

Learning the Temporal Dynamics of Behavior

Armando Machado
Indiana University Bloomington

This study presents a dynamic model of how animals learn to regulate their behavior under time-based reinforcement schedules. The model assumes a serial activation of behavioral states during the interreinforcement interval, an associative process linking the states with the operant response, and a rule mapping the activation of the states and their associative strength onto response rate or probability. The model fits data sets from fixed-interval schedules, the peak procedure, mixed fixed-interval schedules, and the bisection of temporal intervals. The major difficulties of the model came from experiments that suggest that under some conditions animals may time 2 intervals independently and simultaneously.

Time is not a cause, but a vehicle of causes.

—(Attributed to Jean Piaget)

Many vertebrates show some form of behavioral sensitivity to time, certainly one of the most important variables in the study of learning. Aware of this fact, psychologists have conducted substantial research to understand how time controls behavior (e.g., Ferster & Skinner, 1957; Gallistel, 1990; Gibbon & Allan, 1984; Platt, 1979; Richelle & Lejeune, 1980; Staddon, 1983). Their efforts have culminated in two elegant theories of timing, the scalar expectancy theory (SET) developed by Gibbon and his colleagues (Church, Meck, & Gibbon, 1994; Gibbon, 1977, 1991), and the behavioral theory of timing (BeT) developed by Killeen and Fetterman (1988; also Killeen, 1991).

SET is a cognitive account that postulates various information-processing devices to explain the temporal regularities of learned behavior. A pacemaker generates pulses with a high but variable frequency. An accumulator accumulates the pulses until an important event, such as food for a hungry animal, occurs. At that moment, the number of accumulated pulses is stored in long-term memory. When the animal starts to time an interval, it extracts a sample from its distribution of values in long-term memory and then, throughout the interval, compares the sampled value with the current number in the accumulator. The relative discrepancy between these two numbers controls, through one or more thresholds, the animal's behavior.

The work reported here was initiated when I was a postdoctoral researcher in John Staddon's laboratory at Duke University and was funded by grants from the National Science Foundation and the National Institute of Mental Health.

Parts of this study were presented at the annual meeting of the Society for the Quantitative Analysis of Behavior, Washington, DC, May 1995.

I thank Francisco Silva, Michael Zeiler, and Eliot Hearst for their comments on earlier versions of this article; Peter Killeen for his encouragement and help with quantitative details; and John Staddon for many years of continual inspiration.

Correspondence concerning this article should be addressed to Armando Machado, Department of Psychology, Indiana University, Bloomington, Indiana 47405. Electronic mail may be sent via Internet to amachado@indiana.edu.

BeT is a more behavioral account that conceives of the animal as moving across an invariant series of behavioral classes during interreinforcement intervals. A pulse from an internal pacemaker changes the behavioral class in which the animal is engaged. When a reinforcer occurs for, say, pecking a key, the animal associates the current behavioral class with pecking. According to BeT, the behavioral classes serve as discriminative stimuli, cues that set the occasion for the operant response.

Although BeT's emphasis on the discriminative function of behavioral classes and SET's reliance on adjustable thresholds imply a learning process, neither account deals explicitly with learning. In other words, SET and BeT describe the behavior of animals after prolonged exposure to time-based reinforcement schedules, but remain silent on how animals attain the steady state and on how animals regulate their behavior in real time. The quest for a dynamic, learning model of temporal control is the major goal of this study.

Ultimately, a general theory of the temporal control of behavior will have to address the issues of learning and real-time dynamics. First, some deep problems of temporal control surface only when the dynamic, real-time properties of the problem are specified clearly. A case in point is the credit-assignment problem: How does time, as opposed to some stimulus or response dimension, come to control the behavior of the animal? To date, this question has remained not only unanswered, but essentially unasked. Second, to deal with specific instances, static theories are forced to make ad hoc assumptions about the learning process, but these assumptions are not always mutually compatible. For example, to derive predictions for a variety of experimental procedures, Killeen and Fetterman (1988) had to make assumptions concerning which behavioral classes are conditioned to which responses. Sometimes only one class was assumed to be conditioned to a specific response; at other times it was necessary to assume that multiple classes are conditioned to a single response. No principled argument justified the two different assumptions. Third, dynamic models invite researchers to integrate what are currently disconnected pieces of knowledge. Thus, by asking how the processes of reinforcement and extinction affect temporal regulation, how temporal generalization and discrimination map onto absolute response rate, and how choice rules apply in time-based situations, dynamic models bring together pieces of the same puzzle, pieces that hitherto

have remained apart. In the same vein, Ferster and Skinner (1957) identified a set of variables that are likely to control behavior in time-based schedules: response rate at the moment of reinforcement, total number of responses emitted during the interreinforcement interval, time since the last reinforcer, and the discriminative stimulus function of the reinforcer itself. Dynamic models are one avenue to test ones' guesses of how these variables relate to each other and determine behavior.

The development of a learning model of temporal control faces three major challenges. The first is to describe how temporal regulation develops in real time. Thus, if we assume that reinforcement is a critical component of temporal regulation, then the model must describe in real time how reinforcement plays its role. The second challenge is to specify how the process of temporal regulation maps onto observable behavior. Third, the model must satisfy the empirical constraints identified by the static accounts. That is, the predictions of the dynamic model must converge to the correct steady-state description of behavior.

The present study is a first attempt to develop a dynamic model of how animals learn to regulate their behavior in time-based schedules. Despite the contemporary fashion for probabilistic approaches (e.g., SET and BeT), the present model is *deterministic* and stresses the average behavior of the animal. At present, I am willing to trade off the reduced depth of the model (e.g., its inability to account for trial-by-trial variability in behavior) for its greater generality and ones' greater understanding of its properties. In fact, all of the model's equations have closed-form solutions that greatly enhance ones' understanding of its predictions. It is a *behavioral*, not a cognitive, connectionist, or neurophysiological model, such as those of Church and Broadbent (1992) or Grossberg and Schmajuk (1989), because it emphasizes the role of the behavior of the animal and its interactions with the environment in the development of temporal control. It is *dynamic*, not static, because it stresses the equations of motion in time-based schedules; and, for the most part, it deals with processes occurring in real time, not in some other epoch.

This article is divided into three parts. The first presents the basic elements of the model, their interconnection, and their empirical and theoretical justification. The second compares the model's predictions with the empirical results obtained in a variety of procedures. Even though most data sets come from steady-state analyses, reflecting a dominant trend in the field, I also illustrate how the model illuminates some dynamic aspects of the behavior observed in these procedures. The last part summarizes the major strengths and weaknesses of the model, discusses the behavior of its parameters, and contrasts it with related approaches.

Basic Model: Assumptions and Equations

Figure 1 shows the basic structure of the model. It consists of three major components, a serial organization of behavioral states (top circles), a vector of associative links from the behavioral states to the operant response (middle connections), and the operant response itself (bottom circle).

Behavioral States

I assume that, similar to BeT, each time marker, typically, but not always, food or a reliable predictor of food activates a

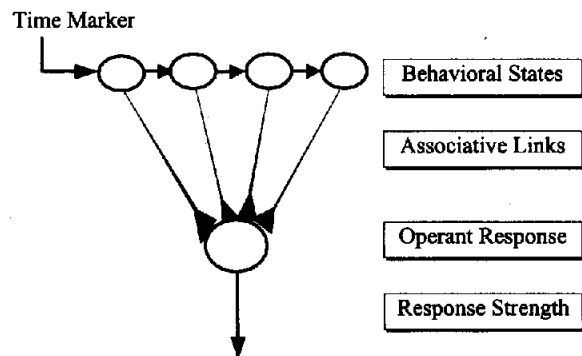


Figure 1. Components of the model. A time marker (e.g., food) initiates a spread of activation across the behavioral states. Each state is associated to some degree with the operant response. The activation of the states multiplied by their associative links determines the strength of the operant response.

series of behavioral states. These states embody our concepts of elicited, induced, adjunctive, interim, and terminal classes of behavior (Falk, 1971; Killeen & Fetterman, 1988; Matthews & Lerer, 1987; Staddon, 1977; Staddon & Simmelhag, 1971; Timberlake & Lucas, 1985). To illustrate, when exposed to periodic presentations of food, pigeons show different types of behavior during different segments of the interfood interval. Elicited behaviors occur immediately after food (e.g., searching for food in the vicinity of the food hopper or pecking around the floor). Interim behaviors occur in the middle of the interval when the probability of food is low (e.g., moving away from the feeder and pacing along the side walls of the chamber). Terminal behaviors occur in the final segment of the interval (e.g., pacing along and pecking at the wall where the food hopper is located). Some authors have emphasized the motivational differences among the states that follow food. Thus, Timberlake and his colleagues (e.g., Timberlake, 1994; Timberlake & Silva, 1995) distinguished general search, focal search, and handling-consuming modes, with each mode corresponding to a different set of stimulus sensitivities, motor patterns, and motor programs. These authors believe that such phylogenetically evolved modes underlie the sequential and temporal organization of behavior. How best to conceptualize the different behaviors observed during the interreinforcement intervals remains an unsettled issue, but for the present approach only the relative invariance of the serial order of the behavioral states is critical.

The activation of each behavioral state is a time-dependent variable with the following dynamics:

$$\frac{d}{dt} X(t, 0) = -\lambda X(t, 0) \quad (1)$$

$$\frac{d}{dt} X(t, n) = \lambda X(t, n-1) - \lambda X(t, n) \quad n \geq 1, \quad (2)$$

where $X(t, n)$ is the activation or strength of state n at time t , and $\lambda > 0$ is a rate parameter that controls how fast the activation spreads across the states. Time t is measured from the preceding time marker. The initial conditions for this system of linear differential equations are

$$X(0, 0) = 1 \quad (3)$$

$$X(0, n) = 0, \quad n > 0, \quad (4)$$

which means that immediately after the time marker (i.e., $t = 0$), state 0 has maximal strength and all remaining states have zero strength. Serial processes with different dynamics are obviously possible, but few would be simpler or more intuitive than the process described in Equations 1–4.

The following analogy may help to interpret the above equations. Imagine a cascade of water compartments as depicted in Figure 2. The compartments, all physically identical and with a hole at the bottom, are the analogues of the behavioral states. The amount of water in each compartment represents the degree of activation of the corresponding state, variable X . The initial conditions, Equations 3 and 4, state that immediately after the time marker, the top compartment is full of water and the others are empty. Behaviorally this means that the time marker elicits a particular behavioral state, numbered 0 in the preceding equations, a state that therefore has full strength, whereas all other states have zero strength.

Water can flow only out of the top compartment, hence Equation 1. For the remaining compartments, water flows both in and out. The amount of water entering compartment n equals the amount leaving compartment $n - 1$, hence Equation 2. Parameter λ , related to the size of the hole, determines how fast the water flows across compartments or, behaviorally, how the wave of activation spreads across states.

Learned Associations

The learning component of the model describes how the behavioral states come to regulate the operant response. Although learning may proceed differently during each state (e.g., the associability with food-related items may be greater for states temporally contiguous with food), my analysis is restricted to a simpler case, namely, that all states follow the same conditioning rules. Specifically, I assume that each state is associated to some degree with the operant response. The degree or strength of the association, represented by variable $W(t, n)$, changes in real time, decreasing during extinction when food is not available and increasing during reinforcement when food is available. However, the amount that $W(t, n)$ changes depends on how active the corresponding state is at time t ; a weakly active state

changes its associative links much less than a strongly active one. These considerations translate into the following equations: during extinction,

$$\frac{d}{dt} W(t, n) = -\alpha X(t, n) W(t, n) \quad (\alpha > 0); \quad (5)$$

during reinforcement,

$$\frac{d}{dt} W(t, n) = \beta X(T, n) [K - W(t, n)] \quad (\beta > 0, t \geq T). \quad (6)$$

K is a parameter related to the incentive value of the reinforcer, and in the present study it always equals 1; α and β are learning rate parameters, and T is the interval from the preceding time marker to the onset of reinforcement. Equation 5 states that during extinction, the associative link between a state and the operant response decreases to 0 at a rate proportional to the current activation of the state. Equation 6 states that during reinforcement that associative link increases to K , its asymptote, at a rate proportional to the activation of the state at the moment the reinforcer was delivered (i.e., T in Equation 6).

Equations 5 and 6 describe one of the oldest, and possibly the simplest, learning rules psychologists have studied, particularly in the domain of choice behavior (e.g., Estes, 1950; Horner & Staddon, 1987; Myerson & Miezin, 1980). They are structurally equivalent to Bush and Mosteller's (1955) linear operator model, although with variable coefficients, and, in the terminology of Restle and Greeno (1970), they are instances of a replacement model for learning.

There are definite advantages to starting with simple linear rules of the form of Equations 5 and 6. First, one can understand better the various properties and predictions of the model. Second, linear rules provide convenient standards against which to compare alternative models. Much the same has happened in the domain of choice theory, for example, where the venerable linear operator model of Bush and Mosteller (1955) has played the useful role of a null hypothesis in assessing the strength of new models (see, e.g., Davis, Staddon, Machado, & Palmer, 1993; Lea & Dow, 1984). Third, in the same way that linear functions provide convenient approximations to nonlinear functions, linear learning models may provide useful approximations to more complex, nonlinear ones.

Response Rule

Finally, the strength of the operant response is obtained by adding the cueing or discriminative function of all states, that is, their associative links, each multiplied by the degree of activation of the corresponding state. In other words, states that are both strongly active and strongly associated with the operant response exert more control over that response than do less active or conditioned states. In quantitative terms, the rate of the operant response equals

$$R(t) = A \sum_{n=1} X(t, n) W(t, n) \quad (A > 0), \quad (7)$$

where A , a scale parameter with units $1/s$, maps the degree of temporal regulation given by the term $\sum_{n=1} X(t, n) W(t, n)$, always less than or equal to 1, onto measurable response rate.

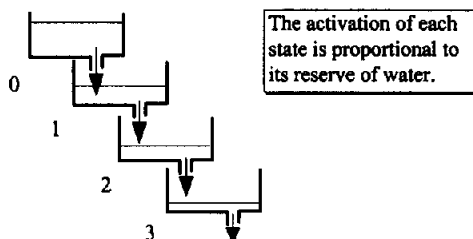


Figure 2. Compartmental model to illustrate the dynamics of the behavioral states. Each compartment corresponds to a behavioral state, and its amount of water corresponds to the degree of activation of the state. Water flows through the compartments according to Equations 1–4.

The summation variable n starts at 1 to force response rate to start at zero, that is, $R(0) = 0$.

Equation 7 permits the distinction between arousal states, variable X , and their influence on operant behavior, variable W , a distinction also emphasized by Killeen (1994) through the concepts of activation and coupling, respectively. Given the multiplicative operation between arousal and associative strength, it follows that states with associative connections close to zero (i.e., $W \approx 0$), even when strongly active (i.e., high X), have no influence on the operant response, although they may control general activity and adjunctive behavior.

In summary, the model assumes that after a time marker, a set of behavioral states is activated in series. During periods of extinction, the states lose their association with the operant response, whereas during reinforcement they increase that association. The strength of the association will therefore come to embody the animal's learning history. At any time, response rate is determined jointly by the profile of the activation of the states and their associative links with the operant response.

Solution of the Equations

The differential equations presented above have closed-form solutions. Thus, Equations 1 to 4 have the solution

$$X(t, n) = \frac{\exp(-\lambda t)(\lambda t)^n}{n!}, \quad (8)$$

which, up to the multiplicative constant λ , is the gamma density function, the cornerstone of BeT (Killeen & Fetterman, 1988). Figure 3 plots this function. Each curve represents the activation of a different state ($n = 1, 2, \dots$) as a function of time, assuming that a time marker occurred at $t = 0$. The activation curves, each with area $1/\lambda$, have mean and variance proportional to n . Because state n is maximally active when $t = n/\lambda$, as time passes different states have a chance to control the operant response. Whether they effectively control the response depends on their associative strength W .

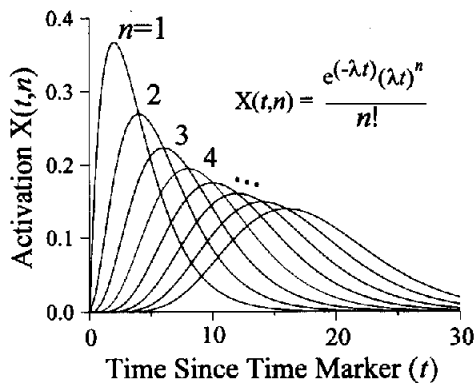


Figure 3. Activation of the states $X(t, n)$ as a function of time since the time marker. Each curve represents a different state: $\lambda = .5$. Each state is maximally active at time $t = n/\lambda$.

Equations 5 and 6 have solutions

$$W(t, n) = W(0, n) \exp \left[-\alpha \int_0^t X(\tau, n) d\tau \right], \quad (9)$$

and

$$W(d, n) = K - (K - W(T, n)) \exp [-\beta X(T, n)d], \quad (10)$$

respectively. The first shows that during extinction the associative strength of a state decays by an amount that depends on the total activation of the state, denoted by the integral in brackets. The second shows that the difference between the associative strength at the onset of reinforcement, that is, $W(T, n)$, and its maximum value K is reduced by an amount that depends on the activation of the state at the time of reinforcement, $X(T, n)$, and the duration of reinforcement, d (in what follows, d is always set at 3 s).

Before comparing the model with experimental data, I will show how it solves the basic problem of credit assignment in time-based schedules, that is, how time eventually controls the behavior of the animal. The explanation also reveals the fundamental building block of the model, one from which all of the remaining properties follow. Figure 4 shows the details.

Credit-Assignment Problem

Suppose that a time marker, presented at Time 0, is followed by 30 s of extinction and then by 3 s of reinforcement. In addition, assume that initially all associative strengths are equal to .8, an arbitrary value chosen for illustrative purposes. Figure 4 shows snapshots of the model at Times $t = 0$ (top panels), $t = 30$ (middle panels), and $t = 33$ (bottom panels). In agreement with Equations 3 and 4, the top left panel shows that state 0 has an initial activation of 1 and that all other states have 0 activation. The horizontal line in the right panel represents the assumption that all associations are initially equal. Because no reinforcer occurs during the next 30 s, Equation 9 applies and $W(0, n)$ decays in proportion to the total activation of state n . The left middle panel plots the total activation of each state at Time $t = 30$,

$$\int_0^{30} X(\tau, n) d\tau,$$

whereas the corresponding $W(30, n)$ values are plotted in the right panel. The earlier states, the most active during the extinction period, lose more associative strength than the later states. Finally, reinforcement occurs for 3 s and, therefore, Equation 10 applies. The activation of the states at the onset of reinforcement, $X(30, n)$, is represented in the left bottom panel. Given the (arbitrary) value of $\lambda = .5$, states 14, 15, and their neighbors are the most active and, consequently, their associative links increase the most. The right bottom panel shows the vector of associative strengths at the end of reinforcement, $W(33, n)$. By comparing the initial and the final distributions of W (top right vs. bottom right), one can see that the states that were more active when food occurred received more credit for the occurrence of food than did the earlier states. As a consequence, they come to exert more control over the operant response. Furthermore, because different states are maximally active at

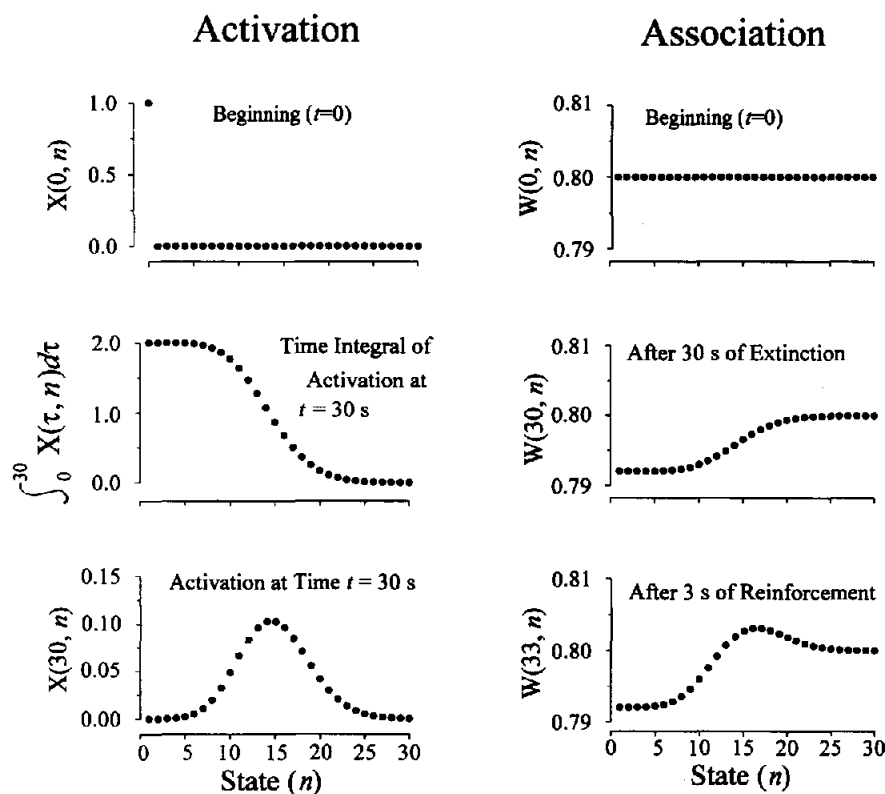


Figure 4. Solving the credit-assignment problem. Top: initial conditions for both the activation of the states (left) and the strength of their association with the operant response (right). Middle: After 30 s of extinction, the strength of the associative links decreased (right) as a function of the integral of the activation of the states (left). Bottom: After 3 s of reinforcement, the associative links increase (right) in proportion to the activation of the states at the onset of reinforcement (left). The most active states receive more credit for the occurrence of reinforcement.

different times since the time marker (see Figure 3), the operant response controlled by them comes to exhibit specific temporal regularities.

Validating the Model

Fixed-Interval Schedules

The simplest environment in which behavior shows temporal properties is possibly the fixed-interval (FI) schedule of reinforcement. In a FI schedule of T seconds, FI- T s for short, a reinforcer follows the first operant response emitted at least T seconds after the previous reinforcer. Typically, at the steady state the animal pauses immediately after the reinforcer and then responds at a constant or slightly increasing rate until the next reinforcer (Ferster & Skinner, 1957; Schneider, 1969; Skinner, 1938). The average response rate in successive segments of the interval generally follows a monotonically increasing, sigmoid-like curve. Figure 5, replotted from Killeen, Hanson, and Osborne (1978), shows representative examples for pigeons responding on FIs of 30, 100, and 300 s (circles).

Predictions of the model. To derive the model's predictions for FI schedules, I first obtained the steady-state distribution of W and then assumed that W does not change appreciably for

the duration of a single trial. This assumption is reasonable whenever the learning parameters α and β are small, a condition I imposed whenever I fit data obtained after prolonged training and sometimes relaxed to explain dynamic effects.

The steady-state distribution of W , derived in Appendix A, equals

$$W(n) = \frac{X(T, n)}{X(T, n) + \gamma \int_0^T X(\tau, n) d\tau}, \quad (11)$$

where $\gamma = \alpha/\beta d$ and $X(t, n)$ is given by Equation 8. The $W(n)$ values depend only on the ratio of the extinction and the reward parameters, γ , henceforth referred to as the learning parameter (by a change of variable, one can eliminate d , reinforcement duration, always equal to 3 s here). By assuming relatively small values for α and β , the number of free parameters is reduced by one.

The response rate across the interval is given by Equation 7, with $W(n)$ given by Equation 11. That is,

$$R(t) = A \sum_{n=1}^{\infty} \frac{X(T, n)}{X(T, n) + \gamma \int_0^T X(\tau, n) d\tau} \times X(t, n), \quad (12)$$

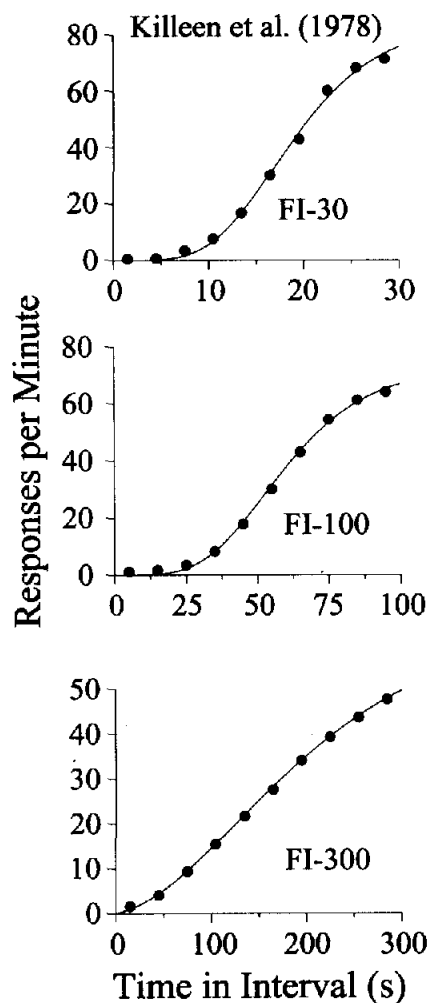


Figure 5. Filled circles: Average response rate of three pigeons in three different fixed-interval (FI) schedules. Curves: predictions of the model with the parameters shown in Table 1. Adapted from "Arousal: Its Genesis and Manifestation as Response Rate," by P. R. Killeen, S. J. Hanson, and S. R. Osborne, 1978, *Psychological Review*, 85, p. 574. Copyright 1978 by the American Psychological Association. Adapted with permission of the author.

where A is the scale factor, the fraction is the total associative strength of the states, and the last term gives the activation-arousal of the states at time t .

The solid lines in Figure 5 show the fit of the model to Killeen et al.'s (1978) data set, and Table 1 shows the parameters used and the variance accounted for (ω^2) by the model. For reasons to be discussed shortly, instead of λ and γ Table 1 reports the dimensionless quantities $\lambda^* = \lambda T$ and $\gamma^* = \gamma T$. Appendix E gives details on the algorithm followed to estimate the parameters.

The model fits the data well. As the FI increased from 30 s to 100 s, the rate and learning parameters remained roughly constant. During the largest FI, λ^* decreased by one order of magnitude, whereas γ^* increased slightly. The disproportionate decrease in λ^* when $T = 300$ s reflects the relative loss of temporal control evidenced by the early rise of the bottom re-

Table 1

Parameters Used to Fit the Data Shown in Figures 5 and 6

Study/condition	λT	γT	A	ω^2
Killeen et al. (1978)				
FI 30	21.36	.381	87.1	.997
FI 100	20.30	.286	75.0	.998
FI 300	7.80	.954	72.8	.999
Dews (1970)				
FI 30	7.95	.648	—	.997
FI 300	7.95	.648	—	.994
FI 3000	7.95	.648	—	.993

Note. FI = fixed interval; λ = rate of transition across the behavioral states; T = duration of interval; γ = relative effects of extinction; A = scale factor; ω^2 = variance accounted for.

sponse rate curve. The scale parameter A did not change appreciably across conditions.

Static models have emphasized one important quantitative property of FI performance. When average response rate along the interval is plotted with both axes normalized, the curves for the same animal during different FI schedules sometimes overlap. Figure 6 from Dews (1970) illustrates this superposition or scalar property of FI performance for intervals ranging from 30 s to 3,000 s.

According to the present model, the response curves for different FIs overlap only if λ and γ remain inversely proportional to the interreinforcement interval. That is, if $\lambda T = k_1$ and $\gamma T = k_2$, where k_1 and k_2 are constants, response rate at Time t will depend only on the ratio t/T (see Appendix A for a proof). For

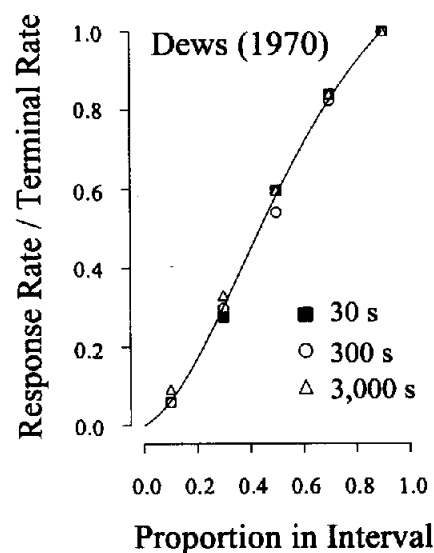


Figure 6. Symbols: relative response rate as a function of relative time into the interval for three different fixed-interval (FI) schedules. Curve: prediction of the model when both the rate and the learning parameters are inversely proportional to the FI value. Parameters in Table 1. Adapted from "The Theory of Reinforcement Schedules," by P. B. Dews, 1970, in W. N. Schoenfeld, *The Theory of Reinforcement Schedules* (p. 48), New York: Appleton & Lange. Copyright 1970 by Appleton & Lange. Adapted with permission.

this reason, I always report the parameters of the model in terms of the quantities λ^* and γ^* . The solid curve in Figure 6 illustrates the fit of the model to Dews's (1970) data, and Table 1 gives the corresponding parameters. As the ω^2 values show, the model fit the data well. (Notice that the superposition property holds for the top two curves in Figure 5 but not for the bottom one, a fact that explains why λ^* decreased significantly at the largest FI.)

Dynamic analysis. A dynamic model is most useful when researchers focus attention on the transient phases of learning. Typical questions concerning those phases are "In an FI schedule, what is the behavioral trajectory followed by the animal before it reaches the steady state?"; "What is performance like during the early stages of extinction following FI training?"; and "Under what conditions is temporal regulation a fast-acting process?"

The behavior of the model during the early moments of exposure to an FI schedule depends on the initial distribution of $W(n)$, which in turn depends on the history of the animal (e.g., the type and amount of its previous training). Consider the

simple case in which an animal is exposed only briefly to continuous reinforcement. A reasonable guess for $W(n)$ at the beginning of FI training is that only the first few states are associated with the operant response. To derive specific predictions, I therefore assumed that the first four states were maximally conditioned, whereas subsequent states had exponentially decaying associative values. Figure 7 shows the development of typical performance in an FI 120-s schedule. The left panel plots response rate across the interval, and the right panel shows corresponding cumulative records. Figure 7 shows four distinct stages of training. On the first stage, corresponding to Trial 1, response rate increases rapidly after reinforcement and then decays to zero; in the cumulative record, one can see an inverted scallop. After 65 trials, the second stage, response rate rises and then remains roughly constant until the end of the interval; in the cumulative record, this profile generates a straight line. After 200 trials, the third stage, response rate does not increase as much immediately after reinforcement, and the cumulative record shows a slight scallop. At the steady state, the fourth stage, the scallop is clearly developed. This sequence of response pro-

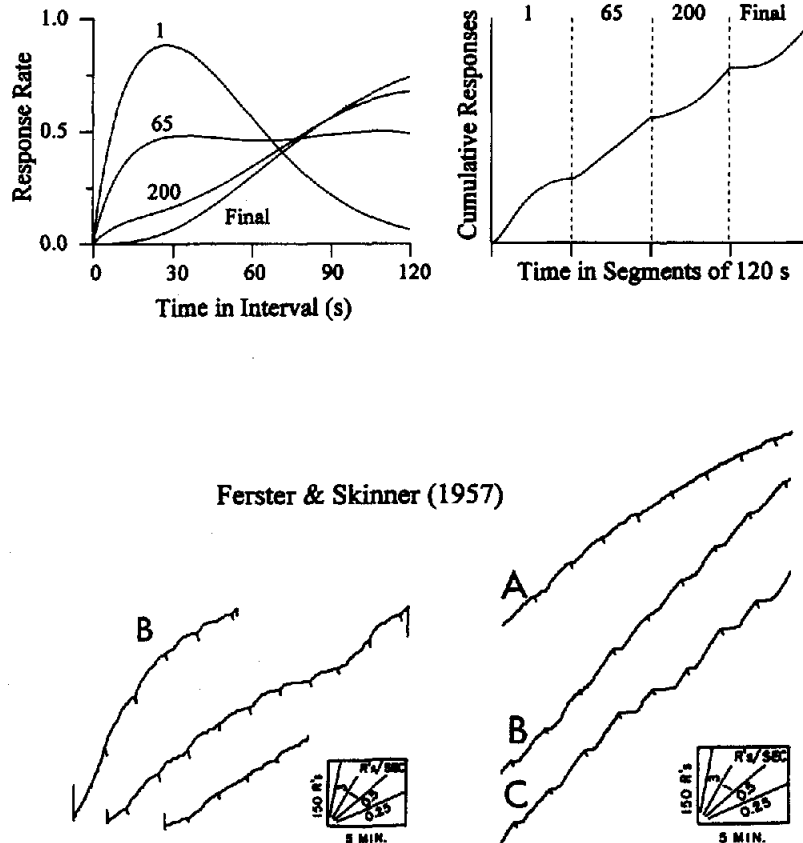


Figure 7. Development of typical fixed-interval (FI) performance. Top left: response rate across a 120-s FI schedule during Trials 1, 65, and 200 and at the steady state. Top right: corresponding cumulative records. Parameters are $\lambda = .1$, $\alpha = .001$, $\beta = .05$, and $A = 1$. Bottom left: cumulative records of one pigeon during the first session of a 120-s FI schedule. The initial inverted scallops change into linear curves. Bottom right: cumulative records of the same bird during the second (at A), fourth (at B), and sixth (at C) sessions. The linear curves at A are replaced by slight scallops at B and full scallops at C. Adapted from *Schedules of Reinforcement* (pp. 137 and 144) by C. Ferster and B. F. Skinner, 1957, New York: Appleton & Lange. Copyright 1957 by Appleton & Lange. Adapted with permission.

files reproduces the characteristic stages of FI acquisition in pigeons, as reported by Ferster and Skinner (1957). The bottom panel of Figure 7 provides an example. The cumulative curves on the left show the transition from continuous reinforcement to a 120-s FI schedule, and the right curves show the behavior of the same pigeon during Sessions 2, 4, and 6 (A, B, and C, respectively). The initial, inverted scallops (see the first two segments on the left) are followed by straight lines (see last segment on the left and segment A on the right), then by slight scallops (segment B on the right) and finally by well-marked scallops (at C). This pattern of effects agrees with the four phases predicted by the model.

Now consider extinction. If an animal on an FI schedule had learned only the time to reinforcement, T , then its response rate should decrease when the time since the last reinforcer exceeds T . However, after prolonged exposure to an FI schedule, response rate in extinction does not decrease for a long time (e.g., Ferster & Skinner, 1957). The present model predicts this result because, as demonstrated in Appendix A, the steady-state values of $W(n)$ increase monotonically with n . Hence, when later states become active in extinction, their strong associative links maintain a high response rate. In addition, if for whatever reason the cascade of behavioral states is reset to its initial conditions, Equations 3 and 4, the model predicts another scallop when responding resumes, although with a lower terminal rate, which again agrees with Ferster and Skinner's description: "The most obvious interpretation [of extinction following FI training] is that the interval performance follows whenever the bird has been pausing for any length of time" (p. 197).

Under some conditions, pigeons adjust quickly to changes in interreinforcement intervals. For example, Higa, Wynne, and Staddon (1991) exposed pigeons to 100 FIs, 99 of which were 15 s long, and one that was 5 s long. The short, or impulse, interval occurred unpredictably in the session. The authors found that 3 birds had reliably shorter postfood pauses just in the 15-s interval after the short impulse; 1 bird also paused less after the impulse, but took a few more trials to recover its baseline pause duration. Unfortunately, Higa et al. did not report average rate data, but the following reasoning shows how the present model can accommodate their findings. Figure 8 presents the details.

The dotted lines in the top panels show response rate curves during the 15-s interval preceding the impulse. The remaining curves show response rate during the 15-s intervals after the impulse, with the leftmost, fastest growing curve (labeled 1 in the graphs) corresponding to the interval immediately after the impulse. On the left panel, the extinction parameter is relatively large ($\alpha = .6$), and therefore the rate curve returns quickly to its baseline value (the curve for the third postimpulse interval is indistinguishable from the baseline dotted curve and for that reason is not presented). On the right, α is smaller and the return to baseline takes longer. If we assume that the duration of the pause is proportional to some value of response strength (say .25, the horizontal lines in Figure 8), then the model predicts the observed pattern of pauses.¹ As the bottom panel shows, when α is large, the pause returns to its baseline value in one or two trials (filled circles); when α is small, the pause takes longer to recover (empty circles).

The effects displayed in Figure 8 depend on the magnitude of α and β and the level of arousal at the end of the short

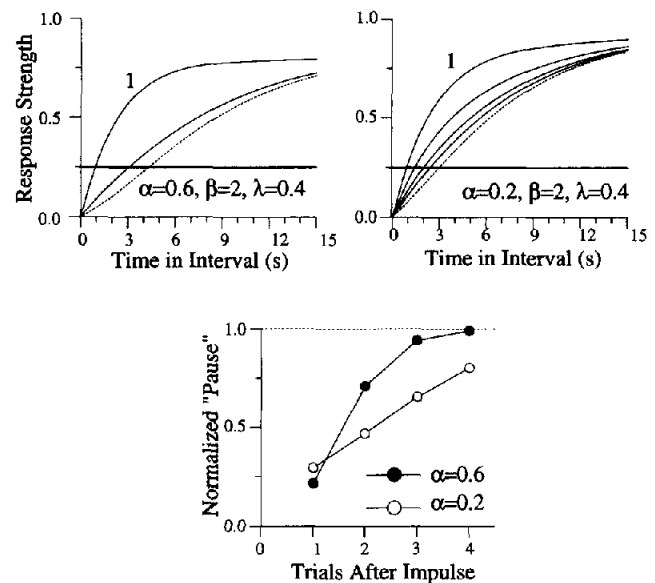


Figure 8. Predictions of the model for Higa et al.'s (1991) impulse experiment. Top: response strength during the last 15-s interval preceding the 5-s impulse (dotted line) and during the 15-s intervals following the impulse (solid lines). The extinction parameter α is larger on the left, and therefore, response strength returns quickly to its baseline. Bottom: If the pause is proportional to a specific value of response strength (e.g., .25, horizontal lines in top panels), then the pause during the postimpulse intervals recovers quickly for large, and slowly for small, α .

interval.² As can be seen from the equations that govern the magnitude of the changes within an interval, Equations 5 and 6, if α and β are very small or if the states are weakly active, then the effects of the shorter interval will be greatly reduced. These two conditions are likely to occur after long impulse intervals, and therefore, I predict that Higa et al.'s (1991) findings will not hold at large FIs.

Peak Procedure

Another procedure commonly used to study temporal regulation is the peak procedure (Catania, 1970). In its simplest version, two types of trials occur in random order. All of the trials begin with the onset of a stimulus (e.g., a keylight). On food trials, the first response after a fixed interval of time (T_1 s) is rewarded; on empty trials, the stimulus remains on for a longer period of time (T_2 s), at the end of which the stimulus is simply turned off and no food is delivered. The experimenter analyzes the behavior of the animal during the empty trials.

After prolonged training, response rate increases from the

¹ The same argument coupled with the assumption that λ^* and γ^* remain constant predicts a linear relation between the duration of the postreinforcement pause and the FI value.

² The role of (high) arousal in engendering the dynamic effects reported by Higa et al. (1991) has not escaped the authors' attention: "Because temporal discrimination is itself so poorly understood—we really have no idea of the processes that allow it to occur—we cannot be sure that arousal and temporal discrimination do not involve overlapping processes" (p. 288).

beginning of the empty trial to the moment when food is primed on food trials and then decreases, sometimes with a positive skew. The top left panel in Figure 9, replotted from Catania (1970), illustrates these findings with pigeons. In this experiment, food and empty trials were 10 s and 48 s long, respectively. In one condition (filled circles) 90% of the trials were food trials ($p = .9$); in another condition (empty circles), only 10% of the trials ended with food ($p = .1$). As expected, overall response rate decreased when fewer trials were reinforced, but the peak of maximum responding remained close to the time when food was primed on food trials (10 s). In the condition $p = .9$, the peak occurred slightly before 10 s, though.

Roberts (1981) replicated Catania's (1970) experiment and obtained the results displayed in the top right panel of Figure 9. In this experiment, rats experienced a mix of 80% food trials and 20% empty trials in the presence of one cue (e.g., tone; filled circles), but the opposite mix in the presence of a different cue (light; empty circles). Food trials were 20 s long, and empty trials averaged 160 s. As before, despite substantial changes in the peak rate of responding, the peak time, around 20 s, did not change significantly.

It is noteworthy that in both Catania's (1970) and Roberts's (1981) studies, the empirical curves obtained with the low p

values are not simple scale transforms of the curves for the high p values. In fact, after scaling the curves along the y-axis, one observes that the low- p curves are wider than the high- p curves, a result that suggests that more than simple changes in overall response rate may be involved.

The bottom panels of Figure 9 illustrate that peak time can change without affecting peak rate substantially. In both experiments, two cues signaled different times to food on food trials, 15 s and 30 s on the left panel (from Roberts, Cheng, & Cohen, 1989), and 20 s and 40 s on the right panel (from Roberts, 1981). Average response rate in the presence of one cue increased with time into the trial, peaked close to the expected time of food in the presence of that cue, and decreased thereafter. It is interesting that the rate curve from Roberts et al.'s study, condition $T_1 = 30$ s, increases again at the end of the empty trial. I will return to this finding in the section on Dynamic Analysis.

These four sets of results show that pigeons and rats can learn to regulate their behavior on the basis of intervals of time, that this learning is resistant to changes in motivational variables because peak rate can vary without changing peak time, and that temporal regulation can be placed under stimulus control because different peak times can be controlled by different cues.

Predictions of the model. The steady-state distribution of $W(n)$ in a peak procedure is approximately equal to

$$W(n) = \frac{X(T_1, n)}{X(T_1, n) + \gamma \int_0^{T_1} X(\tau, n) d\tau + \gamma \frac{1-p}{p} \int_0^{T_2} X(\tau, n) d\tau} \quad (13)$$

where T_1 and T_2 are the durations of food and empty trials, respectively, and p is the proportion of food trials. Appendix B may be consulted for a derivation of Equation 13 and for a study of the roles of the parameters λ and γ . Response rate is obtained from Equation 7, with the values of W coming from Equation 13.

The solid lines in Figure 9 illustrate the fit of the model to the four data sets from the peak procedure, and Table 2 gives the corresponding parameters. Consider Catania's (1970) data. When $p = .1$, the model accounted for the data well. However, when $p = .9$, the predicted curve (not shown) peaked close to 10 s, the time when food was effectively primed on food trials, but clearly to the right of the data points, centered around 7 s. One way to interpret the observed anticipation of the moment of food is to assume that the time required to initiate the activation of the behavioral states after the time marker (Equations 1 and 2) is necessarily positive, whereas so far, I have assumed it to be zero. If this assumption is correct, then the effective duration of a food trial (i.e., the interval from the beginning of the serial process to food) could be less than 10 s. I therefore assumed an average effective interval of 7 s. The new prediction of the model, the solid line through the filled circles, fit the data well.

The model also fit well Roberts's (1981) replication of Catania's (1970) study (see curves on top right panel of Figure 9). The first four rows of Table 2 show that when reinforcement rate decreased in both Catania's and Roberts's studies, the rate

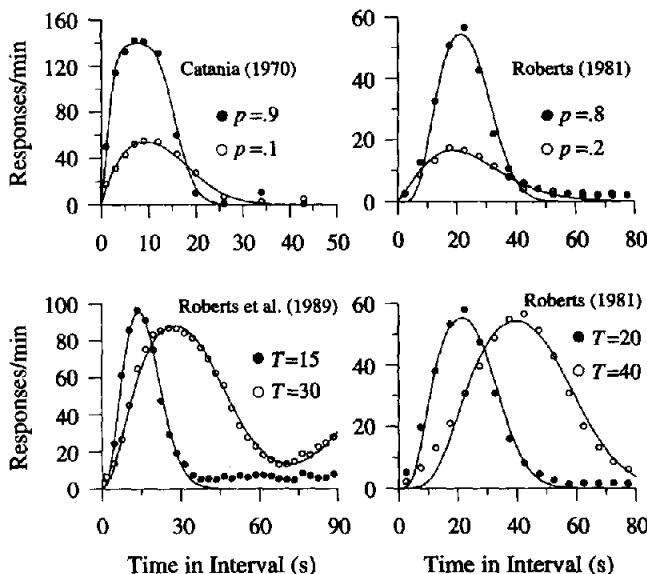


Figure 9. Symbols: data from the peak procedure. In the top panels, the probability p of food trials varied across conditions; in the bottom panels, two cues signaled different intervals to food, T . Curves: predictions of the model with the parameters shown in Table 2. Adapted from "Reinforcement Schedules and the Psychophysical Judgments: A Study of Some Temporal Properties of Behavior," by A. C. Catania, 1970, in W. N. Schoenfeld, *The Theory of Reinforcement Schedules* (p. 11), New York: Appleton & Lange. Copyright 1970 by Appleton & Lange. Adapted with permission. Adapted from "Isolation of an Internal Clock," by S. Roberts, 1981, *Journal of Experimental Psychology: Animal Behavior Processes*, 7, p. 245. Copyright 1981 by the American Psychological Association. Adapted with permission of the author. Adapted from "Timing Light and Tone Signals in Pigeons," by W. A. Roberts, K. Cheng, and J. S. Cohen, 1989, *Journal of Experimental Psychology: Animal Behavior Processes*, 15, p. 27. Copyright 1989 by the American Psychological Association. Adapted with permission of the author.

Table 2
Parameters Used to Fit the Data Shown in Figure 9

Study	Condition	T_1	λT_1	γT_1	A	ω^2
Catania (1970)	$p = .9$	7 ^a	12.95	5.56×10^{-3}	140.0	.992
	$p = .1$	10	6.97	9.74×10^{-3}	63.3	.980
Roberts (1981)	$p = .8$	20	21.40	2.76×10^{-1}	63.9	.976
	$p = .2$	20	5.00	3.96×10^{-1}	64.0	.951
Roberts (1981)	$p = .8$	20	16.80	1.75×10^{-1}	62.6	.982
	$p = .8$	37 ^a	19.24	3.23×10^{-1}	65.7	.970
Roberts et al. (1989)	$p = .75$	13 ^a	10.66	5.92×10^{-1}	148	.976
	$p = .75$	26 ^a	10.14	1.82×10^{-1}	105	.993

Note. T_1 = duration of the shorter interval; λ = rate of transition across the behavioral states; γ = relative effects of extinction; A = scale factor; ω^2 = variance accounted for; p = reinforcement probability at the end of the shorter interval.

^a Duration of food intervals slightly smaller than the real value.

parameter λ^* decreased but the learning parameter γ^* did not change appreciably. However, in Catania's study, parameter A decreased markedly with reinforcement rate, whereas in Roberts's study, A remained unchanged. It is interesting that a similar behavioral output, the substantial change in response rate across conditions, was achieved by different sets of parameter changes.

The predictions of the model for the experiments in which two cues signaled different times to food are the curves in the bottom panels of Figure 9. On the right panel and for $T_1 = 20$ s, the model fit the data well. However, in all other cases, the model predicted a response curve (not shown) to the right of the observed data points, a deviation analogous to that observed for Catania's (1970) data. As before, the fit improved substantially when I assumed a duration of food trials slightly smaller than the real value (see Footnote a in Table 2, column T_1). The parameter values for these experiments did not change appreciably across conditions, except the scale factor A in Roberts et al.'s (1989) experiment that decreased in the presence of the cue associated with the larger interfood interval.

Dynamic analysis. A peak procedure with food trials T_1 s long is closely related to a T_1 -s FI schedule, and this fact is revealed in the present model by the similarity of the corresponding steady-state equations for $W(n)$, Equations 11 and 13. The only difference between the two equations is the rightmost term in the denominator of Equation 13. When the proportion p of food trials equals 1, the peak procedure reduces to an FI - T_1 s schedule and Equation 13 simplifies to Equation 11, as it should. The similarities and differences between the two equations are further illustrated in Figure 10. The top panel plots the steady-state values of W in an FI schedule and in a comparable peak procedure. In the FI schedule, the associative values increase monotonically with the state number, but in the peak procedure, they increase to a maximum and then decrease to zero. This difference in the behavior of W happens because in the peak procedure the later states become active during the long periods of extinction, the empty trials, and therefore they lose more associative strength than they gain during reinforcement (Appendix B provides a proof of this statement). In the FI schedule, later states have no opportunity to lose their association, which gradually increases to 1. Furthermore, given that the initial states have similar associative values in the FI and in the peak procedure, the model predicts similar response rate curves up to the moment when food is primed (i.e., $t = T_1$ s).

The bottom panel in Figure 10 illustrates this prediction. Although the curve for FI schedules stays high if we allow the trial to continue past the usual reinforcement time, the curve for the peak procedure decreases in a slightly asymmetrical fashion.

The preceding analysis helps one to understand the tail of the response rate curve observed in the bottom left panel of Figure

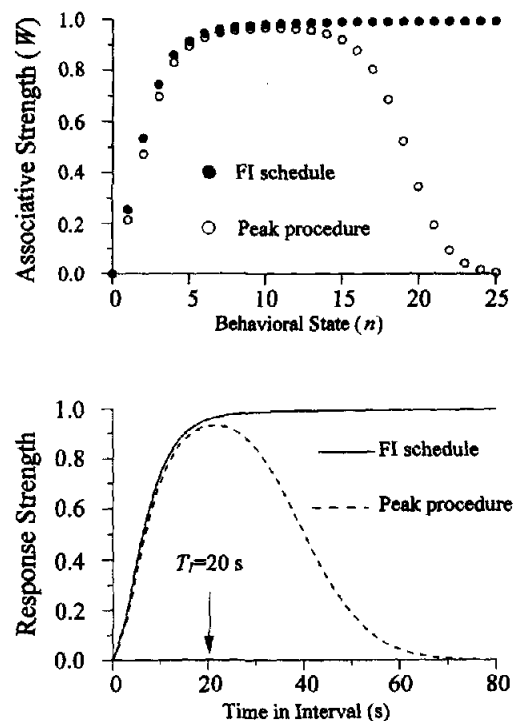


Figure 10. Comparing the predictions of the model for simple fixed-interval (FI) schedules and for the peak procedure. Top: distribution of associative values in a 20-s FI schedule and in a peak procedure with 20-s food trials, 160-s empty trials, and a .8 probability of a food trial. Associative strength increases monotonically with n in the FI schedule, but is bitonic in the peak procedure. Bottom: response rate for the two schedules. Whereas responding in the FI remains high after the usual time of reinforcement, it decreases with some skew in the peak procedure. Parameters: $\lambda = .5$, $\gamma = .003$, $A = 1$.

8 (empty circles). This result, obtained also in other studies with the peak procedure (e.g., Cheng & Roberts, 1991), can be predicted by the present model if we take into account the experimental history of Roberts et al.'s (1989) birds. Initially, the pigeons experienced an FI 30-s schedule in the presence of one cue. Next, during the peak procedure, the same cue signaled either a 30-s food or a 90-s empty trial. The peak procedure was run for 25 sessions of 24 trials each (another 24 trials involved the cue associated with the 15-s food trials and are not discussed here). Hence, the birds experienced a total of 600 (24×25) trials, a number that may have been insufficient to attain the new steady state. If behavior is not at the steady state, then response rate should increase twice on empty trials because the later states did not lose completely the association they gained during the FI training. To check the preceding account, I derived the predictions of the model after 600 trials, with the assumption that the behavioral states were fully conditioned at the end of the FI schedule (see Appendix B for details). As the solid curve through the empty circles shows (see left bottom panel of Figure 9), the model accounted well for the response rate profile. If one assumes that the steady state has been reached, the predicted response curve remains the same up to 70 s into the trial, but thereafter decays to zero instead of increasing.

Finally, by following the same steps that led to the derivation of Dews's (1970) results (see Figure 6 and Appendix A), one can also show that response curves obtained with different food trial durations will superpose when plotted in relative time only if λ and γ remain inversely proportional to the duration of the food trials.³ In other words, both in the FI schedule and in the peak procedure, the scalar property requires proportionately slower transitions across behavioral states and smaller relative effects of extinction as reinforcers become less frequent.

Timing Two Intervals

Some experiments show that animals can also track two intervals of time in the absence of external cues. For example, Catania and Reynolds (1968) presented pigeons with two FI schedules in mixed order. After each reward, the programming apparatus decided whether the next FI would be T_1 or T_2 s long ($T_1 < T_2$). The authors varied the length of the short interval, T_1 , and its relative frequency, p , across conditions and measured the average response rate in successive segments of the long interval.

Figure 11 shows the effect on the behavior of a single bird of varying T_1 while keeping p constant at .05. From the top to the bottom panel, T_1 equaled 30, 90, 150, 210, and 240 s (see left asterisks in Figure 11); T_2 was always 240 s (right asterisks). In the bottom panel, $T_1 = T_2$, and therefore the schedule is equivalent to a simple FI 240-s schedule. Three features of the data are worth noting: As T_1 increases, responding starts later in the interval; response rate at T_1 s is always lower than the terminal rate at 240 s; and when the short interval is 30 s long (top panel), response rate clearly shows two peaks, centered at T_1 and T_2 , which means that pigeons are capable of learning more than one interfood interval and regulate their behavior appropriately.

Figure 12 shows the results for another bird, but here the probability of the short interval equaled .5. From top to bottom, $T_1 = 30, 90, 210$, and 240 s. Results illustrate the same features mentioned above, except that when T_1 equaled 30 s or 90 s,

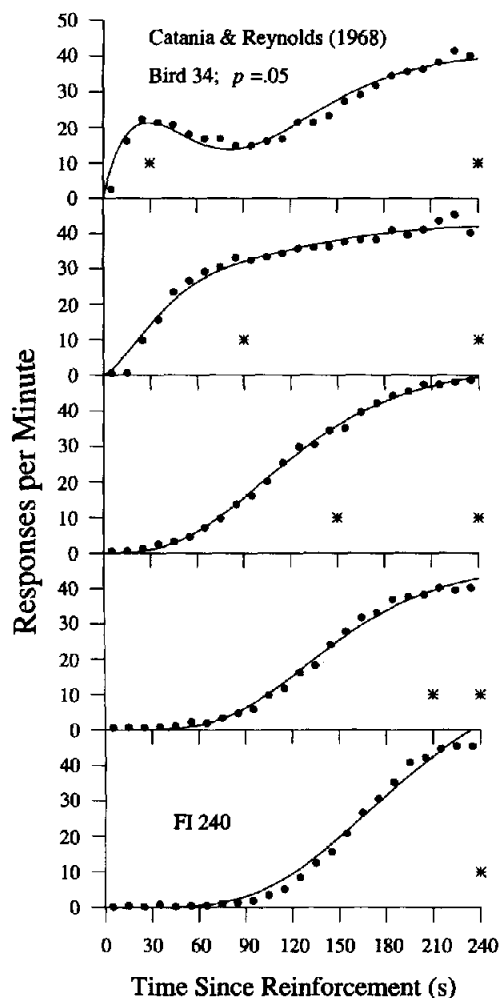


Figure 11. Circles: response rate of one pigeon in mixed fixed-interval–fixed-interval (Mixed FI–FI) schedules. The asterisks indicate the duration of the two FIs. The shorter FI occurred with probability $p = .05$. The bottom panel is equivalent to a simple 240-s FI. Curves: prediction of the model with the parameters shown in Table 3. Adapted from "A Quantitative Analysis of the Responding Maintained by Interval Schedules of Reinforcement," by A. C. Catania and G. S. Reynolds, 1968, *Journal of the Experimental Analysis of Behavior*, 11, p. 360. Copyright 1968 by the Society of the Experimental Analysis of Behavior, Inc. Adapted with permission.

response rate around T_1 s was equal to, or slightly greater than, the terminal response rate.

Figure 13 allows a direct comparison of both types of manipulation, a change in either the relative frequency or the duration of the short interval. In the top panel, p varies while T_1 remains constant at 30 s; in the bottom panel, T_1 varies while p remains

³ Strictly speaking, the scalar property also requires that the duration of the empty trials be a constant multiple of the duration of food trials. However, when empty trials are much longer than food trials (as is typically the case), this requirement becomes irrelevant for practical purposes because it affects only the far right tail of the curves, where not much responding is observed.

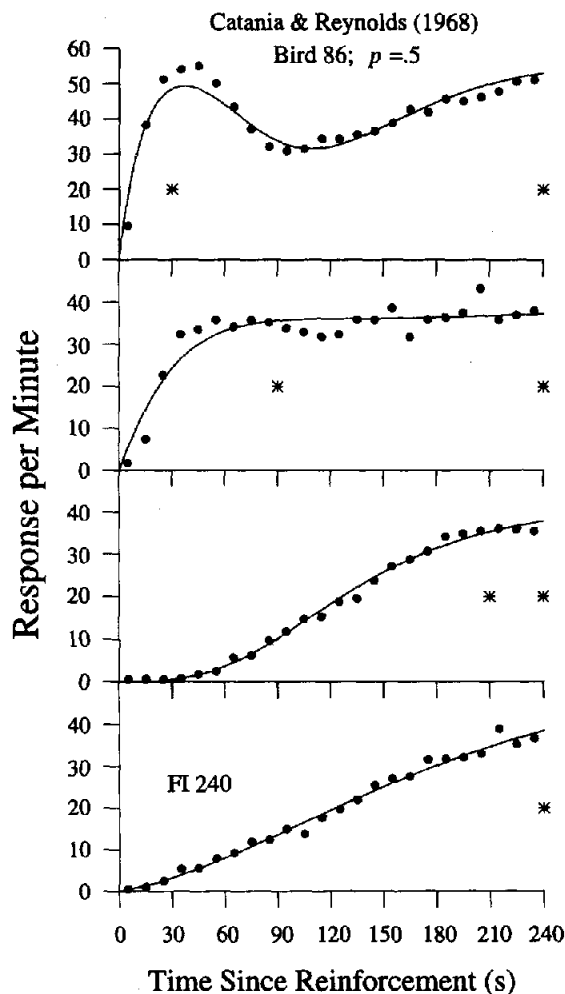


Figure 12. Circles: response rate of one pigeon in mixed fixed-interval–fixed-interval (Mixed FI–FI) schedules. The asterisks indicate the duration of the two FIs. The shorter FI occurred with probability $p = .5$. Curves: prediction of the model with the parameters shown in Table 3. Adapted from "A Quantitative Analysis of the Responding Maintained by Interval Schedules of Reinforcement," by A. C. Catania and G. S. Reynolds, 1968, *Journal of the Experimental Analysis of Behavior*, 11, p. 361. Copyright 1968 by the Society of the Experimental Analysis of Behavior, Inc. Reproduced with permission.

constant at .05. The filled circles are equivalent to an FI 240-s schedule. The data show that increasing the relative frequency of the short interval or decreasing its length produced earlier onsets of responding.

Predictions of the model. The steady-state distribution of $W(n)$ in a mixed FI T_1 -s–FI T_2 -s schedule is derived in Appendix C and equals

$$W(n) = [pX(T_1, n) + (1 - p)X(T_2, n)] / [p[X(T_1, n) + \gamma \int_0^{T_1} X(\tau, n) d\tau] + (1 - p)[X(T_2, n) + \gamma \int_0^{T_2} X(\tau, n) d\tau]] \quad (14)$$

Notice the symmetry of the equation: The terms multiplied by p and $1 - p$ correspond to the equations for FIs T_1 and T_2 s long, respectively. Hence, when $p = 0$ or 1 , the mixed FI–FI schedule reduces to a simple FI schedule, and Equation 14 reduces to Equation 11, as it should. Response rate is obtained from Equation 7, with the W values given by Equation 14.

The solid lines in Figures 11, 12, and 13 show the fit of the model to Catania and Reynolds's (1968) data, and Table 3 shows the corresponding parameters and ω^2 values. Because there were two interfood intervals, I used the average time to food (\bar{T}) to calculate λ^* and γ^* . In Figure 11, the rate of reinforcement did not change appreciably and λ^* and A remained approximately constant (Table 3, Bird 34). With one exception, parameter γ^* tended to increase with T_1 . The ω^2 values indicate that the model fits the data well when p is small.

In Figure 12, reinforcement rate varied more substantially (see Table 3, Bird 86), but the rate parameter λ^* changed only in the last condition when the degree of temporal control was clearly reduced (compare the initial segments of the response rate curves for conditions $T_1 = 150$ and $T_1 = 240$). The variation of γ^* was similar to the preceding case—as the intervals be-

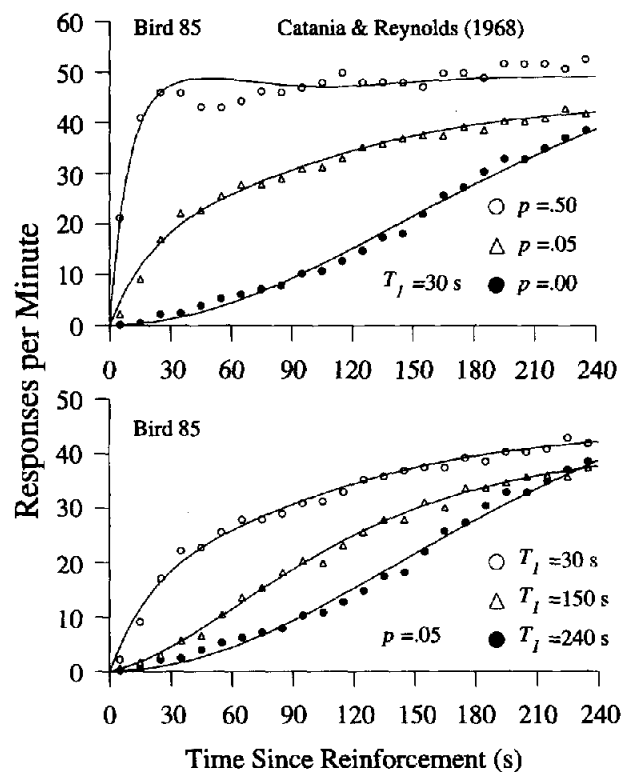


Figure 13. Symbols: response rate of one pigeon in mixed fixed-interval–fixed-interval (Mixed FI–FI) schedules. Top: The duration of the short FI, T_1 , was 30 s and it occurred with probabilities .5, .05, and 0. Bottom: The short FI was 30-, 150-, and 240-s long and occurred with probability .05. Curves: prediction of the model with parameters shown in Table 3. Adapted from "A Quantitative Analysis of the Responding Maintained by Interval Schedules of Reinforcement," by A. C. Catania and G. S. Reynolds, 1968, *Journal of the Experimental Analysis of Behavior*, 11, p. 361. Copyright 1968 by the Society of the Experimental Analysis of Behavior, Inc. Adapted with permission.

Table 3
Parameters Used to Fit Catania and Reynolds's (1968) Data
(See Figures 11, 12, and 13)

T_1	p	Rft/hr	$\lambda\bar{T}$	$\gamma\bar{T}$	A	ω^2
Pigeon 34						
30	.05	15.69	19.28	1.41×10^{-1}	42.0	.958
90	.05	15.48	19.76	4.72×10^{-2}	43.1	.977
150	.05	15.29	17.90	2.18×10^{-1}	54.3	.998
210	.05	15.09	20.27	3.46×10^{-1}	48.7	.994
240	.05	15.00	20.40	1.76	79.7	.989
Pigeon 86						
30	.5	26.67	11.21	1.17×10^{-1}	58.4	.898
90	.5	21.82	13.20	6.39×10^{-2}	38.4	.880
210	.5	16.00	14.40	4.68×10^{-1}	44.3	.994
240	.5	15.00	6.72	1.37	64.0	.989
Pigeon 85						
30	.5	26.82	14.58	1.76×10^{-3}	49.5	.840
30	.05	15.69	8.49	1.04×10^{-1}	44.7	.989
150	.05	15.48	9.18	2.66×10^{-1}	42.9	.996
240	.05	15.00	8.40	2.40	77.1	.992

Note. T_1 = duration of the shorter interval; \bar{T} = average interreinforcement interval; p = reinforcement probability at the end of the shorter interval; Rft/hr = observed reinforcement rate per hour; λ = rate of transition across the behavioral states; γ = relative effects of extinction. A = scale factor; ω^2 = variance accounted for.

came more similar or, equivalently, as the variance of the interfood intervals decreased, the relative effects of extinction increased. Again, A did not change appreciably with T_1 .

The preceding interpretation of the parameters must be taken with some reservation because, as the ω^2 values indicate, the fit of the model was not very good for the conditions $T_1 = 30$ and $T_1 = 90$. In the latter condition, the data themselves are noisy (see horizontal segment of the curve), but the same reason cannot explain the relatively poor fit in condition $T_1 = 30$. I will return to this problem shortly.

In Figure 13, λ^* decreased disproportionately when the overall reinforcement rate decreased (cf. first two rows of Bird 85, Table 3) but then remained constant, whereas γ^* increased as the variance of the interfood intervals decreased. Parameter A did not change appreciably across conditions. From the ω^2 values, one can see that the fit was good when $p = .05$ but was not as good when $p = .5$, because in addition to the noise in the data, the observed and predicted curves are slightly out of phase.

In summary, the model fit the data well when the probability of the short interval was low. When that probability was high, however, the model (a) could not generate a response rate at T_1 s greater than the terminal rate, and (b) predicted a profile of response rate across the interval slightly out of phase with the profile of the bird's data. No combination of parameter values can correct these errors because, as I explain next, they stem from the basic serial structure of the model.

Dynamic analysis. According to the present model, the procedure with two distinct FIs presented in random order is analogous to the combination of two other procedures.⁴ To see this, consider one of the late behavioral states (roughly, state n , with

n of order λT_2). When the reinforcer occurs at the end of the short FI, this state is not strongly active and therefore its associative strength changes little. However, when the reinforcer comes after the long FI, this state will be active and its association with the operant response will increase. Thus, for later states, the procedure is similar to a simple FI T_2 s schedule. Consider now an early state, one that is mostly active at the end of the short FI (roughly, state n , with n of order λT_1). For this state, the procedure is analogous to the peak procedure, with the long FIs playing the role of empty trials; when the reinforcer happens at the end of a long FI, this early state is weakly active and so its associative strength does not increase appreciably. If the preceding analysis is correct, performance during a mixed FI- T_1 -s-FI- T_2 -s schedule should be derivable from the predictions of the model for a simple FI- T_2 s long and a peak procedure with food trials T_1 s and empty trials T_2 s long. Figure 14 shows that this is indeed the case.

The top panel reproduces the data of Bird 34 in Catania and Reynolds's (1968) study, condition $T_1 = 30$, $T_2 = 240$, and $p = .05$. It also shows three theoretical curves: the curve for the mixed FI-FI schedule, the curve for a simple FI 240-s schedule, and the curve for a peak procedure with food trials 30 s long, empty trials 240 s long, and a proportion of food trials equal to .05. All of the curves used the parameters from the first row of Table 3. The bottom panel shows the same curves when the short interval is more frequent ($p = .5$), the case that the model could not fit adequately. The parameters were arbitrarily set at $\lambda = .05$, $\gamma = 0.001$, and $A = 1$.

Initially, when the early states control behavior, the curves for the mixed schedule follow closely the curves for the peak procedure; later, when subsequent states exert greater control, they approach the curves for the simple FI schedules. To understand the interrelation of the three curves in Figure 14, notice that, for small n , $X(T_2, n)$ in Equation 14 is not only close to zero but is also much smaller than the remaining terms; setting $X(T_2, n) = 0$ yields the equation for the peak procedure, Equation 13. Conversely, for large values of n , all terms involving $X(T_1, n)$ are much smaller in magnitude than the terms involving $X(T_2, n)$; eliminating them yields the equation for an FI- T_2 s, Equation 11.

From the preceding analysis, two important consequences follow. First, given that later states are always more conditioned than earlier states in FI schedules (see Figure 10, top panel), response rate is always maximal at the end of the trial. This finding explains the model's first difficulty, to predict cases in which response rate is higher at T_1 s than at T_2 s. Second, the dip of the curve for the mixed schedule, a local minimum of response rate, is close to the point at which the curves for the peak procedure and the simple FI schedule intersect (see Figure 14). When $p = 1/2$, this point is basically invariant to changes in either the learning parameter γ or the rate parameter λ . Although each of these parameters changes the shape of the two component curves, the changes are in opposite directions and therefore cancel each other. For example, increasing λ narrows

⁴ Catania and Reynolds (1968) concluded that performance in mixed FI-FI schedules is a combination of performance under two separate FIs with corresponding values. The analysis that follows shows that this conclusion is incorrect.

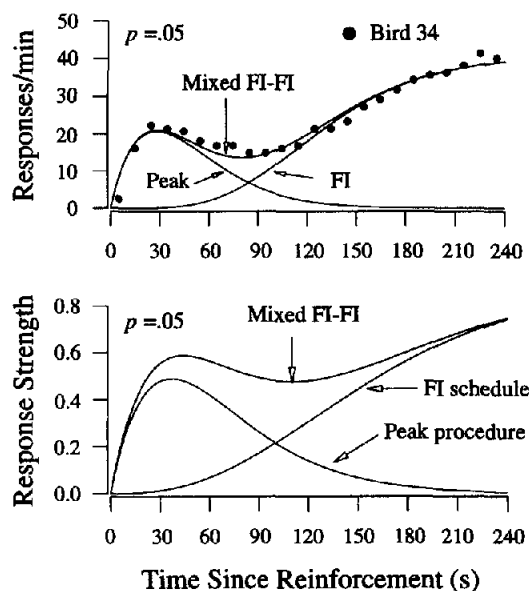


Figure 14. Derivation of the curve for a mixed fixed-interval (30 s)-fixed-interval (240 s) schedule (Mixed FI-FI), with $p = .05$ (top) and $p = .5$ (bottom), from the curves for a simple FI 240 s and a peak procedure with food trials 30 s long, empty trials 240 s long, and the same p values. Parameter values are $\lambda = .084$, $\gamma = .000617$, $A = 42$ (top), and $\lambda = .05$, $\gamma = .0033$, $A = 1$ (bottom). Adapted from "A Quantitative Analysis of the Responding Maintained by Interval Schedules of Reinforcement," by A. C. Catania and G. S. Reynolds, 1968, *Journal of the Experimental Analysis of Behavior*, 11, p. 360. Copyright 1968 by the Society of the Experimental Analysis of Behavior, Inc. Adapted with permission.

the curve for the peak procedure by shifting its descending limb to the left, but it also shifts the curve for the simple FI schedule to the right; the intersection point does not move.

More specifically, I show in Appendix C that when $p = 1/2$, the curves for the peak procedure and the simple FI intersect approximately at Time t^* , equal to

$$t^* = \frac{T_2 - T_1}{\ln(T_2/T_1)} \quad (15)$$

Notice that Equation 15 is parameter free. The invariance of the local minimum of response rate is the source of the model's second difficulty: If the bird shows a dip in its response rate profile at another location, the model cannot accommodate it. I discuss the implications of these "errors" in the final section of the study.

Bisection Experiment

So far I have examined situations in which the animal had to time intervals as they elapsed. A different set of procedures studies how behavior may be controlled by past stimulus dura-

tions. In what follows, I describe the major findings obtained with one such procedure, the bisection of temporal intervals, and then extend the model to deal with this new situation. Another set of studies has examined how animals time their own responses (e.g., how rats adjust the duration of their lever presses to some reinforcement rule; see Platt, 1979, for a review), but they will not be addressed here.

In the bisection experiment, the animal is exposed to either a short- or a long-duration stimulus. To obtain food, the animal must then report during a two-choice phase which duration has just occurred (e.g., if the stimulus was *short*, press the right lever; if the stimulus was *long*, press the left lever; Catania, 1970; Church & Deluty, 1977; Platt & Davis, 1983; Stubbs, 1968). Once the animal discriminates between the two stimuli, the experimenter introduces intermediate durations on test trials. The test duration that produces the same number of choices of the left and right levers, that is, the duration judged to be halfway between the two training stimuli, is referred to as the bisection, or indifference, point.

Catania (1970) reported one of the first experiments with the bisection procedure. To obtain food, a squirrel monkey had to press a left lever three times after a 1-s tone and a right lever three times after a 4-s tone. Control by intermediate durations was assessed on probe trials with no food. The top panel of Figure 15 (circles) shows the results. The proportion of short responses was close to 1 after the short stimulus and close to 0 after the long stimulus. The bisection point, obtained by linear interpolation, was close to 2.3 s, less than the arithmetic average (2.5 s) but slightly greater than the geometric average (2.0 s) of the two training stimuli.

The middle panel of Figure 15 shows the average results obtained by Fetterman and Killeen (1991) with 4 pigeons trained with 4-s and 12-s stimuli. The vertical bars show the standard error of the means. The obtained mean bisection point equaled 7.31, again less than the arithmetic mean (8.00) but slightly greater than the geometric mean (6.92) of the two training stimuli.

Church and Deluty (1977) performed a more extensive study with the bisection procedure. Rats learned to press the left lever after one signal duration and the right lever after another duration. One signal was always 4 times longer than the other. The bottom panel of Figure 15 shows the average results for four pairs of short and long stimulus durations. The proportion of short responses is plotted as a function of the relative duration of the signal, with the short duration as the standard. The superposition of the data points suggests a Weber-like property for temporal discrimination: Equal ratios yield equal discriminabilities (Gibbon, 1981). Furthermore, the bisection point in all four conditions is very close to the geometric mean (2.0) of the training stimuli.

Extending the model. To deal with the bisection results, the basic model was modified slightly. Obviously, two responses and a rule to determine which response occurs during the choice phase are needed. In addition, given that the animal can respond only after the signal terminates, learning is restricted to the choice phase. To accommodate this fact, I changed the learning rule from continuous to discrete time and applied it only after each choice response. The extended model is illustrated in Figure 16 and works as follows: The onset of the training stimulus, the time marker, triggers the transitions across the behavioral

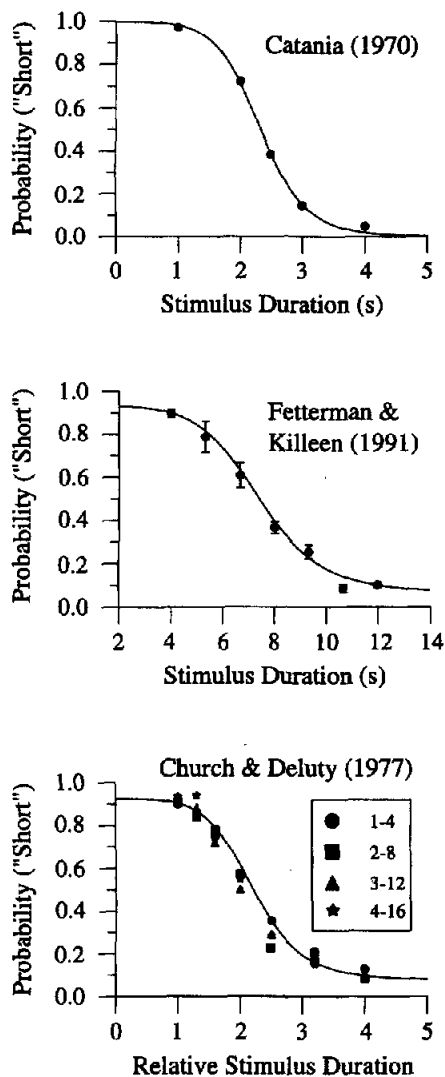


Figure 15. The symbols show the proportion of short responses as a function of stimulus duration in three experiments. The left and rightmost durations are the two training stimuli. The vertical bars in the middle panel show the standard error of the means. In the bottom panel, the data from four conditions (training stimuli 1-4, 2-8, 3-12, and 4-16 s) are shown with the abscissa expressed in terms of relative stimulus duration, with the short stimulus as the standard. Animals were a squirrel monkey (top panel), 4 pigeons (middle panel), and 6 rats (bottom panel). Curves: predictions of the model with the parameters shown in Table 5. Adapted from "Reinforcement Schedules and the Psychophysical Judgments: A Study of Some Temporal Properties of Behavior," by A. C. Catania, 1970, in W. N. Schoenfeld, *The Theory of Reinforcement Schedules* (p. 9), New York: Appleton & Lange. Copyright 1970 by Appleton & Lange. Adapted with permission. Adapted from "Adjusting the Pacemaker," by J. G. Fetterman and P. R. Killeen, 1991, *Learning and Motivation*, 22, p. 234. Copyright 1991 by Academic Press. Adapted with permission. Adapted from "Bisection of Temporal Intervals," by R. M. Church and M. Z. Deluty, 1977, *Journal of Experimental Psychology: Animal Behavior Processes*, 3, p. 220. Copyright 1977 by the American Psychological Association. Adapted with permission of the author.

states described by Equation 8 (see top circles in Figure 8). When the stimulus ends, the animal responds left or right (bottom circles). If the response is reinforced, then the association between each state and the choice response increases and the association between each state and the opposite response decreases by the same amount. These associations are represented by the two sets of middle arrows in Figure 16. If the choice response is incorrect, then the association between that response and each state decreases and the association between the states and the opposite response increases.

Table 4 presents the discrete-time equations of the model. Let the left (L) response be the correct choice after the short (S) stimulus and the right response (R) be the correct choice after the long (L) stimulus. The top left cell of Table 4 states that when the animal responds left after the short stimulus and therefore receives food, the association between state n and the left response increases by an amount equal to $\Delta WL(n) = \beta \cdot X(S, n) \cdot [1 - WL(n)]$, where β is a reinforcement parameter, $X(S, n)$ is the activation of state n after the short stimulus, and $WL(n)$ is the current association between behavioral state n and the left response. The association between state n and the right response decreases by the same amount. If after stimulus S, the animal responds right and therefore does not receive food (top right cell), the association between state n and the right response decreases by an amount equal to $\Delta WR(n) = -\alpha \cdot X(S, n) \cdot WR(n)$, where α is an extinction parameter. The association between state n and the left response increases by the same amount. The remaining cells of the table have similar interpretations. In summary, this reward-following-learning rule, structurally identical to Equations 5 and 6, states that responses compete for a fixed amount of associative strength and that associative strength changes with reinforcement and extinction.

After a stimulus with duration t , the strengths of the left and right responses equal

$$RL(t) = \sum_{n=0} WL(n)X(t, n) \quad (16)$$

and

$$RR(t) = \sum_{n=0} WR(n)X(t, n) = 1 - RL(t).$$

Time Marker

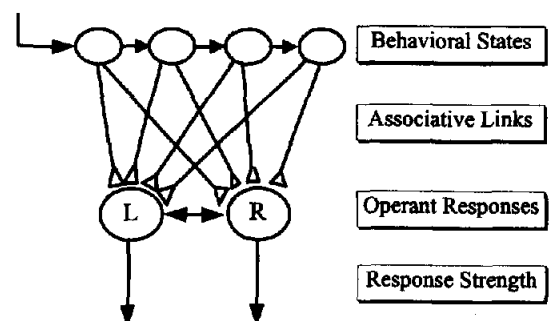


Figure 16. Extended model to deal with choice situations such as the bisection of temporal intervals. A time marker initiates the activation of the behavioral states (top). The states become associated with both responses (middle arrows). The choice responses (bottom) compete for the final common path (horizontal double arrow).

Table 4
Change in the Associative Strength $WL(n)$ According to the Stimulus and the Response That Occurred on the Trial

Stimulus	Left response	Right response
Short	$\Delta WL(n) = \beta X(S, n)(1 - WL(n))$	$\Delta WL(n) = \alpha X(S, n)(1 - WL(n))$
Long	$\Delta WL(n) = -\alpha X(L, n)WL(n)$	$\Delta WL(n) = -\beta X(L, n)WL(n)$

Note. $WL(n)$ = the association between state n and the left response; $\Delta WL(n)$ = change in $WL(n)$ at the end of the trial; α = extinction parameter; β = reinforcement parameter; $X(S, n)$ = activation of state n at the end of the short stimulus; $X(L, n)$ = activation of state n at the end of the long stimulus.

On short trials, $t = S$; on long trials, $t = L$; and on probe trials, t may take any value, generally somewhere between these two.

The quantities $RL(t)$ and $RR(t)$ describe the strength of the two responses at the time that the signal ends and the choice phase begins. The actual choice predicted by the model depends on how these two quantities are combined in a decision rule. I propose the following rule⁵ for the probability of a left response after a stimulus with duration t ,

$$P(L|t) = \frac{\exp[\delta \cdot RL(t)]}{\exp[\delta \cdot RL(t)] + \exp[\delta \cdot RR(t)]} = \frac{1}{1 + \exp[-\delta(1 - 2RL(t))]}, \quad (17)$$

where $\delta > 0$ is a dimensionless sensitivity parameter that, together with λ , determines the steepness of the psychometric curve. Parameter δ may be related to factors such as the stimulus modality, the distance between the two response keys, and other structural features of the situation; hence, as long as these features remain constant, δ should not change.

Equation 17 has the right set of properties: When the strength of the left response is much greater than the strength of the right response, $P(L|t)$ is close to 1; when the two responses have equal strength, $P(L|t)$ equals $1/2$. Hence, the bisection point predicted by the model does not depend on δ ; when the right response is much stronger than the left, $P(L|t)$ approaches 0. The equation has the additional advantage of generalizing easily to situations with more than two alternatives.

Predictions of the model. Computer simulations showed that the steady-state behavior of the model does not change appreciably when the extinction parameter α equals the reinforcement parameter β . Therefore, in what follows $\alpha = \beta$. This restriction has the additional advantage of leading to a steady-state distribution of W that is independent of the decision rule. In fact, it can be shown that the expected value of $\Delta WL(n)$ after one trial is equal to

$$E[\Delta WL(n)] = \alpha p X(S, n) - \alpha [p X(S, n) + (1 - p) X(L, n)] WL(n),$$

where p is the probability of presenting the short duration stimulus. At the steady state and for the common value of $p = 1/2$,

$$E[WL(n)] = \frac{X(S, n)}{X(S, n) + X(L, n)} = \frac{1}{1 + \left(\frac{L}{S}\right)^{\lambda} \exp[-\lambda(L - S)]}, \quad (18)$$

and

$$E[WR(n)] = 1 - E[WL(n)].$$

The final distribution of W depends on the ratio and the difference of the two stimulus durations, but it does not depend on the learning parameter α . This is an intuitively plausible result because the effects of reinforcement and extinction are likely to be identical for both responses and, consequently, their effects should cancel in a symmetric, two-choice situation. The final model is therefore composed of Equation 18, the steady-state value of the associative links; Equation 16, the strength of the two responses at the beginning of the choice period; and Equation 17, the probability of each response. It uses two free parameters: λ , included in X , and δ .

The solid curves in Figure 15 illustrate the predictions of the model for the three experiments reported before, and Table 5 gives the corresponding parameters. The theoretical curves describe all data sets well. In the bottom panel, δ and λ^* remained constant across conditions (see Table 5). I selected these parameter values because they were close to the best fitting parameters and, more important, because they predict a single curve when the probability of a short response is plotted against relative stimulus duration. In fact, Equation 18 implies that, for a constant ratio of stimulus durations, when λ is inversely proportional to the average duration of the signals, the steady-state values of W , RL , RR , and $P(L|t)$ do not change across conditions.

It is clear from Figure 15 that the model does not predict a bisection point exactly at the geometric mean of the training stimuli, but at a slightly greater value. As shown in Appendix D, this value, denoted by g^* , is approximately equal to

$$g^* = \frac{L - S}{\ln(L/S)}. \quad (19)$$

⁵ Computer simulations showed that the simpler ratio rule $P(L|t) = RL(t)/[RL(t) + RR(t)]$ could not account for the data presented above.

Table 5
Parameters Used to Fit the Data Sets Shown in Figure 15

Study	S	L	$\lambda \frac{S+L}{2}$	δ	ω^2
Catania (1970)	1	4	1.2	11.8	.999
Fetterman & Killeen (1991)	4	12	9.0	2.7	.993
Church & Deluty (1977)	1	4	8.2	2.5	.988
	2	8	8.2	2.5	.978
	3	12	8.2	2.5	.968
	4	16	8.2	2.5	.982

Note. S = duration of short stimulus; L = duration of long stimulus; λ = rate of transition across the behavioral states; δ = sensitivity parameter; ω^2 = variance accounted for.

The relative discrepancy between g^* and the geometric mean g is given by

$$\frac{g^* - g}{g} = \frac{r - 1}{\ln(r)\sqrt{r}} - 1, \quad (20)$$

where $r = L/S$. For small ratios, g^* is close to g , but as the ratio of signal durations increases so does the discrepancy. In the studies described earlier, r equaled 3 or 4, which means that the relative discrepancy ranged from .05 to .082, values that are too small to be easily detected. It can also be shown that g^* is always less than the arithmetic mean of the signals. Hence, the model predicts an indifference point that lies always between the geometric and the arithmetic means of the training stimuli, as illustrated by Catania (1970) and Fetterman and Killeen's (1991) data.

Dynamic analysis. Remarkably, perhaps, the bisection point of the model has the same equation as the local minimum of response rate in mixed FI-FI schedules (cf. Equations 15 and 19). This identity is more than coincidental and highlights a basic similarity between the two procedures. In fact, when the steady state is reached, the bisection experiment may be seen as a mixed FI-FI schedule, with the short and long stimuli playing the role of the two FIs. Furthermore, the middle state that in the mixed schedule is the least conditioned to the operant response is equivalent to the middle state that in the bisection experiment is equally conditioned to both responses; the former determines the local minimum of response rate, the latter the indifference point.

Church and Deluty (1977) also asked whether the rats in the bisection procedure learn to choose on the basis of the relative or the absolute durations of the stimuli. To answer this question, they performed the experiment schematized in Table 6. During Phase 1, all of the animals were trained to choose left after the short stimulus (T_1) and right after the long stimulus (T_2). In Phase 2, T_2 was presented with an even longer stimulus (T_3), so that the long stimulus of Phase 1 became the short stimulus in Phase 2. For one half of the rats (Group I), the left lever remained associated with the short stimulus, whereas for the other rats (Group II), the association between relative stimulus durations and levers was reversed. The results showed that Group I learned the new discrimination faster than Group II, which suggests that responding in Phase 1 was controlled by the relative, not the absolute, durations of the stimuli.

Table 6
Design of Church and Deluty's (1977) Experiment

Group	Phase 1		Phase 2	
	Stimulus	Correct response	Stimulus	Correct response
I: No reversal	T_1 (short)	Left	T_2 (short)	Left
	T_2 (long)	Right	T_3 (long)	Right
II: Reversal	T_1 (short)	Left	T_2 (short)	Right
	T_2 (long)	Right	T_3 (long)	Left

Note. T = duration of the stimulus.

To understand this finding within the current model, look at Figure 17. The filled circles show the strength of the associations between the behavioral states and the left response at the end of Phase 1, that is, $WL(n)$; the associations with the right response, that is, are not shown because they equal $1 - WL(n)$. The empty circles show the values of $WL(n)$ for Group I (left panel) and Group II (right panel) at the end of Phase 2. The difference between the filled and empty circles (see arrows) is the amount of reconditioning that will take place during Phase 2. Clearly, this difference is greater in the right panel, which means that when the experimenters reversed the assignment of the levers with respect to the relative durations of the stimuli, they left the animals in Group II farther from their steady state—hence, the greater amount of time required to learn the new discrimination. (The preceding argument is even more compelling if parameter λ decreases during Phase 2 because of a decrease in the overall reinforcement rate.)

Another way to understand the results of Phase 2 is to realize that, according to the model, Group I (no reversal) makes incorrect choices mainly after the short stimulus; the long stimulus is already correctly associated with the right response. However, Group II (reversal) makes incorrect choices mainly after the long stimulus; the short stimulus is correctly associated with the left response. But Table 4 shows that the magnitudes of ΔWL and ΔWR (i.e., the speed of learning) are proportional to X ,

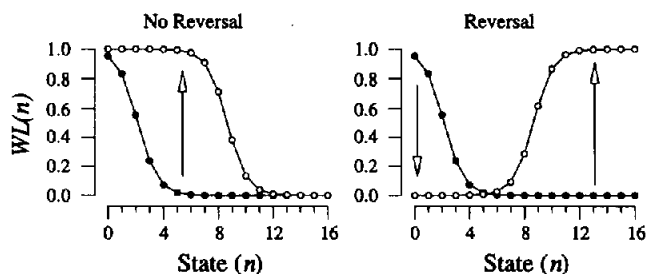


Figure 17. Predictions of the model for the experiment schematized in Table 6. The filled circles show the association of the states with the left response, $WL(n)$, at the end of Phase 1 or the beginning of Phase 2. The empty circles show the steady-state values of $WL(n)$ at the end of Phase 2. The arrows indicate the distance between the initial and final values of $WL(n)$ for both groups. The no-reversal group has to travel a shorter distance and should therefore learn the new discrimination faster.

the activation of the states at the beginning of the choice phase, and this activation is greater after short than long stimuli (see Figure 3).

Discussion

The temporal regulation of behavior is a derivative of three fundamental processes: the sequence of behavioral states, presumably of phylogenetic origin, that fills the interreinforcement interval; the process of reinforcement-extinction that modulates the association between the states and the operant response; and the motivational-incentive process that maps the degree of temporal regulation onto observable behavior. The theoretical model proposed in this study emphasizes mainly the second process, the much neglected learning component of temporal regulation. At the heart of the model is the belief that animals do not passively tell time by using central, all-purpose internal clocks; rather, they act on and interact with their environment, and in the process their behavior changes. In some environments, these changes yield behavior the properties of which lead human observers to talk about clocks. In other words, the behavioral clock evidenced in environments such as the peak procedure emerges from the interplay of the three basic processes mentioned earlier and the characteristic features of those environments.

The four procedures used to validate the model are a representative sample of the simplest environments in which behavior shows temporal regularities. The FI schedule and the peak procedure show how animals adjust their behavior to single intervals of time between reinforcers. In agreement with the data, the model predicts that response rate increases monotonically with time into the trial in the FI schedule and remains high if we allow the trial to continue past the usual reinforcement time. In the peak procedure, response rate increases up to the time of reinforcement but then it decreases because during the empty trials the later states extinguish their association with the operant response. The experiments with the mixed schedules show that animals can also adapt to two different temporal intervals. Again, the model reproduces some major features of the data, the dependence of the response rate profile on the duration and frequency of the short interval. Finally, the bisection procedure shows that behavior may also be controlled by the duration of past events. A discrete-time version of the model correctly predicts the shape of the psychometric function and yields an indifference point in close agreement with the empirical results. In sum, the model not only highlighted the strong family resemblance that exists between these four distinct environments, but also described the main features of the steady-state behavior observed in them.

In the area of temporal control, the large volume of existing steady-state data is comparable only to the scarcity of experimental and quantitative analyses of the corresponding acquisition processes. One likely reason for this asymmetry is the lack of theoretical frameworks sufficiently developed to guide research and organize empirical findings on the topic of temporal learning. Hence, the relevance of the examples of dynamic analyses presented. From acquisition, extinction, and impulse effects in FI schedules, the occasional tails of the response rate function in the peak procedure and the local minimum of responding in the mixed schedules, to the effects of reversing

the stimulus-response assignments in the bisection procedure, all these illustrative cases hint at the potential of the model to enhance our understanding of the dynamic aspects of temporal regulation.

Parameters of the Model

Each of the three basic processes assumed to underlie the temporal regulation of behavior depended on a single parameter. The rate of transition across the behavioral states depended on the rate parameter λ , the relative effects of extinction and reinforcement depended on the learning parameter γ , and the mapping from temporal regulation onto response rate depended on the scale parameter A . In the bisection experiment, I included the sensitivity parameter δ . The four parameters have straightforward interpretations: λ sets the maximum achievable degree of temporal regulation, the potential; γ translates the potential into an actual degree of temporal regulation (see Appendix B for the effects of λ and γ on the quality of temporal control); A determines absolute response rate; and δ determines the accuracy of the discrimination between two stimulus durations.

For the most part, the parameters of the model behaved orderly. Figure 18 presents the values of the parameters for all of the experiments reported earlier. Parameters λ^* and A are read on the left axis, and γ^* and δ are read on the right axis; both axes are logarithmic. Each set of connected points corresponds to the same experiment.

The rate parameter λ^* (filled circles) remained roughly constant across experimental conditions. Thus, BeT's traditional assumption that λ is proportional to the overall density of food (Bizo & White, 1994; Fetterman & Killeen, 1991; Killeen & Fetterman, 1988; Morgan, Killeen, & Fetterman, 1993) was observed in most fits. The behavior of the learning parameter γ (empty circles) was less straightforward. In FI schedules and the peak procedure, γ^* remained approximately constant across conditions, which agrees with the common observation that extinction becomes less effective as the average interval between reinforcers increases. In the mixed FI-FI schedules, however,

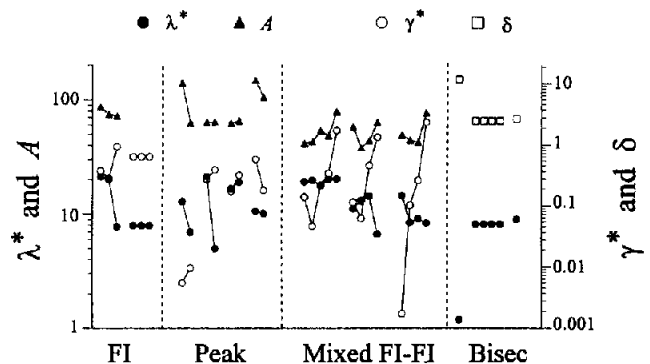


Figure 18. Parameter values used to fit the experiments reported in the text. Each set of connected points corresponds to one experiment. Note the logarithmic scale on the y-axes. The data for the fixed-interval (FI), peak procedure (Peak), mixed fixed-interval-fixed-interval (Mixed FI-FI), and the bisection (Bisec) segments come from Tables 1, 2, 3, and 5, respectively.

γ^* tended to increase as the variance of the interreinforcement intervals decreased (for each set of connected points in Figure 18, section Mixed FI-FI, the variance decreases from left to right). The overall behavior of γ^* is thus consistent with the partial-reinforcement extinction effect, the greater resistance to extinction after partial than continuous reinforcement (Mackintosh, 1974; but see Nevin, 1988) and with a discrimination theory of extinction (Skinner, 1950) that implies that of two schedules with the same average but different variances of interreinforcement intervals, the one with the greater variance will engender greater resistance to extinction.

In general, parameter A (filled triangles) did not vary substantially or consistently across experiments or conditions. And finally, the results from the bisection experiments do not allow any strong statements about the sensitivity parameter δ (empty squares). However, if my interpretation of δ is correct, then a new and testable prediction is possible: The steepness of the psychometric function should change whenever the discriminability of two training durations is varied, but these changes should not affect the animal's indifference point, because as I have shown, the indifference point predicted by the model does not depend on δ . Ways to manipulate the discriminability of the stimuli may include varying the distance between the response keys or increasing the response requirement during the choice phase. The latter happened in Catania's (1970) experiment, which may explain the high value for δ required to fit his results (see Figure 15 and Table 5).

Difficulties of the Model

Although the model fit well (i.e., $\omega^2 > .95$) 30 of the 33 data sets to which it was applied, it faced some difficulties. First, to deal with the scalar property, the model requires a direct relation between reinforcement rate and its two parameters, λ and γ . How problematic this restriction is remains to be seen. On the one hand, no study has quantified the dependency of γ on reinforcement density; on the other hand, the empirical evidence concerning λ is mixed. Several studies have reported the predicted changes in λ , but the changes have been less than proportional (e.g., Bizo & White, 1994; Fetterman & Killeen, 1991; Killeen & Fetterman, 1988; see also Leak & Gibbon, 1995, and Zeiler & Powell, 1994).

Second, to fit data from the peak procedure, I had to assume in one half of the cases that the effective interval from the time marker to the next reinforcer was slightly shorter than the real interval, an assumption also required to fit other data sets (e.g., Church et al., 1994). The required change was in the range of 2 to 4 s, which may reflect a short latency before the onset of the timing process. As also assumed by SET and BeT, a latency period could be due to inhibitory properties of the time marker or, more generally, to variations in its effectiveness to initiate the serial activation of the behavioral states. The fact that in the peak procedure the time marker is typically a stimulus change (e.g., the onset of a tone) rather than the reinforcer itself, gives some plausibility to this explanation.

Third, in the bisection procedure, the model predicted indifference at a stimulus duration slightly greater than the geometric mean (cf. Zeiler, 1985). Again, the empirical evidence reviewed in my study is not conclusive on this issue because the prediction

of the model and the geometric mean are too close to be easily detected with current procedures.

Fourth, the model faced its greatest difficulty in one condition of Catania and Reynolds's (1968) study, namely, a mixed FI 30-s-FI 240-s schedule with a probability of .5 for the short FI (see Figures 11 and 12). Recent results reported by Leak and Gibbon (1995), using the same procedure, suggest that when the probability of the short interval is high, animals behave as if they time the two intervals independently and simultaneously. This interpretation is based on the peaks and the widths of the response rate curve during the long interval: The average response rate curve rises quickly and peaks at the short interval, falls to a smaller value, and then increases gradually until the end of the trial. Also, in contrast with Catania and Reynolds's data, the response rate at the short interval is much higher than the terminal rate.

These results challenge both serial and parallel models but in different ways. Thus, to fit the initial, narrow peak, a serial model such as the present one (or BeT) has to assume a high value for the rate parameter λ , a value that then fails to predict the gradual increase in response rate observed subsequently. Conversely, to fit the later portions of the response profile, the model has to assume a small λ that does not fit the narrow, initial segment of the response curve. If time regulation is a serial process, then the preceding results seem to demand changes in the parameter λ within one interreinforcement interval and changes in λ that are under the control of the animal's own behavior. The latter hypothesis echoes Ferster and Skinner's (1957) claim that in many reinforcement schedules the animal's current behavior becomes a discriminative stimulus for subsequent behavior (see also Skinner, 1950). Only further research can decide whether these new assumptions about λ constitute spurious epicycles designed to rescue fundamentally flawed serial theories, or alternatively, whether they present a more realistic portrait of the complexities of temporal regulation.

On the other hand, if we assume two parallel and independent processes of temporal regulation, one estimating the short and the other the long interval, then we need to answer a different set of questions. First, how do the two processes determine response rate? Second, if response rate is a function of the relative discrepancy between the current time estimate (t_i) and a reference time to food (T_1 for the short or T_2 for the long interval), as SET typically assumes, then at all times t_1 and t_2 in which $(t_1 - T_1)/T_1 = (t_2 - T_2)/T_2$, response rates should be equal. This equality certainly holds at the peaks (i.e., $t = T_1$ and $t = T_2$), but at those times response rates are clearly different. Third, and more important from a dynamic viewpoint, if the memory stores containing the two sets of estimates of T_1 and T_2 are independent, then how does the animal determine the store in which to save a just experienced reinforced time? (To say that large values go into one store and small values go into the other simply begs the question.) Fourth, Innis and Stadon (1971) showed that pigeons can track complex FI schedules in which the successive interreinforcement intervals follow a sinusoidal pattern. Does it follow that the birds formed independent estimates of the large number of FIs involved in these cyclic schedules?

In a strictly Popperian view of science, progress is achieved by refuting conjectures (e.g., Popper, 1972). Accordingly, the difficulties identified earlier could be seen as decisive and the

serial model should then be abandoned. But if we adopt a different viewpoint toward the role of models in science, then an alternative is possible. Rather than hypotheses to be only rejected, dynamic models may also help to organize empirical findings, to suggest similarities where they are not apparent, and to identify clearly the strengths and limitations of current approaches. The present model satisfied these purposes. Although initially conceived to deal with performance during FI schedules, the model accounted reasonably well for the major results obtained with other procedures. More to the point, its predictions for one procedure helped us to understand the behavior in a different procedure. Thus, performance in the peak procedure was related to performance in simple FI schedules, performance in mixed FI-FI schedules was derived from the combination of performance in single FI schedules and the peak procedure, and performance in the bisection procedure was related to performance in mixed FI-FI schedules. Hence, behavioral phenomena that looked very different were organized into a common and coherent framework. Furthermore, the model also revealed clearly how little is currently known about the parameters that regulate the learning processes, response rate, and response probability.

General Theory of Timing Behavior

Church and Broadbent (1992) summarized some of the challenges facing a general theory of timing behavior thus:

The first fact that any theory of temporal generalization must account for is the smooth peak function in which the mean probability of a response gradually increases to a maximum near the time of reinforcement and then decreases in a slightly asymmetrical fashion.

The second fact that any theory of temporal generalization must account for is that performance on individual trials, unlike the mean function, is characterized by an abrupt change from a state of low responding to a state of high responding and finally another state of low responding.

The third fact that any theory of temporal generalization must account for is that the mean functions are very similar with time shown as a proportion of the time of reinforcement. (p. 58)

As the predictions for the peak procedure show (Figure 9), the present model clearly satisfies the first requirement. At the steady state, response rate increases smoothly up to the time of reinforcement and then decreases in an asymmetrical way. The model satisfies the third requirement whenever the rate parameter λ and the learning parameter λ are directly proportional to overall reinforcement rate. The second requirement falls outside the scope of the present model because, being both deterministic and linear, the model restricts its domain to the average behavior of the animal. Two reasons mitigate the potential negative impact of this restriction, however. First, the abrupt change in behavior mentioned by Church and Broadbent (1992) occurs only after prolonged exposure to time-based schedules; early in training, response rate changes more gradually, in closer agreement with the present model (e.g., Ferster & Skinner, 1957; Schneider, 1969). Second, as I will explain, the formal structure of the model and its equations are easily cast in a stochastic framework, a framework that may be able to deal with the forms of nonlin-

earity identified by Church and Broadbent. The stochastic framework also captures some of the verbal statements made by the proponents of BeT (see Killeen & Fetterman, 1988).

Specifically, the cascade of behavioral states assumed by the present model is mathematically equivalent to a Poisson process, the cornerstone of BeT. In other words, if a pulse from an internal pacemaker, occurring randomly in time, moves the animal from state n to state $n + 1$, then the differential equations for the probability of being in state n at time t equal Equations 1 to 4. Also, if we assume that the association between the animal's current state and the operant response decreases during extinction and increases during reinforcement, then Equations 5 and 6 could describe the conditioning process; variable $X(t, n)$ would be set to 1 if the animal is in state n at time t , and to 0 otherwise. Finally, the associative strength of a state could be interpreted as the tendency to respond when in that state, a tendency mapped onto response rate by the multiplier A . Hence, the behavior of the present model is formally equivalent to the mean performance predicted by BeT with a linear learning rule and a constant scaling factor to generate response rate. The adequacy and full implications of this stochastic framework⁶ (e.g., its predictions concerning trial-by-trial variability in pause duration; see Hanson & Killeen, 1981; Schneider, 1969) remain to be investigated.

As the foregoing remarks show, the present model bears a close resemblance to BeT. In fact, it instantiates BeT as a dynamic, real-time model of temporal regulation. By providing BeT with an explicit learning rule and an explicit mapping from temporal regulation onto response rate, or probability, the model eliminates BeT's most dubious assumptions. In all single-response schedules, it is no longer necessary to assume, somewhat ad hoc, that only one state or a subset of states becomes conditioned to the operant response. Similarly, in experiments on the bisection of temporal intervals, it is no longer necessary to assume that only two states control the animal's choices (e.g., Bizo & White, 1994). In fact, one wonders what learning rules would lead to such extreme forms of conditioning.

In addition to its extensions of BeT, the present model also suggests that Church and Broadbent's (1992) list of challenges to a general theory of timing is incomplete; at least a fourth one is needed. A general theory of timing must also explain the behavioral trajectory, or trajectories, followed by animals as they approach the steady state. This general challenge boils down to specific questions: How does learning proceed when important events in the life of an animal occur periodically or quasi-periodically? How do animals solve the credit-assignment problem in time-based reinforcement schedules? How do the processes of reinforcement and extinction affect temporal regulation? Having identified a set of critical, albeit currently dormant, questions central to our understanding of timing behavior, and having advanced plausible answers to them, the model described in this article has also fulfilled its heuristic role.

⁶ Another way to deal with the variability in the temporal control of behavior is to assume that the two major parameters of the model are random variables with means λ and γ and some finite variance. This assumption would open the door to an important source of multiplicative variance in the model, one that, according to Gibbon's (1992) analysis, would rapidly swamp the Poisson variance engendered by a constant λ .

References

- Bizo, L. A., & White, K. G. (1994). The behavioral theory of timing: Reinforcer rate determines pacemaker rate. *Journal of the Experimental Analysis of Behavior*, 61, 19–33.
- Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. New York: Wiley.
- Catania, A. C. (1970). Reinforcement schedules and the psychophysical judgments: A study of some temporal properties of behavior. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 1–42). New York: Appleton-Century-Crofts.
- Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 11, 327–383.
- Cheng, K., & Roberts, W. A. (1991). Three psychophysical principles of timing in pigeons. *Learning and Motivation*, 22, 112–128.
- Church, R. M., & Broadbent, H. A. (1992). Alternative representations of time, number, and rate. In C. R. Gallistel (Ed.), *Animal cognition* (pp. 55–81). Cambridge, MA: MIT Press.
- Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes*, 3, 216–228.
- 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.
- Davis, D. G. S., Staddon, J. E. R., Machado, A., & Palmer, R. G. (1993). The process of recurrent choice. *Psychological Review*, 100, 320–341.
- 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.
- Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, 57, 94–107.
- Falk, J. L. (1971). The nature and determinants of adjunctive behavior. *Physiology and Behavior*, 6, 577–588.
- Ferster, C., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
- Fetterman, J. G., & Killeen, P. R. (1991). Adjusting the pacemaker. *Learning and Motivation*, 22, 226–252.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84, 279–325.
- Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology*, 24, 58–87.
- Gibbon, J. (1991). Origins of scalar timing theory. *Learning and Motivation*, 22, 3–38.
- Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical Psychology*, 36, 283–293.
- Gibbon, J., & Allan, L. (Eds.). (1984). *Annals of the New York Academy of Sciences: Vol. 423. Timing and time perception*. New York: New York Academy of Sciences.
- Grossberg, S., & Schmajuk, N. (1989). Neural dynamics of adaptive timing and temporal discrimination during associative learning. *Neural Networks*, 2, 79–102.
- Hanson, S. J., & Killeen, P. R. (1981). Measurement and modeling of behavior under fixed-interval schedules of reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 129–139.
- Higa, J. J., Wynne, C. D. L., & Staddon, J. E. R. (1991). Dynamics of time discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 281–291.
- Horner, J. M., & Staddon, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioral Processes*, 15, 59–92.
- Innis, N. K., & Staddon, J. E. R. (1971). Temporal tracking on cyclic-interval reinforcement schedules. *Journal of the Experimental Analysis of Behavior*, 16, 411–423.
- Killeen, P. R. (1991). Behavior's time. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 295–334). New York: Academic Press.
- Killeen, P. R. (1994). Mathematical principles of reinforcement. *Behavioral and Brain Sciences*, 17, 105–172.
- Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review*, 95, 274–285.
- Killeen, P. R., Hanson, S. J., & Osborne, S. R. (1978). Arousal: Its genesis and manifestation as response rate. *Psychological Review*, 85, 571–581.
- Lea, S. E. G., & Dow, S. M. (1984). The integration of reinforcement over time. *Annals of the New York Academy of Sciences*, 423, 269–277.
- Leak, T. M., & Gibbon, J. (1995). Simultaneous timing of multiple intervals: Implications of the scalar property. *Journal of Experimental Psychology: Animal Behavior Processes*, 21, 3–19.
- Mackintosh, N. J. (1974). *The psychology of animal learning*. London: Academic Press.
- Matthews, T. J., & Lerer, B. E. (1987). Behavior patterns in pigeons during autoshaping with an incremental conditioned stimulus. *Animal Learning & Behavior*, 15, 69–75.
- Morgan, L., Killeen, P. R., & Fetterman, J. G. (1993). Changing rates of reinforcement perturbs the flow of time. *Behavioral Processes*, 30, 259–272.
- Myerson, J., & Miezian, F. M. (1980). The kinetics of choice: An operant systems analysis. *Psychological Review*, 87, 160–174.
- Nevin, J. A. (1988). Behavioral momentum and the partial reinforcement effect. *Psychological Bulletin*, 103, 44–56.
- Platt, J. R. (1979). Temporal differentiation and the psychophysics of time. In M. D. Zeiler & P. Harzem (Eds.), *Advances in the analysis of behavior: Vol. 1. Reinforcement and the organization of behavior* (pp. 1–29). New York: Wiley.
- Platt, J. R., & Davis, E. R. (1983). Bisection of temporal intervals by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 9, 160–170.
- Popper, K. R. (1972). *Objective knowledge*. New York: Oxford University Press.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in C: The art of scientific computing* (2nd ed.). New York: Cambridge University Press.
- Restle, F., & Greeno, J. G. (1970). *Introduction to mathematical psychology*. London: Addison-Wesley.
- Richelle, M., & Lejeune, H. (1980). *Time in animal behavior*. Oxford, England: Pergamon Press.
- Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 242–268.
- Roberts, W. A., Cheng, K., & Cohen, J. S. (1989). Timing light and tone signals in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 15, 23–35.
- Schneider, B. (1969). A two-state analysis of fixed-interval responding in the pigeon. *Journal of the Experimental Analysis of Behavior*, 12, 667–687.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, 57, 193–216.
- Staddon, J. E. R. (1977). Schedule-induced behavior. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 125–152). Englewood Cliffs, NJ: Prentice-Hall.
- Staddon, J. E. R. (1983). *Adaptive behavior and learning*. Cambridge, England: Cambridge University Press.
- Staddon, J. E. R., & Simmelhag, V. L. (1971). The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review*, 78, 3–43.

- Stubbs, D. A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior*, 11, 223–238.
- Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychonomic Bulletin & Review*, 1, 405–420.
- Timberlake, W., & Lucas, G. A. (1985). The basis of superstitious behavior: Response contingency, stimulus substitution, or appetitive behavior? *Journal of the Experimental Analysis of Behavior*, 44, 279–299.
- Timberlake, W., & Silva, K. M. (1995). Appetitive behavior in ethology, psychology, and behavior systems. In N. S. Thompson (Ed.), *Perspectives in ethology: Vol. 11. Behavioral design* (pp. 211–253). New York: Plenum Press.
- Zeiler, M. D. (1985). Pure timing in temporal differentiation. *Journal of the Experimental Analysis of Behavior*, 43, 183–193.
- Zeiler, M. D., & Powell, D. G. (1994). Temporal control in fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 61, 1–9.

Appendix A

Fixed Interval Schedules

The following functions are used to simplify the notation in subsequent equations:

$$X(t, n) = \frac{\exp(-\lambda t)(\lambda t)^n}{n!},$$

$$f(T, n) = \exp[-\beta d X(T, n)],$$

$$g(T, n) = \exp\left[-\alpha \int_0^T X(\tau, n) d\tau\right],$$

where d is reinforcement duration, α and β are learning parameters, and λ is the rate parameter that controls the activation of the behavioral states. The following restrictions apply: $0 < \alpha, \beta, \lambda$, and $d = 3$ s.

To derive the steady-state distribution of W for FI schedules, I divided the FI cycle into its two segments, the first T s of extinction and the next d s of reinforcement. The reinforced response separates the two segments. Next, I considered the m th FI cycle and related the three values of $W(t, n)$ illustrated in the following diagram:

$$W(0, n)^{(m)} \xrightarrow{T \text{ s of extinction}} W(T, n)^{(m)} \xrightarrow{d \text{ s of reinforcement}} W(T + d, n)^{(m)}.$$

$W(0, n)^{(m)}$ is the associative strength of state n at the beginning of the m th cycle; similarly, $W(T, n)^{(m)}$ and $W(T + d, n)^{(m)}$ are the associations of state n after the T s of extinction and the d s of reinforcement, respectively. From Equations 9 and 10 (with $K = 1$) and remembering the earlier definitions of f and g , we get

$$W(T, n)^{(m)} = W(0, n)^{(m)} g(T, n),$$

$$W(T + d, n)^{(m)} = 1 - [1 - W(0, n)^{(m)} g(T, n)] f(T, n).$$

But $W(T + d, n)^{(m)} = W(0, n)^{(m+1)}$ because the associative strength of state n at the end of cycle m equals the associative strength of state n at the beginning of cycle $m + 1$. Hence, a linear difference equation relates two successive values of $W(0, n)$:

$$W(0, n)^{(m+1)} = 1 - f(T, n) + f(T, n) g(T, n) W(0, n)^{(m)}.$$

At the steady state,

$$W(0, n)^{(\infty)} = \frac{1 - f(T, n)}{1 - f(T, n) g(T, n)}$$

$$= \frac{1 - \exp(-\beta d X(T, n))}{1 - \exp\left[-\alpha \int_0^T X(\tau, n) d\tau - \beta d X(T, n)\right]}.$$

When the learning parameters α and β are small, the above expression

may usefully be simplified by expanding the exponentials and retaining only first-order terms. After rearranging,

$$W(0, n)^{(\infty)} \approx \frac{X(T, n)}{X(T, n) + \gamma \int_0^T X(\tau, n) d\tau}, \text{ where } \gamma = \frac{\alpha}{\beta d}.$$

This is Equation 11 in the text.

Limit of $W(n)$ for Large n

To determine the value of W for large n , an important result to predict response rate early in extinction after FI training, we divide numerator and denominator by $X(T, n)$ and then analyze the ratio

$$\frac{\int_0^T X(\tau, n) d\tau}{X(T, n)} = \frac{\exp(\lambda T)}{T^n} \int_0^T \exp(-\lambda \tau) \tau^n d\tau$$

$$< \frac{\exp(\lambda T)}{T^n} \int_0^T \tau^n d\tau = \frac{k}{n+1}.$$

As n grows, this ratio tends to 0 and consequently $W(n)$ approaches 1.

The Scalar Property

Equation 12 in the text is equivalent to

$$R(t) = A \sum_n \frac{\exp(-\lambda t)(\lambda t)^n / n!}{1 + \gamma \int_0^t \exp[-\lambda(\tau - T)] \left(\frac{\tau}{T}\right)^n d\tau}.$$

By the change of variable $z = (\tau/T)$, we get

$$= A \sum_n \frac{\exp(-\lambda t)(\lambda t)^n / n!}{1 + \gamma T \int_0^1 \exp(-\lambda T(z - 1)) z^n dz}.$$

If $\lambda = k_1/T$ and $\gamma = k_2/T$, then $R(t)$ depends only on the ratio $t/T = t^*$,

$$= A \sum_n \frac{\exp(-k_1 t^*)(k_1 t^*)^n / n!}{1 + k_2 \int_0^1 \exp(-k_1(z - 1)) z^n dz}.$$

The normalization along the Y axis cancels the effect of parameter A . Hence, response rate during the interval depends exclusively on t^* .

Appendix B

The Peak Procedure

In the peak procedure, $W(n)$ is a random variable the value of which at the end of one cycle depends on the type of trial—food or empty—in effect during the cycle. Hence, I approximate the distribution of $W(n)$ by its expected value (computer simulations show that the approximation is extremely good when the learning parameters are small). With probability $1 - p$, the trial is empty and the value of $W(0, n)$ decays for T_2 s; with probability p , the trial is a food trial and, therefore, after T_1 s of extinction, d s of reinforcement follow. Hence, the expected value of $W(0, n)$ on trial $m + 1$ equals

$$W(0, n)^{(m+1)} = p \underbrace{[1 - (1 - W(0, n)^{(m)})g(T_1, n))f(T_1, n)]}_{\text{food trial}} + (1 - p) \underbrace{W(0, n)^{(m)}g(T_2, n)}_{\text{empty trial}}$$

Hence,

$$W(0, n)^{(m+1)} = p[1 - f(T_1, n)] + [pf(T_1, n)g(T_1, n) + (1 - p)g(T_2, n)]W(0, n)^{(m)}$$

Let

$$a(n) = p[1 - f(T_1, n)]$$

and

$$b(n) = [pf(T_1, n)g(T_1, n) + (1 - p)g(T_2, n)]$$

Then,

$$\begin{aligned} W(0, n)^{(m)} &= a(n) \sum_{i=0}^{m-1} [b(n)]^i + [b(n)]^m W(0, n)^{(0)} \\ &= \frac{a(n)}{1 - b(n)} + \left[W(0, n)^{(0)} - \frac{a(n)}{1 - b(n)} \right] [b(n)]^m \end{aligned}$$

The last equation with $m = 600$ was used to predict the non-steady-state data from Roberts et al.'s (1989) study, condition $T = 30$ s. The fit is shown in the left bottom panel of Figure 9 (see solid curve through empty circles).

In the steady state,

$$W(0, n)^{(\infty)} = \frac{p[1 - f(T_1, n)]}{1 - pf(T_1, n)g(T_1, n) - (1 - p)g(T_2, n)}$$

Again, by expanding the exponentials and retaining only first-order terms, the preceding expression simplifies to

$$W(0, n)^{(\infty)} \approx \frac{X(T_1, n)}{X(T_1, n) + \gamma \int_0^{T_1} X(\tau, n) d\tau + \gamma \frac{1-p}{p} \int_0^{T_2} X(\tau, n) d\tau}$$

This is Equation 13 in the text.

The Effects of λ and γ on the Quality of Temporal Control

To illustrate the effects of the rate and learning parameters on the quality of temporal control, assume a peak procedure with 20-s food trials, 160-s empty trials, and $p = .8$. The top panel of Figure B1 shows the effect of varying λ while keeping γ constant. As λ decreases, the response rate curves broaden and the degree of temporal control is reduced—small λ implies greater variance. The bottom panel shows the effect of holding λ constant and varying γ . When the relative effects of extinction decrease, two consequences follow: (a) Response strength increases, and (b) the rate curves broaden and, therefore, temporal con-

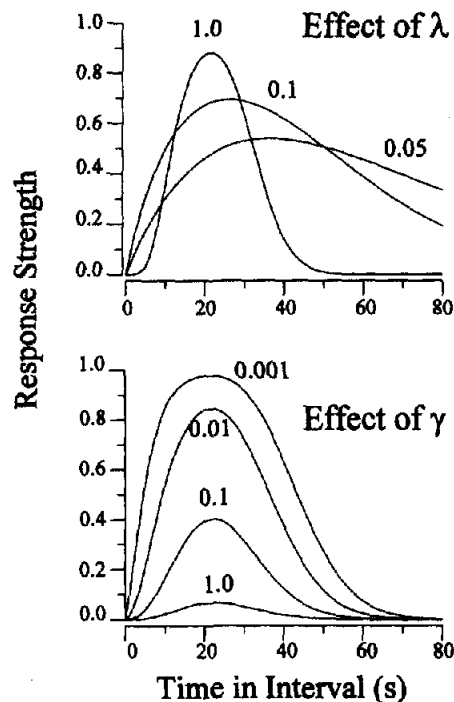


Figure B1. The effect of the rate (λ) and learning (γ) parameters on the shape of the response curves in the peak procedure. Top: For constant γ , increasing λ increases the degree of temporal discrimination (variance is reduced). Bottom: For constant λ , decreasing γ increases response strength but also decreases the degree of temporal discrimination (variance increases).

rol is reduced. In the extreme case of $\gamma = 0$ (i.e., extinction has no effect), the animal would respond throughout the interval and show no temporal control.

Notice that if the learning parameter γ can increase the variance of a response rate curve, it cannot decrease that variance below the minimum value imposed by parameter λ . In other words, the rate parameter λ sets the maximum degree of temporal control (i.e., the lowest variance); changes in γ can only decrease that control (i.e., increase the variance).

Limit of $W(n)$ for large n

In contrast with the FI schedule, in the peak procedure the limit of $W(n)$ for large n is 0. To see this, divide both numerator and denominator of the preceding equation by $X(T_1, n)$, and then consider the ratio

$$\begin{aligned} \frac{\int_0^{T_1} X(\tau, n) d\tau}{X(T_1, n)} &= \frac{\exp(\lambda T_1)}{(T_1)^n} \int_0^{T_1} \exp(-\lambda \tau) \tau^n d\tau \\ &> \frac{\exp(\lambda T_1)}{(T_1)^n} \exp(-\lambda T_2) \int_0^{T_2} \tau^n d\tau = k \frac{\left(\frac{T_2}{T_1}\right)^{n+1}}{n+1} \end{aligned}$$

As n grows, this ratio tends to $+\infty$ and therefore $W(n)$ tends to 0.

(Appendixes continue on next page)

Appendix C

Mixed Schedules

To derive the predictions of the model for mixed FI-FI schedules, I followed the same steps as for the peak procedure, except that here the second interval ends with food. Hence, the expected value of $W(0, n)$ at the beginning of cycle $(m+1)$ equals

$$W(0, n)^{(m+1)} = p \underbrace{[1 - (1 - W(0, n)^{(m)})g(T_1, n))f(T_1, n)]}_{\text{FI } T_1 \text{ schedule}} + (1-p) \underbrace{[1 - (1 - W(0, n)^{(m)})g(T_2, n))f(T_2, n)]}_{\text{FI } T_2 \text{ schedule}},$$

where T_1 and T_2 are the durations of the two FIs and p is the probability of selecting the FI T_1 s long. At the steady state,

$$W(0, n)^{(\infty)} = [1 - pf(T_1, n) - (1-p)f(T_2, n)] / [1 - pf(T_1, n)g(T_1, n) - (1-p)f(T_2, n)g(T_2, n)]$$

Expanding the exponentials and retaining only first-order terms yields

$$W(0, n)^{(\infty)} \approx [pX(T_1, n) + (1-p)X(T_2, n)] / \left\{ p \left[X(T_1, n) + \gamma \int_0^{T_1} X(\tau, n) d\tau \right] + (1-p) \left[X(T_2, n) + \gamma \int_0^{T_2} X(\tau, n) d\tau \right] \right\}.$$

This is Equation 14 in the text. Note that when $p = 0$ or $p = 1$, the preceding equation reduces to the equation for an FI- T_1 s, or an FI- T_2 s, respectively.

The Local Minimum

The local minimum of response rate may be approximated by determining first the most active state when the curves for the peak procedure and the FI schedule intersect (see Figure 14):

$$\frac{X(T_1, n)}{X(T_1, n) + \gamma \int_0^{T_1} X(\tau, n) d\tau + \gamma \frac{1-p}{p} \int_0^{T_2} X(\tau, n) d\tau} = \frac{X(T_2, n)}{X(T_2, n) + \gamma \int_0^{T_2} X(\tau, n) d\tau},$$

which simplifies to

$$\exp(\lambda s)r^{-n} = \frac{\int_0^{\lambda T_1} \exp(-\tau)\tau^n d\tau}{\int_0^{\lambda T_2} \exp(-\tau)\tau^n d\tau} + \frac{1-p}{p},$$

where $s = T_2 - T_1$ and $r = T_2/T_1$. The ratio of the two integrals is always between 0 and 1, and therefore

$$\frac{\lambda s}{\ln(r)} + \frac{\ln(p)}{\ln(r)} \leq n \leq \frac{\lambda s}{\ln(r)} - \frac{\ln\left(\frac{1-p}{p}\right)}{\ln(r)}.$$

When $p = 1/2$, as in the bottom panel of Figure 14, the above expression simplifies to

$$\frac{\lambda s}{\ln(r)} - \frac{\ln(2)}{\ln(r)} \leq n \leq \frac{\lambda s}{\ln(r)}.$$

The second term on the left is typically much smaller than 1 because r is much greater than 2. Therefore, and to an approximation,

$$n \approx \frac{\lambda s}{\ln(r)}.$$

The time t^* when state n is the most active state equals

$$t^* = \frac{n}{\lambda} \approx \frac{s}{\ln(r)}.$$

This is Equation 15 in the text. In summary, the minimum of response rate when $p = 1/2$ and $r \gg 2$ is basically independent from the parameters of the model. The foregoing conclusion is valid only if λ is not too small (i.e., $\lambda > 1/T_1$), but this condition is not too restrictive because when λ is very small, there simply is no local minimum of response rate.

Appendix D

Bisection Procedure

The exact value of the bisection point predicted by the model can only be computed numerically, but the following argument provides a close approximation. After a short duration signal, the early states are the most active, and because they are strongly conditioned to the left response, that response occurs. Similarly, after a long duration signal, the later states are the most active, and because they are mainly associated with the right response, that response occurs. Figure D1 illustrates these remarks for the case $S = 1$ s, $L = 4$ s, and $\lambda = 1.44$. The filled and empty circles show the distributions of WL and WR , respectively. As expected, early states are mostly conditioned to the left response, whereas later states are mostly conditioned to the right response. The point at which the two curves intersect corresponds to a state that is approximately equally conditioned to the left and right responses. When this state and its neighbors are the most active states at the moment of choice, the animal is indifferent.

The two curves in Figure D1 intersect when

$$E[WL(n)] = E[WR(n)]$$

$$X(S, n) = X(L, n)$$

$$n = \lambda \frac{L - S}{\ln\left(\frac{L}{S}\right)}.$$

This result was also obtained by Killeen and Fetterman (1988) following a different approach. State n will be the most active state for a stimulus duration equal to

$$g^* = \frac{n}{\lambda} = \frac{L - S}{\ln\left(\frac{L}{S}\right)}.$$

This is Equation 19 in the text. To compare g^* with the geometric

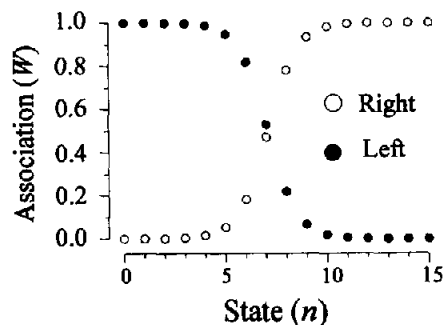


Figure D1. Predicted associative connections of the behavioral states to the left and right responses when the short and long stimuli have durations of 1 and 4 s, respectively, and $\lambda = 1.44$. State 7, almost equally associated with both responses, predicts the bisection point of the model.

mean g , we express both values in terms of the difference between the signal durations ($s = L - S > 0$) and their ratio ($r = L/S > 1$):

$$g^* = \frac{s}{\ln(r)}$$

$$g = \frac{s\sqrt{r}}{r - 1}.$$

The relative discrepancy between g^* and g is then given by

$$\frac{g^* - g}{g} = \frac{r - 1}{\ln(r)\sqrt{r}} - 1.$$

This is Equation 20 in the text.

Appendix E

Parameter Estimation

The parameters of the model were estimated by the nonlinear simplex algorithm as described by Press, Teukolsky, Vetterling, and Flannery (1992). For 4 of the 33 data sets, the algorithm converged to unreasonable parameter values (e.g., extremely high values of A or $\gamma \leq 0$, or both). In these cases, I imposed restrictions on the parameters on the

basis of visual inspection of the data. The variance accounted for with the new parameters was reduced by less than .01.

Received January 12, 1996

Revision received July 11, 1996

Accepted July 23, 1996 ■