

9. *Why is the latency of the CR proportional to the latency of reinforcement?*

Standard answer: There is no widely accepted answer to this question in associative theory.

Timing answer: Because the animal remembers the reinforcement latency and compares a currently elapsing interval to that remembered interval.

10. *What happens when more than one CS is present during reinforcement?*

Standard answer: The CSs compete for a share of a limited increment in associative strength, or selective attention to one CS denies other CSs access to the associative mechanism (CS processing deficits), or predicted USs lose the power to reinforce (US processing deficits).

Timing answer: The rate of reinforcement is partitioned among reinforced CSs in accord with the additivity and predictor minimization constraints.

11. *How does conditioned inhibition arise?*

Standard answer: The omission of an otherwise expected US (the occurrence of a no US) strengthens inhibitory associations.

Timing answer: The additive solution to the rate estimation problem yields a negative rate of reinforcement.

12. *What happens when a CS follows a reinforcer rather than precedes it?*

Standard answer: Nothing, or an inhibitory connection between CS and US is formed.

Timing answer: A negative CS–US interval is recorded or, equivalently, a positive US–CS interval. (More precisely, subjective intervals, like objective intervals, are signed.)

13. *How does a secondary CS acquire potency?*

Standard answer: An association forms between the secondary CS and the primary CS so that activation may be conducted from the secondary CS to the primary CS and then to the US by way of the primary association.

Timing answer: The signed interval between the secondary and the primary CS is summed with the signed interval between the primary CS and the US to obtain the expected interval between the secondary CS and the US.

14. *How is CS–US contingency defined?*

Standard answer: By differences in the conditional probability of reinforcement.

Timing answer: By the ratio of the rates of reinforcement.

15. *What is the fundamental experiential variable in operant conditioning?*

Standard answer: Probability of reinforcement.

Timing answer: Rate of reinforcement.

Contrasting Basic Assumptions

Central to the timing framework is the assumption that the nervous system times the durations of the intervals marked off by the events in a conditioning protocol, stores records of these intervals in memory, cumulates successive intervals of exposure to the same CS, and generates CRs through the agency of decision processes, which take stored intervals and currently elapsing intervals as their inputs. None of these elements are found in associative analyses of conditioning. There is no provision for the timing of intervals. There is no provision for the summing of intervals. There is no memory process that stores the result of an interval timing process, and there are no decision processes.

Conversely, none of the elements of associative models are found in timing models. There is no associative bond—no learned signal-conducting connection—thus also no strengthening of connections through repetition and, hence, no associability parameters. There is no notion of a learning trial, and the probability of reinforcement, whether of stimuli or responses, plays no role. Thus, the two conceptual frameworks have no fundamental elements in common. Timing models of conditioning have more in common with psychophysical models in vision and hearing than with associative models in learning. Like models in psychophysics, they focus on quantitative aspects of the experimental data. Like modern perceptual theories, they embody computational principles that resolve ambiguities inherent in the input to yield an unambiguous percept (representation) of the state of the world.

Elementary acquisition event in associative models. The elementary event in the associative conception of acquisition is a change in the strength of a connection. Repetitions of the same learning experience—for example, repetitions of the temporal pairing of a tone and a puff of air directed at the eye—strengthen the connection. Thus, what the associative conception appears to require at the neurobiological level is a mechanism for altering the strength of a synaptic connection between neurons. The enduring appeal of the associative conception is this neurobiological transparency. A second strong appeal is the straightforwardness of its explanation for the (presumed) gradual increase in the strength of the CR. Thus, on the one hand, the strengthening of the CR is naturally seen as a consequence of successive increments in the underlying connection strengths.

On the other hand, the assumption of a conductive connection whose strength is incremented over successive repetitions places serious obstacles in the way of an associative explanation for the fact that animals learn the durations of intervals, the magnitudes of reinforcements, the intensities of CSs and USs, and other parameters of the experimental protocol The size of an increment in associative strength is a function of more than one aspect of the events on a given trial—the magnitude and delay of reinforcement, the pretrial strength of the association, strengths of other associations to the same US, and so on. Because of this confounding of experiential variables, none of these variables are recoverable from the current strength of the connection. Put more formally, the strength of an associative connection is a many–one function of different properties of the conditioning protocol, and many–one functions are not invertible; you cannot get from the one back to the many.[6]

---

[6] It might, however, be possible to get from many back to the many; that is, a manifold of associative connections might—with a very careful choice of connection-forging processes—be such that one can recover from that manifold the values of the many experiential variables that created it. However, the associative processes invoked to explain animal conditioning do not have the properties required to make an associative manifold invertible. It is not clear that it is possible to modify the association-forming process in such a way as to make the associative manifold invertible without eliminating the neurobiological transparency and straightforward explanation of gradual response acquisition that account for much of the appeal of associative models.

Elementary acquisition event in timing models. The elementary event in the timing conception of conditioning is the measuring and recording of an elapsed interval. Learning is the product of an information-gathering system that functions automatically. Each repetition of the same experience (e.g., each trial in an eyeblink conditioning experiment) lays down a new record (cf. Logan, 1988). The system also keeps running totals for the cumulative durations of the salient stimuli. Thus, what the timing conception requires at the neurobiological level is a mechanism capable of cumulating and storing the values of a large number of distinct variables.

It is not difficult to suggest cellular mechanisms capable of storing the values of variables. Miall (1996) proposed network models for accumulation, a central feature in timing models of acquisition, and for storage and comparison processes (see also Fiala, Grossberg, & Bullock, 1996; Grossberg & Merrill, 1996, for network timing models). Although these models do not have anything like the complexity of the systems we have proposed, they do illustrate the feasibility, in principle, of a neural-net kind of representation of accumulated elapsed time. Enduring changes in synaptic strength could, in principle, be used to record the values of variables, albeit at the cost of sacrificing those properties that make the change constitute a change in associative strength, as traditionally understood.

In sum, what timing models require is a memory functionally analogous to a conventional computer memory, a mechanism for storing and retrieving the values of variables. Although it is not immediately apparent how to create such a memory from the currently understood elements of neurobiology, such a memory is both physically possible (witness computer memory) and biologically possible. Genes provide the proof of biological possibility; they store the values of variables for readout in response to gene-activating stimuli. Note that genes are repositories of information, not paths for signal flow. Genes are read by gene-activating signals, just as computer memories are read by memory-activating signals. Fifty years ago, the physicochemical mechanisms for these genetic operations were deeply mysterious. What timing models require neurobiologically is a selectively activatable repository for the values of experientially determined variables. The neurobiological mysteriousness of this requirement is no greater than the biochemical mysteriousness of the gene theory's requirement for a self-replicating molecule was in, say, 1950.

Decision mechanisms versus no decision mechanisms. Associative models do not have decision mechanisms that take remembered values as their inputs and generate CRs as their outputs, whereas decision mechanisms are central to the timing perspective. (Miller and Schachtman's [1985] comparator hypothesis is an exception to this generalization, which is why it represents a step in the direction of the kind of model for which we argue.) Timing models of conditioning share with psychophysical models of vision and hearing the explicit, formal specification of decision processes. Associative models, by contrast, have long lacked an explicit specification of the process by which the strengths of associations translate into observed behavior (Miller & Matzel, 1989; Wasserman & Miller, 1997). The lack of decision mechanisms in associative models goes hand-in-hand with the lack of a mechanism for storing and retrieving the values of variables because decision mechanisms take the values of remembered variables as their inputs.

Challenges

Quite aside from its intrinsic value as an explanatory framework that integrates a wide variety of findings and suggests many new experiments, our alternative to the associative conceptual framework is valuable because it brings into clear focus several enduring puzzles that have arisen in the standard associative explanations of conditioned behavior. This challenge may stimulate the further development of associative models.

Understanding Partial Reinforcement

The fact that partial reinforcement does not affect the number of reinforcements to acquisition and does not reduce the number that must be omitted to produce extinction is a puzzle. None of the formally specified associative models with which we are familiar can account for this because they all assume that nonreinforced trials weaken the net excitatory effects of reinforced trials. Therefore, the number of reinforced trials required to reach a given level of excitatory effect ought to increase as the schedule of reinforcement gets thinner, but it does not, or does so only slightly.

Also, the net excitatory effect after a given number of reinforcements ought to be weaker the thinner the schedule of reinforcement is. Therefore, the number of omitted reinforcements required to produce extinction ought to be reduced as the schedule of reinforcement is thinned, but it is not, or only slightly. Gibbon (1981a) showed that the asymptotic strength of the CS–US association in the Rescorla–Wagner theory is

$$\frac{p\beta_l}{p\beta_l + (1-p)\beta_e}, \tag{6}$$

where p is the probability that the CS is reinforced, and β_l and β_e are the learning and extinction rate parameters, respectively. When the rates of learning and extinction are equal, then the βs cancel out, and Equation 6 reduces to p. Regardless of the values of the βs, the asymptotic strength of the association declines as the probability of reinforcement declines. Trials (and omitted reinforcements) to extinction should be reduced correspondingly, but, in fact, partial reinforcement increases trials to extinction and does not change omitted reinforcements to extinction.

As the extinction rate (β_e) is reduced relative to the learning rate (β_l), the amount by which partial reinforcement reduces asymptotic associative strength is reduced. However, the effect remains large for any plausible ratio of extinction rates to learning rates. The extinction rate cannot be made too much slower than the learning rate because the lower the ratio of the extinction rate to the learning rate, the longer extinction should take relative to learning, which brings us to a second problem.

It is puzzling that the number of reinforcements that must be omitted to produce extinction can be substantially less than the number of reinforcements required for acquisition. To avoid catastrophic effects of partial reinforcement on acquisition, associative models generally assume that the rate of extinction is less than the rate of acquisition (associations get stronger faster than they get weaker). In that case, extinction ought to take longer than acquisition, which is not generally the case. Indeed, by reducing the ratio of the average intertrial interval, *I*, to the average trial duration, *T*, to 1.5:1, one can create a protocol in which the number of reinforcements required for acquisition is more than twice the number of omitted reinforcements required for extinction. At this *I*/*T* ratio, subjects can still be conditioned on a 10:1 partial-reinforcement schedule—with no more reinforcements than are required under continuous reinforcement! These findings appear to challenge the foundational assumption that reinforcement and nonreinforcement have opposing effects on the net excitatory effect of CS–US associations.

It has been pointed out that the basic assumptions of associative conditioning theories about the strengthening and weakening effects of reinforcement and nonreinforcement fail to account for the microstructure of performance under partial reinforcement (the trial-by-trial pattern; Gormezano & Coleman, 1975; Prokasy & Gormezano, 1979; see also Capaldi & Miller, 1988). They also fail to account for the macrostructure, when trials to extinction are considered alongside trials to acquisition.

Timescale Invariance

None of the formally specified associative models with which we are familiar account for the timescale invariance of the acquisition process. They all assume that delaying reinforcement reduces associability. Indeed, neurobiologically oriented associative theories often take a narrow window of associability as the signature of the associative mechanism (e.g., Gluck & Thompson, 1987; Grossberg & Schmajuk, 1991; Usherwood, 1993). The problem with the assumption that delaying reinforcement reduces associability is that delaying reinforcement has no effect if the intertrial interval is increased proportionately. This is a manifestation of timescale invariance: Changing the relevant time intervals by some scaling factor does not affect the results.

Another basic problem for associative theories is that conditioning depends on a contingency between the CS and the US (or the instrumental response and the reinforcement). Associative theories, at least those that have influence in neurobiological circles, assume that conditioning is driven by temporal pairing. Contingency is a global statistical property of the animal's experience. Like all such properties, it is timescale-invariant. It is difficult to see how the operation of a mechanism that is activated by temporal pairing can be timescale-invariant, because the concept of temporal pairing would seem to be a clear example of a concept that is dependent on timescale. Defining temporal pairing in a manner that is timescale-invariant would necessitate radical rethinking of the mechanisms that could mediate the effects of this temporal pairing.

The inverse proportionality between reinforcements to acquisition and the trial-to-trial interval also lacks an explanation. The trial-spacing effect is explained qualitatively by the assumption that it gives more scope for the extinction of conditioning to the background (Durlach, 1989; Rescorla & Durlach, 1987). However, Gibbon (1981a) showed that the Rescorla–Wagner theory predicted only a weak effect of intertrial interval on reinforcements to acquisition while predicting a strong effect of partial reinforcement—the opposite of what is observed empirically.

We believe that other associative models would make the same predictions regarding the relative potencies of these two basic variables if they were modified so as to make predictions regarding the effect of the intertrial interval. It is hard to say what most associative models, as they now stand, would predict about the relative effects of varying partial reinforcement and the intertrial interval on the rates of conditioning and extinction, because they cannot predict the effect of the intertrial interval at all without making use of the assumption of multiple background "trials" during one intertrial interval. This assumption brings up the "trial problem."

The Trial Problem

The notion of a trial is a fundamental but insufficiently scrutinized notion in most associative models. A trial is a discrete interval of time during which events occur that cause the updating of associative connections. In Pavlovian conditioning, if both a CS and a US occur on a trial, it is a reinforced trial; if a CS occurs but not a US, it is an unreinforced trial. In instrumental conditioning, a trial is an interval during which there is an input (a stimulus), an output (a response), and, finally, an error-correcting feedback (reinforcement or nonreinforcement). This latter conception of a trial also applies to the many contemporary associative network models that use an error-correcting algorithm to update associative strengths ("supervised learning algorithms"). The notion of a trial is intimately linked to the notion that it is the probability of reinforcement that drives learning because probabilities cannot be defined in the absence of trials (Granger & Schlimmer, 1986).

We discern three different traditions governing the identification of the theoretician's trial with the elements of a conditioning protocol. In one conception, which one might call the neurobiological conception of Pavlovian conditioning, the beginning of a trial corresponds to the onset of a CS, and the termination of a trial corresponds to the closing of the window of associability. The onset of a CS opens the window of associability: If a reinforcement occurs while the window remains open, it is a reinforced trial; if a reinforcement does not occur while the window is open, it is an unreinforced trial. Another trial does not begin until there is a further CS onset.

In the second, more pragmatic conception of a Pavlovian trial, a trial begins when a CS comes on, and it terminates when the CS goes off (or soon thereafter). If a reinforcement occurs while the CS is on (or soon thereafter), it is a reinforced trial; if it does not, it is an unreinforced trial.

Finally, in the instrumental tradition, a trial begins when the animal makes a response. The response opens a window of associability. If reinforcement is delivered while the window is open, the association between the response and the stimuli present when it was made is strengthened. If reinforcement does not occur soon after the response is made, it is an unreinforced trial.

The trouble with the first conception is that it is known to be empirically indefensible. It has not been possible to define by experiment when the window of association opens or how long it stays open (see Rescorla, 1972, for a review of such efforts). That is, it has never been possible to define temporal pairing in the simple way that this traditional conception suggests that it should be defined. Indeed, the timescale invariance of the acquisition process would appear to be irreconcilable with such an assumption.

The trouble with the more pragmatic notion of a Pavlovian trial is that it cannot be applied in the case of background conditioning, or in any of the conditioning protocols in which there are many reinforcements during a sustained CS presentation. Such protocols are common in operant paradigms, in which a stimulus present at the time of reinforcement is called a secondary reinforcer, a discriminative stimulus, or both. The stimuli reinforced in operant paradigms (the stimuli projected onto the keys that the pigeons peck) are often continuously present, as, for example, in concurrent schedules of reinforcement. In Pavlovian conditioning, background or context conditioning is empirically well-established and theoretically important: Reinforcements that occur while an Aplysia, a pigeon, or a rat is in an experimental chamber establish a CR to the chamber (Baker & Mackintosh, 1977; Balsam, 1985; Balsam & Schwartz, 1981; Colwill, Absher, & Roberts, 1988; Rescorla, 1968, 1972). The more frequent the reinforcer, the stronger the background conditioning is (Mustaca, Gabelli, Papine, & Balsam, 1991).

When the pragmatic conception of a trial is applied to a background-conditioning protocol, each experimental session constitutes one trial because the CS (the background) "comes on" at the beginning of the session, when the subject is placed in the apparatus, and terminates with the end of the session, when it is removed. How to deal with the effects of reinforcer frequency is then problematic. To get around this problem, Rescorla and Wagner (1972) posited a trial "clock," which carved time into purely subjective trials. This clock permitted them to treat the intertrial intervals, when the background alone was present, as consisting of sequences of multiple internally timed event-independent trials (autotrials, for short), with the association between the background and the US strengthened or weakened accordingly as a reinforcer did or did not happen to occur during such a "trial." Their immensely influential analysis of the effect of background conditioning on conditioning to a transient CS depended on this autotrial assumption just as strongly as on the much better known assumption that associations to a given US compete for an asymptotically limited total associative strength. But Rescorla and Wagner seem to have regarded this assumption as a temporary theoretical expedient. There has been no attempt to explore its consequences. We believe that such an attempt would uncover unfortunate implications (cf. Granger & Schlimmer, 1986).

The problem is that, in many cases, it appears necessary to assume that autotrials are very short (on the order of 1 s or less). If they are allowed to be as long as Rescorla and Wagner (1972) assumed them to be for the purposes of their analysis—2 min—then, in many conditioning protocols, one again encounters the problem that a single trial encompasses more than one presentation of both the CS and the US. And two CSs (or a CS and a US) that did not in fact coincide are counted as coinciding because they both occurred during one autotrial. However, if autotrials are assumed to be very short, then many protocols have very large numbers of unreinforced autotrials. Unreinforced trials weaken excitatory associations (as in the Rescorla–Wagner theory); or they strengthen inhibitory associations, which negate the behavioral effects of excitatory associations; or they reduce attention to the CS. In any event, the numerous unreinforced trials introduced into the analysis by the (very short) autotrials assumption would seem to make the buildup of any appreciable net excitatory effect impossible. The effects of rare reinforced autotrials are swamped by the effects of the frequent unreinforced autotrials. One is forced to assume that the effects of nonreinforcement are extremely weak relative to the effects of reinforcement, but then it becomes difficult to explain the results of experiments on extinction and conditioned inhibition. These experiments show that conditioned inhibition can develop rapidly (Nelson & Bouton, 1997) and that extinction can occur more rapidly than acquisition. For an indication of the difficulty that even very complex associative theories confront in explaining context conditioning, see Mazur and Wagner (1982, p. 33f).

The operant conception of a trial has similar problems. The notion of the probability of response reinforcement is not definable in the absence of a discrete interval of specified duration following the response, during which a reinforcement either is or is not received. However, there are no data that indicate what this interval might be. One can delay reinforcement for long intervals, provided that intertrial intervals are made correspondingly long. Indeed, the phenomenon of autoshaping, as we have already noted, calls into question the distinction between the Pavlovian and operant paradigms. It is not clear that conditioning in any paradigm actually depends on the animal's making a response. Conditioning in operant paradigms, as in Pavlovian paradigms, appears to be driven by the temporal relations between events, including events initiated by the subject, such as key pecks, leverpresses, and chain pulls. Insofar as an operant response is irrelevant in operant conditioning—or relevant only insofar as it constitutes a distinguishable kind of event—then the operant conception of a trial (response-initiated trial) becomes inapplicable.

If the trial notion is theoretically indispensable but operationally undefinable—that is, if one cannot say when a given segment of a protocol constitutes a trial—then one must ask whether associative theories can actually be applied to the phenomena they seek to explain. If these models are to be used as a guide to mechanisms to be sought at the neurobiological level of analysis, they are going to have to be more specific about what constitutes a trial.

Paradoxical Effects of Reinforcement Magnitude

An animal's preference for one concurrent VI schedule of reinforcement over another is proportional to the relative magnitudes of the reinforcements (Catania, 1963; Keller & Gollub, 1977; Leon & Gallistel, 1998; Mark & Gallistel, 1993). The natural associative interpretation of this is that bigger reinforcements produce bigger increments in associative strength and, therefore, a bigger asymptotic associative strength. In the Rescorla–Wagner theory, for example, the parameter lambda, which is the asymptotic net associative strength that a US can support, is generally taken to be a function of the magnitude of reinforcement. That assumption has the just-specified consequence, at least qualitatively—greater preference for bigger rewards. However, if bigger reinforcements increase the size of the increments in associative strength, they ought to increase the rate of acquisition, but they do not. Thus, from an associative perspective, there is a paradox between the strong effect of reinforcement magnitude on preference (Figure 28) and its negligible effect on the rate of acquisition (Figure 12).

The Problem of No Unconditioned Stimulus

Extinction and conditioned inhibition require an associative change in response to the failure of a US to occur. For a CS to become an inhibitor, it must signal the nonoccurrence of an otherwise expected stimulus (LoLordo & Fairless, 1985; Rescorla, 1988). It is unclear how to conceptualize this within an associative context. Both Pavlov (1928) and Hull (1943) wrestled with the question of how the nonoccurrence of a stimulus could be the cause of something. We do not think the question has been solved in the decades since. Dickinson (1989) referred to this as the "no-US problem."

As Dickinson (1989) pointed out, there is nothing inherently puzzling about a process set in motion by the failure of something to occur. In a device like a computer, which is capable of comparing input values against internally generated or stored values (i.e., against an *expectation,* in the sense that this term is used in this article), the failure of an input to occur generates a signal from the comparison process. This signal is proportional to the difference or ratio between the input magnitude (zero, in the case of failure) and the comparison magnitude (the expectation). The discrepancy signal initiates whatever events are to be consequent on the failure. This is the nature of our explanation of extinction: The decision to stop responding is based on the ratio between a currently elapsing unreinforced interval and an expected interreinforcement interval. Indeed, an expected interval between reinforcements is the denominator in each of the decision ratios that appear in the timing analysis of acquisition, extinction, and response timing.

The basic problem is that, in associative theory, the nonoccurrence of reinforcement (the occurrence of a no US) is generally treated as an event. This event sets in motion the mechanism that responds to nonreinforcement. Even if one adds to an associative model the machinery necessary to have an expectation (as is implicitly or explicitly done in the Rescorla–Wagner [Rescorla & Wagner, 1972] model and many other contemporary associative models), it is still unclear how to make the associative analysis go through, because an event has to have a time of occurrence, and there does not appear to be a principled way of saying when a nonreinforcement has occurred. Acquisition and extinction proceed normally when reinforcements are delivered by random rate or Poisson scheduling mechanisms, and a random rate of background reinforcement is generally used in the explicitly unpaired protocol for the production of conditioned inhibition. In these schedules, reinforcement is no more likely at any one moment than at any other. Thus, there is no basis for expecting the reinforcement to occur at any particular moment. How can the changes that underlie extinction and inhibition be set in motion by the failure of an expected reinforcement to occur if that reinforcement is no more likely at any one moment than at any other? It would seem that either the failure must be deemed to happen at every moment, or it must be deemed never to happen. In either case, it is unclear how to make a physically realizable, real-time model of extinction and inhibitory conditioning within the associative conceptual framework, without recourse to the trials assumption, which is itself deeply problematic.

Directionality

In associative models, a special class of stimuli, reinforcing stimuli, sets the associative mechanism in motion and confers directionality on the associative process. If there is no stimulus that can be identified as a reinforcing stimulus, then there is nothing to set the associative process in motion (nothing to cause learning), and there is no way of specifying directionality, that is, no way of saying whether one is dealing with forward or backward conditioning. This makes the phenomenon of sensory preconditioning an awkward phenomenon for associative theories to come to terms with. On the face of it, if one does not believe that there is a special class of stimuli called reinforcing stimuli, then sensory preconditioning is simply conditioning: Two stimuli, say, a tone and a light, are temporally paired, and they become associated. The problem arises because most theories of the associative process assume that it matters which stimulus comes first, the CS or the US. If the CS comes first, it is forward conditioning. If the US comes first, it is backward conditioning. The associative mechanism itself is assumed to be sensitive to temporal order. Different orderings of the stimuli being associated produce different associative effects. If they did not, then it would not matter whether Pavlov rang the bell before or after presenting food to his dogs, which we know to be false. However, when two neutral stimuli are paired, there is no way to specify whether one has to do with forward or backward conditioning.

From a timing perspective, reinforcers play no privileged role in learning per se. Like other stimuli, their onsets and offsets may mark either end or both ends of a timed interval. The memory of reinforcement plays a privileged role only in determining whether a decision mechanism is used to determine a response. Neutral stimuli—stimuli that do not elicit stimulus-specific unconditioned responses—do not evoke CRs because they are not intrinsically important to any behavior system (Fanselow, 1989; Timberlake & Lucas, 1989). For a behavior-determining decision to be made, a behavior system must operate. That is, the animal must be motivated to make use of the CSs in the control of its behavior. Reinforcers are simply motivationally significant stimuli, stimuli whose occurrence motivates behavior that anticipates that occurrence. They are not important for learning, only for performance.

Second-Order Conditioning

In the Backward, Second-Order, and Trace Conditioning section, we reviewed recent findings on trace conditioning versus delay conditioning, backward conditioning versus forward conditioning, inhibitory conditioning, and second-order conditioning from Ralph Miller's laboratory (Barnet & Miller, 1996; Cole et al., 1995a, 1995b). These findings pose a challenge to traditional conceptions.

The challenge is to explain the reversals in the apparent relative strengths of primary CS–US associations when their strengths are tested directly and indirectly (Barnet & Miller, 1996; Cole et al., 1995a, 1995b). In the direct tests, when the strength of the CR to the primary CS is measured, backwardly conditioned and trace-conditioned stimuli appear to have considerably weaker associations with the US than forwardly conditioned and delay-conditioned stimuli. But in the indirect tests, when the strength of the CR to CSs that have been linked to the primary CSs through identical second-order conditioning is measured, the backwardly conditioned and trace-conditioned primary stimuli appear to have much stronger associations with the US than the forwardly conditioned and delay-conditioned primary stimuli. The results from Miller's laboratory seem to require the kind of assumption that is the foundation of the timing perspective, namely, that in the course of conditioning, the subject acquires quantitative knowledge of the temporal relations between events, and that it is this quantitative knowledge and inferences drawn from it that determine the CR.

Reinstatement

Associative models conceive of extinction in one of two ways: (a) as the weakening of an association (e.g., Pearce, 1994; Rescorla & Wagner, 1972) or (b) as the development of an opposing association (or negating process of some kind) whose effects cancel the effects of the excitatory association. The first assumption is almost universally adhered to in connectionist models. However, it is now widely recognized by experimentalists to be irreconcilable with a large literature demonstrating that conditioning is forever: No subsequent experience can expunge the memories (associations–connections) implanted (strengthened) by earlier conditioning (Bouton, 1991, 1993; Bouton & Ricker, 1994; Brooks, Hale, Nelson, & Bouton, 1995; Kaplan & Hearst, 1985; Mazur, 1996; Rescorla, 1992, 1993; Rescorla & Heth, 1975; see Rescorla, 1998, for a review). To be sure, subsequent experience can lead to the disappearance of the CR, but a variety of reinstatement procedures show that the "associations" (the memories implanted during earlier conditioning) remain when the CR has disappeared. The simplest of the reinstatement procedures is to wait a while after extinction before testing for the presence of the CR. The response comes back (e.g., Rescorla, 1996)—the phenomenon that Pavlov (1928) dubbed "spontaneous recovery." Changing the spatial context (the experimental chamber) also brings the response back (Bouton & Ricker, 1994).

Bringing back the CR after its extinction by changing the temporal or spatial context is called *reinstatement.* The finding that it is almost always possible to reinstate an extinguished (or even counterconditioned) response has led specialists in conditioning to favor the view that extinction involves the development of an association with an opposing (or canceling) effect rather than the weakening of the originally conditioned association—a view favored by Pavlov himself. However, it is not clear how this solves the problem. The problem is to explain why reinstatement procedures make old excitatory associations prevail over newer inhibitory (or canceling) associations.

From the perspective of the analysis of extinction expounded here, a precondition for extinction is that the animal not forget the path to its present estimates of the state of the world. In this analysis, extinction is a consequence of a decision mechanism designed to detect a change in the rate of reinforcement attributed to the CS, that is, to decide whether recent experience is inconsistent with earlier experience. This requires that recent experience be compared with earlier experience. A precondition for that comparison is that recent experience be kept distinct from more remote experience in memory. To detect a change, the system must not represent its experience solely by means of a running average, which is a common assumption in associative models of conditioning. It cannot use a simple running average of its experience to determine its behavior because a segment of the recent past must be compared with the earlier past.

Devenport and Devenport's (1994) temporal weighting rule points the way to the treatment of reinstatement phenomena from a timing perspective. The general principle indicated by a variety of findings (e.g., Mazur, 1996) is that animals track their experience on several different timescales. When recent experience indicates one state of affairs but more remote, albeit more extensive, experience indicates a different state of affairs, the animal favors the more recent experience so long as it is indeed recent. As that experience becomes less recent, the animal begins to base its behavior more on the more remote but more extensive past. In other words, an important variable in determining which part of its remembered experience dominates an animal's behavior is the time elapsed since the experience, not because of forgetting but because the recency of one experience relative to conflicting or different experiences is itself an important decision variable. It is used to decide which of two conflicting previous experiences provides the more plausible current expectation. This account assumes that a representation of how long it has been since a particular conditioning experience is an important determinant of conditioned behavior (cf. Clayton & Dickinson, 1998).

A phenomenon similar to reinstatement has recently been demonstrated in operant choice. Mazur (1995) showed that when pigeons experience a change in the relative rate of reward after a long period of stability, they adjust to the change within the session in which they encounter it, but at the beginning of the next few sessions, they show clear reversion to the prechange time-allocation ratio, followed by an increasingly rapid return to the postchange time-allocation ratio. These results and Devenport and Devenport's (1994) theory suggest that there is yet another decision mechanism to be investigated, namely, the mechanism that decides, in effect, which context should be the basis of the animal's current behavior—an older context or a more recent context. This decision will presumably depend on the animal's previous experience with changes in rates of reinforcement, which will indicate, among other things, how long such changes may be expected to last.

Conclusion

In strictly behaviorist conceptual frameworks, conditioned behavior does not depend on the acquisition of symbolic knowledge of the world. The conditioning experience rewires the nervous system so that the animal behaves more adaptively, but the nervous system does not thereby acquire any internal structure that specifies objective properties of the experience that caused the rewiring. Not only was this the burden of Skinner's (1938) argument; it also was the main point of Hull's early writings (e.g., Hull, 1930). Arguments of a similar nature have been advanced much more recently by some connectionists (Churchland & Sejnowski, 1992; Smolensky, 1986). However, much recent experimental research involving multiple stimuli, multiple types of responses, and multiple kinds of reinforcers has demonstrated that the subjects in conditioning experiments do remember which stimuli and which responses lead to which reinforcements (Colwill, 1993; Rescorla & Colwill, 1989). Timing experiments demonstrate that subjects also learn the temporal intervals in the protocols. Animal subjects have also been shown to learn the time of day at which they were conditioned (see Gallistel, 1990, for a review) and to know which reinforcers were cached where and how long ago (Clayton & Dickinson, 1998, 1999; Sherry, 1984). Clearly, then, conditioning causes changes in the internal structure of the brain that permit the animal to recall objective properties of the conditioning experience.

We believe it is time to rethink conditioning in light of these findings, because the strengths of associative bonds, at least as traditionally conceived, cannot readily serve to specify objective facts about the animal's past experience. In this review, we have attempted to show that timing theory provides a powerful and wide-ranging alternative to the traditional conceptual framework. There is, of course, much that remains to be explained. We believe, however, that we have covered enough territory to demonstrate the power and advantages of this conceptual framework.

Timing theory brings theories of conditioning into the same conceptual and methodological domain as theories of sensory processing and perception—the domain of information processing and decision theory. And it invites us to consider quantitative aspects of the conditioning data much more closely than they have so far been considered. Contrary to the popular impression, simple and profoundly important quantitative regularities are discernible in the experimental data, for example, the timescale invariance of the conditioning process and of the distributions of CRs.

References

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 33B,* 109–121.

Anderson, N. H. (1969). Variation of CS–US interval in long-term avoidance conditioning in the rat with wheel turn and with shuttle tasks. *Journal of Comparative and Physiological Psychology, 68,* 100–106.

Annau, Z., & Kamin, L. (1961). The conditioned emotional response as a function of intensity of the US. *Journal of Comparative and Physiological Psychology, 54,* 428–432.

Aronson, L., Balsam, P. D., & Gibbon, J. (1991, November). *Context value modulates levels of response recovery after random control training.* Paper presented at the meeting of the Psychonomic Society, San Francisco.

Autor, S. M. (1969). The strength of conditioned reinforcers as a function of frequency and probability of reinforcement. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 127–162). Homewood, IL: Dorsey Press.

Baker, A. G., & Mackintosh, N. J. (1977). Excitatory and inhibitory conditioning following uncorrelated presentations of the CS and US. *Animal Learning and Behavior, 5,* 315–319.

Baker, A. G., & Mercier, P. (1989). Attention, retrospective processing and cognitive representations. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 85–116). Hillsdale, NJ: Erlbaum.

Balsam, P. (1985). The functions of context in learning and performance. In P. Balsam & A. Tomie (Eds.), *Context and learning* (pp. 1–21). Hillsdale, NJ: Erlbaum.

Balsam, P. D., & Payne, D. (1979). Intertrial interval and unconditioned stimulus durations in autoshaping. *Animal Learning and Behavior, 7,* 477–482.

Balsam, P., & Schwartz, A. L. (1981). Rapid contextual conditioning in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 382–393.

Barela, P. B. (1999). Theoretical mechanisms underlying the trial-spacing effect in Pavlovian fear conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 177–193.

Barnet, R. C., Arnold, H. M., & Miller, R. R. (1991). Simultaneous conditioning demonstrated in second-order conditioning: Evidence for similar associative structure in forward and simultaneous conditioning. *Learning and Motivation, 22,* 253–268.

Barnet, R. C., Cole, R. P., & Miller, R. R. (1997). Temporal integration in second-order conditioning and sensory preconditioning. *Animal Learning and Behavior, 25,* 221–233.

Barnet, R. C., Grahame, N. J., & Miller, R. R. (1993a). Local context and the comparator hypothesis. *Animal Learning and Behavior, 21,* 1–13.

Barnet, R. C., Grahame, N. J., & Miller, R. R. (1993b). Temporal encoding as a determinant of blocking. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 327–341.

Barnet, R. C., & Miller, R. R. (1996). Second-order excitation mediated by a backward conditioned inhibitor. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 279–296.

Bateson, M., & Kacelnik, A. (1995). Preferences for fixed and variable food sources: Variability in amount and delay. *Journal of the Experimental Analysis of Behavior, 63,* 313–329.

Bateson, M., & Kacelnik, A. (1996). Rate currencies and the foraging starling: The fallacy of the averages revisited. *Behavioral Ecology, 7,* 341–352.

Baum, W. M. (1981). Optimization and the matching law as accounts of instrumental behavior. *Journal of the Experimental Analysis of Behavior, 36,* 387–403.

Baum, W. M. (1992). In search of the feedback function for variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 57,* 365–375.

Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior, 12,* 861–874.

Belke, T. W. (1992). Stimulus preference and the transitivity of preference. *Animal Learning and Behavior, 20,* 401–406.

Bevins, R. A., & Ayres, J. J. B. (1995). One-trial context fear conditioning as a function of the interstimulus interval. *Animal Learning and Behavior, 23,* 400–410.

Bouton, M. E. (1991). Context and retrieval in extinction and in other examples of interference in simple associative learning. In L. Dachowski & C. R. Flaherty (Eds.), *Current topics in animal learning* (pp. 25–53). Hillsdale, NJ: Erlbaum.

Bouton, M. E. (1993). Context, ambiguity, and classical conditioning. *Current Directions in Psychological Science, 3,* 49–53.

Bouton, M. E., & Ricker, S. T. (1994). Renewal of extinguished responding in a second context. *Animal Learning and Behavior, 22,* 317–324.

Brelsford, J., & Theios, J. (1965). Single-session conditioning of the nictitating membrane in the rabbit: Effect of intertrial interval. *Psychonomic Science, 2,* 81–82.

Brooks, D. C., Hale, B., Nelson, J. B., & Bouton, M. E. (1995). Reinstatement after counterconditioning. *Animal Learning and Behavior, 23,* 383–390.

Brown, P. L., & Jenkins, H. M. (1968). Autoshaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior, 11,* 1–8.

Brunner, D., Gallistel, C. R., Fairhurst, S., & Gibbon, J. (2000). *Pretraining determines choice strategy in concurrent choice.* Manuscript in preparation.

Brunner, D., & Gibbon, J. (1995). Value of food aggregates: Parallel versus serial discounting. *Animal Behaviour, 50,* 1627–1634.

Brunner, D., Gibbon, J., & Fairhurst, S. (1994). Choice between fixed and variable delays with different reward amounts. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 331–346.

Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 3–17.

Catania, A. C. (1963). Concurrent performances: A baseline for the study of reinforcement magnitude. *Journal of the Experimental Analysis of Behavior, 6*, 299–300.

Catania, A. C. (1970). Reinforcement schedules and psychophysical judgments. In W. N. Schoenfeld (Ed.), *Theory of reinforcement schedules* (pp. 1–42). New York: Appleton-Century-Crofts.

Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition, 37,* 55–81.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 216–228.

Church, R. M., & Gibbon, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes, 8,* 165–186.

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 135–155.

Church, R. M., Miller, K. D., Meck, W. H., & Gibbon, J. (1991). Symmetrical and asymmetrical sources of variance in temporal generalization. *Animal Learning and Behavior, 19,* 207–214.

Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain.* Cambridge, MA: MIT Press.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature, 395,* 272–274.

Clayton, N. S., & Dickinson, A. (1999). Memory for the content of caches by scrub jays (*Aphelocoma coerulescens*). *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 82–91.

Cole, R. P., Barnet, R. C., & Miller, R. R. (1995a). Effect of relative stimulus validity: Learning or performance deficit? *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 293–303.

Cole, R. P., Barnet, R. C., & Miller, R. R. (1995b). Temporal encoding in trace conditioning. *Animal Learning and Behavior, 23,* 144–153.

Colwill, R. M. (1993). Signaling the omission of a response-contingent outcome reduces discriminative control. *Animal Learning and Behavior, 21,* 337–345.

Colwill, R. M., Absher, R. A., & Roberts, M. L. (1988). Context–US learning in *Aplysia californica. Journal of Neuroscience, 8,* 4434–4439.

Colwill, R. M., & Delamater, B. A. (1995). An associative analysis of instrumental biconditional discrimination learning. *Animal Learning and Behavior, 23,* 218–233.

Colwill, R., & Rescorla, R. (1986). Associative structures in instrumental learning. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 20, pp. 55–104). New York: Academic Press.

Commons, M. L., Herrnstein, R. J., & Rachlin, H. (Eds.). (1982). *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts.* Cambridge, MA: Ballinger Press.

Cooper, L. D., Aronson, L., Balsam, P. D., & Gibbon, J. (1990). Duration of signals for intertrial reinforcement and nonreinforcement in random control procedures. *Journal of Experimental Psychology: Animal Behavior Processes, 16,* 14–26.

Davis, D. G., Staddon, J. E., Machado, A., & Palmer, R. G. (1993). The process of recurrent choice. *Psychological Review, 100,* 320–341.

Davison, M. C. (1969). Preference for mixed-interval versus fixed-interval schedules. *Journal of the Experimental Analysis of Behavior, 12,* 247–252.

Davison, M. C. (1988). Concurrent schedules: Interactions of reinforcer and frequency and reinforcer duration. *Journal of the Experimental Analysis of Behavior, 49,* 339–349.

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

Devenport, L. D., & Devenport, J. A. (1994). Time-dependent averaging of foraging information in least chipmunks and golden-mantled squirrels. *Animal Behaviour, 47,* 787–802.

Dews, P. B. (1970). The theory of fixed-interval responding. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 43–61). New York: Appleton-Century-Crofts.

Dickinson, A. (1989). Expectancy theory in animal conditioning. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 279–308). Hillsdale, NJ: Erlbaum.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 49B,* 60–80.

Dickinson, A., Hall, G., & Mackintosh, N. J. (1976). Surprise and the attenuation of blocking. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 213–222.

Durlach, P. J. (1983). Effect of signaling intertrial unconditioned stimuli in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 374–389.

Durlach, P. J. (1986). Explicitly unpaired procedure as a response elimination technique in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 12,* 172–185.

Durlach, P. J. (1989). Role of signals for unconditioned stimulus absence in the sensitivity of autoshaping to contingency. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 202–211.

Edmon, E. L., Lucki, I., & Grisham, M. G. (1980). Choice responding following multiple schedule training. *Animal Learning and Behavior, 8,* 287–292.

Fanselow, M. S. (1986). Associative vs topographical accounts of the immediate-shock freezing deficit in rats: Implications for the response selection rules governing species-specific defensive reactions. *Learning and Motivation, 17,* 16–39.

Fanselow, M. S. (1989). The adaptive function of conditioned defensive behavior: An ecological approach to Pavlovian stimulus-substitution theory. In R. J. Blanchard, P. F. Brain, D. C. Blanchard, & S. Parmigiani (Eds.), *Ethoexperimental approaches to the study of behavior* (MATO ASI Series D, Vol. 48, pp. 151–166). Boston: Kluwer Academic.

Fanselow, M. S. (1990). Factors governing one-trial contextual conditioning. *Animal Learning and Behavior, 18,* 264–270.

Fanselow, M. S., DeCola, J. P., & Young, S. L. (1993). Mechanisms responsible for reduced contextual conditioning with massed unsignaled unconditional stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 121–137.

Fanselow, M. S., & Stote, D. (1995, November). *Temporal vs. associative accounts of the immediate shock deficit in contextual fear.* Paper presented at the annual meeting of the Psychonomic Society, Los Angeles.

Fantino, E. (1969). Conditioned reinforcement, choice, and the psychological distance to reward. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 163–191). Homewood, IL: Dorsey Press.

Fantino, E., Preston, R. A., & Dunn, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior, 60,* 159–169.

Fiala, J. C., Grossberg, S., & Bullock, D. (1996). Metobotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. *Journal of Neuroscience, 16,* 3760–3774.

Foree, D. D., & LoLordo, V. M. (1973). Attention in the pigeon: Differential effects of food-getting versus shock-avoidance procedures. *Journal of Comparative and Physiological Psychology, 85,* 551–558.

Gallistel, C. R. (1989). Animal cognition: The representation of space, time and number. *Annual Review of Psychology, 40,* 155–189.

Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: Bradford Books/MIT Press.

Gallistel, C. R. (1992a). Classical conditioning as an adaptive specialization: A computational model. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 28, pp. 35–67). New York: Academic Press.

Gallistel, C. R. (1992b). Classical conditioning as a non-stationary, multivariate time series analysis: A spreadsheet model. *Behavior Research Methods, Instruments, and Computers, 24,* 340–351.

Gallistel, C. R. (1999). Can response timing be explained by a decay process? *Journal of the Experimental Analysis of Behavior, 71,* 264–271.

Gamzu, E., & Williams, D. R. (1971). Classical conditioning of a complex skeletal response. *Science, 171,* 923–925.

Gamzu, E., & Williams, D. R. (1973). Associative factors underlying the pigeon's keypecking in autoshaping procedures. *Journal of the Experimental Analysis of Behavior, 19,* 225–232.

Ganesan, R., & Pearce, J. M. (1988). Effect of changing the unconditioned stimulus on appetitive blocking. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 280–291.

Gibbon, J. (1971). Scalar timing and semi-Markov chains in free-operant avoidance. *Journal of Mathematical Psychology, 8,* 109–138.

Gibbon, J. (1972). Timing and discrimination of shock density in avoidance. *Psychological Review, 79,* 68–92.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84,* 279–325.

Gibbon, J. (1981a). The contingency problem in autoshaping. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 285–308). New York: Academic Press.

Gibbon, J. (1981b). On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology, 24,* 58–87.

Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical Psychology, 36,* 283–293.

Gibbon, J. (1995). Dynamics of time matching: Arousal makes better seem worse. *Psychonomic Bulletin and Review, 2,* 208–215.

Gibbon, J., & Allan, L. (Eds.). (1984). *Timing and time perception* (Vol. 423). New York: New York Academy of Sciences.

Gibbon, J., Baldock, M. D., Locurto, C. M., Gold, L., & Terrace, H. S. (1977). Trial and intertrial durations in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 264–284.

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

Gibbon, J., & Church, R. M. (1990). Representation of time. *Cognition, 37,* 23–54.

Gibbon, J., & Church, R. M. (1992). Comparison of variance and covariance patterns in parallel and serial theories of timing. *Journal of the Experimental Analysis of Behavior, 57,* 393–406.

Gibbon, J., Church, R. M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review, 95,* 102–114.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Vol. 423, pp. 52–77). New York: New York Academy of Sciences.

Gibbon, J., & Fairhurst, S. (1994). Ratio versus difference comparators in choice. *Journal of the Experimental Analysis of Behavior, 62,* 409–434.

Gibbon, J., Fairhurst, S., & Goldberg, B. (1997). Cooperation, conflict and compromise between circadian and interval clocks in pigeons. In C. M. Bradshaw & E. Szabadi (Eds.), *Time and behaviour: Psychological and neurobehavioural analyses* (pp. 329–384). Amsterdam: North-Holland/Elsevier.

Gibbon, J., Farrell, L., Locurto, C. M., Duncan, H. J., & Terrace, H. S. (1980). Partial reinforcement in autoshaping with pigeons. *Animal Learning and Behavior, 8,* 45–59.

Gibbon, J., Locurto, C., & Terrace, H. S. (1975). Signal–food contingency and signal frequency in a continuous trials auto-shaping paradigm. *Animal Learning and Behavior, 3,* 317–324.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7,* 170–184.

Gluck, M. A., & Thompson, R. F. (1987). Modeling the neural substrates of associative learning and memory: A computational approach. *Psychological Review, 94,* 176–191.

Goddard, M. J., & Jenkins, H. M. (1987). Effect of signaling extra unconditioned stimuli on autoshaping. *Animal Learning and Behavior, 15,* 40–46.

Gormezano, I., & Coleman, S. R. (1975). Effects of partial reinforcement on conditioning, conditional probabilities, asymptotic performance, and extinction of the rabbit's nictitating membrane response. *Pavlovian Journal of Biological Science, 10,* 13–22.

Grahame, N. J., Barnet, R. C., & Miller, R. R. (1992). Pavlovian inhibition cannot be obtained by posttraining A–US pairings: Further evidence for the empirical asymmetry of the comparator hypothesis. *Bulletin of the Psychonomic Society, 30,* 399–402.

Granger, R. H. J., & Schlimmer, J. C. (1986). The computation of contingency in classical conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 20, pp. 137–192). New York: Academic Press.

Grossberg, S., & Merrill, J. W. L. (1996). The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience, 8,* 257–277.

Grossberg, S., & Schmajuk, N. A. (1991). Neural dynamics of adaptive timing and temporal discrimination during associative learning. In G. A. Carpenter & S. Grossberg (Eds.), *Pattern recognition by self-organizing neural networks* (pp. 637–674). Cambridge, MA: MIT Press.

Hallam, S. C., Matzel, L. D., Sloat, J. S., & Miller, R. R. (1990). Excitation and inhibition as a function of post-training extinction of the excitatory cue used in Pavlovian inhibition training. *Learning and Motivation, 21,* 59–84.

Harper, D. G. C. (1982). Competitive foraging in mallards: Ideal free ducks. *Animal Behaviour, 30,* 575–584.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4,* 267–272.

Herrnstein, R. J. (1964). Secondary reinforcement and rate of primary reinforcement. *Journal of the Experimental Analysis of Behavior, 7,* 27–36.

Herrnstein, R. J. (1991). Experiments on stable sub-optimality in individual behavior. *American Economic Review, 81,* 360–364.

Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior, 24,* 107–116.

Herrnstein, R. J., & Prelec, D. (1991). Melioration: A theory of distributed choice. *Journal of Economic Perspectives, 5,* 137–156.

Herrnstein, R. J., & Vaughan, W. J. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143–176). New York: Academic Press.

Heyman, G. M. (1979). A Markov model description of changeover probabilities on concurrent variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 31,* 41–51.

Heyman, G. M. (1982). Is time allocation unconditioned behavior? In M. Commons, R. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts* (pp. 459–490). Cambridge, MA: Ballinger Press.

Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition, 37,* 105–131.

Holmes, J. D., & Gormezano, I. (1970). Classical appetitive conditioning of the rabbit's jaw movement response under partial and continuous reinforcement schedules. *Learning and Motivation, 1,* 110–120.

Honig, W. K. (1981). Working memory and the temporal map. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 167–197). Hillsdale, NJ: Erlbaum.

Hull, C. L. (1930). Knowledge and purpose as habit mechanisms. *Psychological Review, 37,* 511–525.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Hyman, A. (1969). Two temporal parameters of free-operant discriminated avoidance in the rhesus monkey. *Journal of the Experimental Analysis of Behavior, 12*, 641–648.

James, J. H., & Wagner, A. R. (1980). One-trial overshadowing: Evidence of distributive processing. *Journal of Experimental Psychology: Animal Behavior Processes, 6*, 188–205.

Kamin, L. J. (1954). Traumatic avoidance learning: The effects of CS–US interval with a trace-conditioning procedure. *Journal of Comparative and Physiological Psychology, 47*, 65–72.

Kamin, L. J. (1967). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior: Aversive stimulation* (pp. 9–33). Miami, FL: University of Miami Press.

Kamin, L. J. (1969a). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276–296). New York: Appleton-Century-Crofts.

Kamin, L. J. (1969b). Selective association and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42–64). Halifax, Nova Scotia, Canada: Dalhousie University Press.

Kamin, L. J., & Gaioni, S. J. (1974). Compound conditioned emotional response conditioning with differentially salient elements in rats. *Journal of Comparative and Physiological Psychology, 87*, 591–597.

Kaplan, P., & Hearst, E. (1985). Contextual control and excitatory vs inhibitory learning: Studies of extinction, reinstatement, and interference. In P. D. Balsam & A. Tomie (Eds.), *Context and learning* (pp. 195–224). Hillsdale, NJ: Erlbaum.

Kaufman, M. A., & Bolles, R. C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society, 18*, 318–320.

Kehoe, E. J., & Gormezano, I. (1974). Effects of trials per session on conditioning of the rabbit's nictitating membrane response. *Bulletin of the Psychonomic Society, 4*, 434–436.

Kehoe, E. J., Graham-Clarke, P., & Schreurs, B. G. (1989). Temporal patterns of the rabbit's nictitating membrane response to compound and component stimuli under mixed CS–US intervals. *Behavioral Neuroscience, 103*, 283–295.

Keller, J. V., & Gollub, L. R. (1977). Duration and rate of reinforcement as determinants of concurrent responding. *Journal of the Experimental Analysis of Behavior, 28*, 145–153.

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

Killeen, P. (1975). On the temporal control of behavior. *Psychological Review, 82*, 89–115.

Killeen, P. R. (1982). Incentive theory: II. Models for choice. *Journal of the Experimental Analysis of Behavior, 38*, 217–232.

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review, 94*, 455–468.

Killeen, P. R., & Taylor, T. (in press). How the propagation of error through stochastic counters affects time discrimination and other psychophysical judgments. *Psychological Review.*

Killeen, P. R., & Weiss, N. A. (1987). Optimal timing and the Weber function. *Psychological Review, 94*, 455–468.

Kremer, E. F. (1978). The Rescorla–Wagner model: Losses in associative strength in compound conditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 4*, 22–36.

LaBarbera, J. D., & Church, R. M. (1974). Magnitude of fear as a function of the expected time to an aversive event. *Animal Learning and Behavior, 2*, 199–202.

Lattal, K. M., & Nakajima, S. (1998). Overexpectation in appetitive Pavlovian and instrumental conditioning. *Animal Learning and Behavior, 26*, 351–360.

Leon, M. I., & Gallistel, C. R. (1998). Self-stimulating rats combine subjective reward magnitude and subjective reward rate multiplicatively. *Journal of Experimental Psychology: Animal Behavior Processes, 24*, 265–277.

Levinthal, C. F., Tartell, R. H., Margolin, C. M., & Fishman, H. (1985). The CS–US interval (ISI) function in rabbit nictitating membrane response conditioning with very long intertrial intervals. *Animal Learning and Behavior, 13*, 228–232.

Libby, M. E., & Church, R. M. (1974). Timing of avoidance responses by rats. *Journal of the Experimental Analysis of Behavior, 22*, 513–517.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492–527.

Logue, A. W., & Chavarro, A. (1987). Effect on choice of absolute and relative values of reinforcer delay, amount, and frequency. *Journal of Experimental Psychology: Animal Behavior Processes, 13*, 280–291.

LoLordo, V. M., & Fairless, J. L. (1985). Pavlovian conditioned inhibition: The literature since 1969. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals* (pp. 143–170). Hillsdale, NJ: Erlbaum.

LoLordo, V. M., Jacobs, W. J., & Foree, D. D. (1982). Failure to block control by a relevant stimulus. *Animal Learning and Behavior, 10*, 183–193.

Low, L. A., & Low, H. I. (1962). Effects of CS–US interval length upon avoidance responding. *Journal of Comparative and Physiological Psychology, 55*, 1059–1061.

Mackintosh, N. J. (1971). An analysis of overshadowing and blocking. *Quarterly Journal of Experimental Psychology, 23*, 118–125.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276–298.

Mackintosh, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning and Behavior, 4*, 186–192.

Mackintosh, N. J., & Dickinson, A. (1979). Instrumental (Type II) conditioning. In A. Dickinson & R. A. Boakes (Eds.), *Mechanism of learning and motivation.* Hillsdale, NJ: Erlbaum.

Mackintosh, N. J., & Reese, B. (1970). One-trial overshadowing. *Quarterly Journal of Experimental Psychology, 31*, 519–526.

Mark, T. A. (1997). *The microstructure of choice.* Unpublished doctoral dissertation, University of California, Los Angeles.

Mark, T. A., & Gallistel, C. R. (1993). Subjective reward magnitude of MFB stimulation as a function of train duration and pulse frequency. *Behavioral Neuroscience, 107*, 389–401.

Mark, T. A., & Gallistel, C. R. (1994). The kinetics of matching. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 79–95.

Matzel, L. D., Held, F. P., & Miller, R. R. (1988). Information and expression of simultaneous and backward associations: Implications for contiguity theory. *Learning and Motivation, 19*, 317–344.

Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation, 16*, 398–412.

Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable delays. *Journal of Experimental Psychology: Animal Behavior Processes, 10*, 426–436.

Mazur, J. E. (1986). Fixed and variable ratios and delays: Further test of an equivalence rule. *Journal of Experimental Psychology: Animal Behavior Processes, 12*, 116–124.

Mazur, J. E. (1989). Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior, 51*, 87–99.

Mazur, J. E. (1991). Choice with probabilistic reinforcement: Effects of delay and conditioned reinforcers. *Journal of the Experimental Analysis of Behavior, 55*, 63–77.

Mazur, J. E. (1995). Development of preference and spontaneous recovery in choice behavior with concurrent variable-interval schedules. *Animal Learning and Behavior, 23*, 93–103.

Mazur, J. E. (1996). Past experience, recency, and spontaneous recovery in choice behavior. *Animal Learning and Behavior, 24*, 1–10.

Mazur, J. E. (1997). Choice, delay, probability, and conditioned reinforcement. *Animal Learning and Behavior, 25*, 131–147.

Mazur, J. E., & Wagner, A. R. (1982). An episodic model of associative learning. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Acquisition* (pp. 3–39). Cambridge, MA: Ballinger Press.

McDiarmid, C. G., & Rilling, M. E. (1965). Reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychonomic Science, 2*, 195–196.

Meck, W. H., Church, R. M., & Gibbon, J. (1985). Temporal integration in duration and number discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 591–597.

Miall, C. (1996). Models of neural timing. In M. A. Pastor & J. Artieda (Eds.), *Advances in psychology: Vol. 115. Time, internal clocks and movement* (pp. 69–94). Amsterdam: North-Holland/Elsevier Science.

Miller, R. R., & Barnet, R. C. (1993). The role of time in elementary associations. *Current Directions in Psychological Science, 2*, 106–111.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1992). Responding to a conditioned stimulus depends on the current associative status of other cues that were present during training of that specific stimulus. *Journal of Experimental Psychology: Animal Behavior Processes, 18*, 251–264.

Miller, R. R., & Grahame, N. J. (1991). Expression of learning. In L. Dachowski & C. F. Flaherty (Eds.), *Current topics in animal learning* (pp. 95–117). Hillsdale, NJ: Erlbaum.

Miller, R. R., & Matzel, L. D. (1989). Contingency and relative associative strength. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 61–84). Hillsdale, NJ: Erlbaum.

Miller, R. R., & Schachtman, T. R. (1985). Conditioning context as an associative baseline: Implications for response generation and the nature of conditioned inhibition. In R. R. Miller & N. S. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 51–88). Hillsdale, NJ: Erlbaum.

Mustaca, A. E., Gabelli, F., Papine, M. R., & Balsam, P. (1991). The effects of varying the interreinforcement interval on appetitive contextual conditioning. *Animal Learning and Behavior, 19*, 125–138.

Myerson, J., & Miezin, F. M. (1980). The kinetics of choice: An operant systems analysis. *Psychological Review, 87*, 160–174.

Nelson, J. B., & Bouton, M. E. (1997). The effects of a context switch following serial and simultaneous feature-negative discriminations. *Learning and Motivation, 28*, 56–84.

Nevin, J. A. (1982). Some persistent issues in the study of matching and maximizing. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts* (pp. 153–165). Cambridge, MA: Ballinger Press.

Pavlov, I. V. (1928). *Lectures on conditioned reflexes: The higher nervous activity of animals* (H. Gantt, Trans.). London: Lawrence & Wishart.

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review, 101*, 587–607.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variation in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87*, 532–552.

Penney, T. B., Allan, L. G., Meck, W. H., & Gibbon, J. (1998). Memory mixing in duration bisection. In D. A. Rosenbaum & C. E. Collyer (Eds.), *Timing of behavior: Neural, psychological, and computational perspectives* (pp. 165–193). Cambridge, MA: MIT Press.

Pliskoff, S. S. (1971). Effects of symmetrical and asymmetrical changeover delays on concurrent performances. *Journal of the Experimental Analysis of Behavior, 16*, 249–256.

Prokasy, W. F., & Gormezano, I. (1979). The effect of US omission in classical aversive and appetitive conditioning of rabbits. *Animal Learning and Behavior, 7*, 80–88.

Pubols, B. H. (1962). Constant versus variable delay of reinforcement. *Journal of Comparative and Physiological Psychology, 55*, 52–56.

Rachlin, H. (1974). Self-control. *Behaviorism, 2*, 94–107.

Rachlin, H. (1978). A molar theory of reinforcement schedules. *Journal of the Experimental Analysis of Behavior, 30*, 345–360.

Rachlin, H. (1995). Self-control: Beyond commitment. *Behavioral and Brain Sciences, 18*, 109–121.

Rachlin, H., Logue, A. W., Gibbon, J., & Frankel, M. (1986). Cognition and behavior in studies of choice. *Psychological Review, 93*, 33–45.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior, 55*, 233–244.

Rashotte, M. E., Griffin, R. W., & Sisk, C. L. (1977). Second-order conditioning of pigeon's keypeck. *Animal Learning and Behavior, 5*, 25–38.

Rescorla, R. A. (1967). Inhibition of delay in Pavlovian fear conditioning. *Journal of Comparative and Physiological Psychology, 64*, 114–120.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*, 1–5.

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin, 72*, 77–94.

Rescorla, R. A. (1972). Informational variables in Pavlovian conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 1–46). New York: Academic Press.

Rescorla, R. A. (1988). Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience, 11*, 329–352.

Rescorla, R. A. (1991). Associations of multiple outcomes with an instrumental response. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 465–474.

Rescorla, R. A. (1992). Response-independent outcome presentation can leave instrumental R–O associations intact. *Animal Learning and Behavior, 20*, 104–111.

Rescorla, R. A. (1993). The preservation of response–outcome associations through extinction. *Animal Learning and Behavior, 21*, 238–245.

Rescorla, R. A. (1996). Spontaneous recovery after training with multiple outcomes. *Animal Learning and Behavior, 24*, 11–18.

Rescorla, R. A. (1998). Instrumental learning: Nature and persistence. In M. Sabourin, F. I. M. Craik, & M. Roberts (Eds.), *Proceeding of the XXVI International Congress of Psychology: Vol. 2. Advances in psychological science: Biological and cognitive aspects* (pp. 239–258). London: Psychology Press.

Rescorla, R. A., & Colwill, R. M. (1989). Associations with anticipated and obtained outcomes in instrumental conditioning. *Animal Learning and Behavior, 17*, 291–303.

Rescorla, R. A., & Cunningham, C. L. (1978). Within-compound flavor association. *Journal of Experimental Psychology: Animal Behavior Processes, 4*, 267–275.

Rescorla, R. A., & Durlach, P. J. (1987). The role of context in intertrial interval effects in autoshaping. *Quarterly Journal of Experimental Psychology, 39B*, 35–48.

Rescorla, R. A., & Heth, C. D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes, 104*, 88–106.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York: Appleton-Century-Crofts.

Reynolds, G. S. (1961). Attention in the pigeon. *Journal of the Experimental Analysis of Behavior, 4*, 203–208.

Rilling, M., & McDiarmid, C. (1965). Signal detection in fixed ratio schedules. *Science, 148*, 526–527.

Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes, 7*, 242–268.

Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes, 44*, 147–162.

Schneider, B. (1969). A two-state analysis of fixed-interval responding in the pigeon. *Journal of the Experimental Analysis of Behavior, 12*, 667–687.

Schneiderman, N., & Gormezano, I. (1964). Conditioning of the nictitating membrane of the rabbit as a function of CS–US interval. *Journal of Comparative and Physiological Psychology, 57,* 188–195.

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

Sheafor, P. J., & Gormezano, I. (1972). Conditioning the rabbit's (*Oryctolagus cuniculus*) jaw-movement response. *Journal of Comparative and Physiological Psychology, 81,* 449–456.

Sherry, D. (1984). Food storage by black-capped chickadees: Memory for location and contents of caches. *Animal Behavior, 32,* 451–464.

Shimp, C. P. (1969). Optimal behavior in free-operant experiments. *Psychological Review, 76,* 97–112.

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

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Foundations* (Vol. 1, pp. 194–281). Cambridge, MA: MIT Press.

Staddon, J. E. R., & Higa, J. J. (1993). Temporal learning. In D. Medin (Ed.), *The psychology of learning and motivation* (Vol. 27, pp. 265–294). New York: Academic Press.

Staddon, J. E. R., & Motheral, S. (1979). On matching and maximizing in operant choice experiments. *Psychological Review, 85,* 436–444.

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 497–537). Cambridge, MA: Bradford Books/MIT Press.

Terrace, H. S., Gibbon, J., Farrell, L., & Baldock, M. D. (1975). Temporal factors influencing the acquisition and maintenance of an autoshaped keypeck. *Animal Learning and Behavior, 3,* 53–62.

Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Instrumental conditioning theory and the impact of biological constraints on learning* (pp. 237–275). Hillsdale, NJ: Erlbaum.

Tomie, A. (1976a). Interference with autoshaping by prior context conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 323–334.

Tomie, A. (1976b). Retardation of autoshaping: Control by contextual stimuli. *Science, 192,* 1244–1246.

Usherwood, P. N. R. (1993). Memories are made of this. *Trends in Neurosciences, 16,* 427–429.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation, 25,* 127–151.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Erlbaum.

Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology, 76,* 171–180.

Wasserman, E. A., Franklin, S. R., & Hearst, E. (1974). Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *Journal of Comparative and Physiological Psychology, 86,* 616–627.

Wasserman, E. A., & McCracken, S. B. (1974). The disruption of autoshaped keypecking in the pigeon by food-tray illumination. *Journal of the Experimental Analysis of Behavior, 22,* 39–45.

Wasserman, E. A., & Miller, R. R. (1997). What's elementary about associative learning? *Annual Review of Psychology, 48,* 573–607.

Weidemann, G., Georgilas, A., & Kehoe, E. J. (1999). Temporal specificity in patterning of the rabbit nictitating membrane response. *Animal Learning and Behavior, 27,* 99–107.

Williams, B. A. (1981). Invariance in reinforcements to acquisition, with implications for the theory of inhibition. *Behaviour Analysis Letters, 1,* 73–80.

Williams, B. A. (1982). Blocking the response–reinforcer association. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Acquisition* (pp. 427–447). Cambridge, MA: Ballinger Press.

Williams, B. A., & Royalty, P. (1989). A test of the melioration theory of matching. *Journal of Experimental Psychology: Animal Behavior Processes, 15,* 99–113.

Appendix A

Acquisition and Extinction

Acquisition

The acquisition decision is to respond whenever

$$(\lambda_{CS} + \lambda_b)/\lambda_b \geq \beta, \tag{A1}$$

where β is the threshold. For delay conditioning (see Figure 11), these rates are given by $\lambda_b = n_I/t_I$ and $\lambda_{CS} = (n/t_{CS}) - \lambda_b$, so that the decision rule (Equation A1) is equivalent to

$$(n/t_{CS})/(n_I/t_I) \geq \beta$$

or

$$n/n_I \geq \beta(t_{CS}/t_I), \tag{A2a}$$

with expectation (for equality)

$$N/N_I = \beta(NT/NI) = \beta(I/T)^{-1}, \tag{A2b}$$

which confirms that the ratio of free to CS reinforcers at acquisition must be constant for constant I/T ratios.

Assuming $n_I = 1$, then on average $N = \beta(I/T)^{-1}$. However, the right side of Equation A2a is a ratio of two normal variates, a Geary z variate (see Gibbon, 1977), which does not have a second moment. We therefore derive the semi-interquartile range for n (SIQR $= N_{.75} - N_{.5}$ for the symmetric Geary z) to show that variability around the regression in Figure 6 is also proportional to $\beta(I/T)^{-1}$, and hence constant on the log scale.

Let $x = \beta t_{CS}$ and $y = nt_I$, with means

$$\mu_x = \beta NT,\ \mu_y = N^2 I \tag{A3a}$$

and variances

$$\sigma_x^2 = (\gamma\beta NT)^2,\ \sigma_y^2 = (\gamma N^2 I)^2, \tag{A3b}$$

reflecting the scalar property with sensitivity index γ (the coefficient of variation).

The criterion (Equation A2) is met with probability α or better iff $\exists(N_\alpha)$ such that

$$P(x - y < 0) = \alpha = \Phi(z_\alpha).$$

But assuming normal form for the scalar random variables x and y,

$$x - y < 0 \Leftrightarrow z = (x - y - \mu_{x-y})/\sigma_{x-y} < (-\mu_{x-y}/\sigma_{x-y}) = z_\alpha. \tag{A4}$$

We then can solve Equation A4 for N_α in terms of z_α. Given independence of x, y,

$$\sigma_{x-y}^2 = \sigma_x^2 + \sigma_y^2 = (N\gamma)^2[(\beta T)^2 + (NI)^2],$$

so that

$$(\gamma z_\alpha)^2 = [(N_\alpha I/\beta T) - 1]^2/[(N_\alpha I/\beta T)^2 + 1]. \tag{A5a}$$

For the median, $z_{.5} = 0$, then

$$N_{.5} = N = \beta(I/T)^{-1}, \tag{A5b}$$

confirming the symmetry of the Geary z variate.

Letting $u = N_\alpha I/\beta T$ and setting $K_\alpha = (\gamma z_\alpha)^2$, Equation A5a may be written as

$$u^2 + 2u/(K - 1) + 1 = 0. \tag{A6}$$

For the upper α percentile, the positive root of this equation, given $|K_\alpha - 1| < 1$ (a common range for γ), is

$$u_\alpha = 1/(1 - K_\alpha) + [1/(K_\alpha - 1)^2 - 1]^{1/2}. \tag{A7}$$

Recalling that $u = N_\alpha I/\beta T$ and using the solution (Equation A7) for u, the upper α percentile (N_α) is then

$$N_\alpha = \beta u_\alpha (I/T)^{-1}. \tag{A8}$$

Because $u_{.5} = 1$, we have

$$\text{SIQR} = \beta(I/T)^{-1}[u_{.75} - 1],$$

proportional to $(I/T)^{-1}$.

Extinction

The extinction decision is very similar to that for acquisition, with the new (extinction) CS rate, $\lambda_{CS\ no\ R}$, playing the role of the background. Now, however, the question is not whether this CS rate is different from 0 but whether it is different from the previous CS rate, λ_{CS}. Again, as with all decisions, the criteria are taken in ratio and compared with a threshold. Subjects stop responding to the CS when

$$\lambda_{CS}/\lambda_{CS\ no\ R} \geq \beta, \tag{A9}$$

where λ_{CS} is the previous remembered reinforced rate,

$$\lambda_{CS} = n/t_{CS} - \lambda_b,$$

and $\lambda_{CS\ no\ R}$ is the new, unreinforced CS rate,

$$\lambda_{CS\ no\ R} = 1/t_{CS\ no\ R} - \lambda_b,$$

where $t_{CS\ no\ R}$ is the current accumulated CS time without reinforcement, with expectation, $N_{no\ R}/T$, and scalar variance, $(\gamma N_{no\ R}/T)^2$.

Because acquisition has occurred by the time extinction is instituted, λ_b is small, and therefore the decision rule (Equation A9) is approximately

$$nt_{CS\ no\ R}/t_{CS} \geq \beta. \tag{A10a}$$

It is convenient to write the left side of Equation A10a as

$$n_{no\ R}(t_{CS\ no\ R}/n_{no\ R})/(t_{CS}/n), \tag{A10b}$$

where it is clear the variates in parentheses are both estimates of T. Hence, taking expectations, subjects stop responding on average when

$$N_{no\ R} \geq \beta. \tag{A11}$$

Thus, the extinction decision is independent of I, T, and their ratio.

An explicit form for the SIQR of $N_{no\ R}$ is obtained similarly to acquisition. Let $x = \beta t_{CS}$ and $y = nt_{CS\ no\ R}$, with means

$$\mu_x = \beta NT,\ \mu_y = NN_{no\ R}T \tag{A12a}$$

and variances

$$\sigma_x^2 = (\gamma\beta NT)^2,\ \sigma_y^2 = (\gamma NN_{no\ R}T)^2, \tag{A12b}$$

reflecting the scalar property. The criterion (Equation A9) is met with probability α or better iff $\exists(N^*_{no\ R})$, such that

$$P(x - y < 0) = \alpha = \Phi(z_\alpha).$$

Again assuming normal form for the scalar random variables x and y,

$$x - y < 0 \Leftrightarrow z = (x - y - \mu_{x-y})/\sigma_{x-y} < -\mu_{x-y}/\sigma_{x-y} = z_\alpha.$$

We can solve Equations A12a and A12b for $N^*_{\text{no R}}$ in terms of z_α.

$$z_\alpha = (NN^*_{\text{no R}}T - \beta NT)/[(\gamma\beta NT)^2 + (\gamma NN^{*2}_{\text{no R}}T)^2]^{1/2},$$

so that terms in *NT* cancel, giving

$$(\gamma z_\alpha)^2 = [(N^*_{\text{no R}} - \beta)]^2/(\beta^2 + N^{*2}_{\text{no R}})]. \qquad \text{(A13)}$$

Again setting $u = N^*_{\text{no R}}/\beta$ and $K_\alpha = (\gamma z_\alpha)^2$, the positive upper α solution to the quadratic Equation A13 gives

$$N^*_{\text{no R}} = 2\beta u_\alpha, \qquad \text{(A14)}$$

where $u_\alpha = 1/(1 - K_\alpha) + [1/(K_\alpha - 1)^2 - 1]^{1/2}$, as in Equation A7. Thus,

$$\text{SIQR} = 2\beta(u_{.75} - 1). \qquad \text{(A15)}$$

Variability is also constant, independent of the conditioning protocol.

Appendix B

Rate Estimation

The general solution to the problem of computing true rates of reinforcement $\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}$ for *m* predictive stimuli (CSs) under the assumption that rates of reinforcement combine additively is

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{vmatrix} 1 & \dfrac{t_{1,2}}{t_1} & \cdots & \dfrac{t_{1,m}}{t_1} \\ \dfrac{t_{1,2}}{t_2} & 1 & \cdots & \dfrac{t_{2,m}}{t_2} \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{t_{1,m}}{t_m} & \dfrac{t_{2,m}}{t_m} & \cdots & 1 \end{vmatrix}^{-1} \times \begin{pmatrix} \dfrac{N_1}{t_1} \\ \dfrac{N_2}{t_2} \\ \vdots \\ \dfrac{N_m}{t_m} \end{pmatrix},$$

where t_i is the cumulative duration of stimulus *i*, $t_{i,j}$ is the cumulative duration of the conjunctions of stimuli *i* and *j*, and N_i is the cumulative number of reinforcements delivered in the presence of stimulus *i*. When there are redundant CSs, the determinant of the matrix is zero, so a unique solution does not exist. Unique solutions are then obtained from reduced systems, which are generated from matrices of reduced rank and vectors of correspondingly reduced dimensionality. These reduced matrices and vectors are obtained by deleting one or more CSs from consideration. Deleting different CSs produces different solutions. Among these, the preferred solution is the one that minimizes the sum of the absolute values of the rate estimates. When there is more than one such minimal solution, the choice among them must either be made at random or be based on extraneous considerations, such as the relative salience of the CSs.

Received October 1, 1997
Revision received May 21, 1999
Accepted June 8, 1999 ■