

# Mathematical principles of reinforcement

Peter R. Killeen

Department of Psychology, Arizona State University, Tempe, AZ  
85287-1104

Electronic mail: [killeen@asu.edu](mailto:killeen@asu.edu)

**Abstract:** Effective conditioning requires a correlation between the experimenter's definition of a response and an organism's, but an animal's perception of its behavior differs from ours. These experiments explore various definitions of the response, using the slopes of learning curves to infer which comes closest to the organism's definition. The resulting exponentially weighted moving average provides a model of memory that is used to ground a quantitative theory of reinforcement. The theory assumes that incentives excite behavior and focus the excitement on responses that are contemporaneous in memory. The correlation between the organism's memory and the behavior measured by the experimenter is given by coupling coefficients, which are derived for various schedules of reinforcement. The coupling coefficients for simple schedules may be concatenated to predict the effects of complex schedules. The coefficients are inserted into a generic model of arousal and temporal constraint to predict response rates under any scheduling arrangement. The theory posits a response-indexed decay of memory, not a time-indexed one. It requires that incentives displace memory for the responses that occur before them, and may truncate the representation of the response that brings them about. As a contiguity-weighted correlation model, it bridges opposing views of the reinforcement process. By placing the short-term memory of behavior in so central a role, it provides a behavioral account of a key cognitive process.

**Keywords:** activation; contiguity; contingency; coupling; mechanics; memory; reinforcement; response rate; schedules; trajectories; tuning curves

What does reinforcement strengthen? Responses, to be sure, but responses as we define them or as the animal's nervous system defines them? Beginning with his earliest articles, Skinner (1935) cogently argued that stimuli, responses, and reinforcers – the constituents of the *operant*, and the key terms in an experimental analysis of behavior – must be defined functionally, and thus interdependently. In his experiments, however, he immersed subjects in arbitrary lights and sounds, levers and keys, snacks and sips, confident that the operant would "select" itself out of the stream of events in time. Modern behavior analysis succeeds to the extent that our definitions and instrumentations of these key terms have themselves been selected to respect the animals' definitions, as best we can intuit them (Bolles 1983; Timberlake & Lucas 1990; Zeiler & Thompson 1986). In the extralaboratory setting we have less control of all three terms and make correspondingly greater use of hedges such as "the stimulus as coded," "response class," and "motivational effects." These qualifications, even when apposite, are usually ad hoc. In this target article I derive a functional definition of the response. I then demonstrate the wide-ranging implications of the appropriate use of this, the organism's definition, for it directly gives shape to many laws of behavior.

Reinforcement controls aspects of the response along multiple dimensions, such as locus, force, tempo, and topography. Experimental contingencies – the conditions set by the experimenter as criteria for reinforcement – may shape these aspects either in concert with or in opposition to the fundamental excitatory property of

reinforcement (Zeiler 1979). The present research disentangles these forces with an experimental analysis of the response. With that characterized, it then derives models of how the response is coupled to reinforcement by experimental contingencies. These, in concert with a general model of arousal-elicitation under time constraints, yield a general theory of reinforcement.

## 1. The spread of effect

A reinforcer's effects are not limited just to the response that immediately preceded it. Response rates may be readily increased or decreased by requiring the pause before the last response to exceed some criterion. It follows that the reinforcement must reach back toward the penultimate response. In fact, the contingencies have been shown to extend even further (Catania 1971; Killeen 1969), but with diminishing effectiveness. The decrease in effectiveness with temporal distance is a manifestation of the *delay of reinforcement gradient*, a concept that has played a key role in theoretical accounts of behavior over the years. The delay may be "bridged" by conditioned reinforcers, but without them the decay of strength with time is usually swift.

For operant responses, measurement of the gradient presents paradoxical technical problems (Catania & Keller 1981). Let us suppose we wish to study the effects of delaying reinforcement by two seconds. A response occurs, we start our delay timer – but then another response occurs. If we restart the timer, we are setting up a new

contingency that punishes responses. This will deter animals from making the target response, but the deterrence speaks less to the reduced strength of a 2-sec delay than to the establishment of the target response as a predictor of nonreinforcement, and thus as something to be avoided. We have shaped nonresponding.

Alternatively, we may ignore the second response. Although that is often done (Sizemore & Lattal 1978; Williams 1976), we can no longer assume we are measuring the reinforcer's effects on a response 2-sec distant when another response from the same class has intervened closer to reinforcement. As a third possibility, we might simply prohibit the second response by removing the lever or turning off the key light. But then we have set up a conditioned reinforcer that will directly strengthen the target response (e.g., Richards 1981).

Even if these problems were somehow to be solved, we must face the complementary, if less obvious, problem of what to do about responses that *precede* the target response. Unless the target response "absorbs" all the strength of reinforcement (Williams [1975] shows that it absorbs some), the presence of immediately prior responses that are also being strengthened by the only-slightly-more delayed reinforcer must be taken into account.

Consider an alternative approach: assume that reinforcement increases the probability of those events that are in a subject's recent memory. Then, to the extent that reinforcement is contingent on just those events in memory, and only those, it should be maximally effective in strengthening behavior: conditioning will proceed most rapidly when our definition of the response matches the organism's definition. If we wish to reinforce response patterns in an organism that can remember exactly four responses, making reinforcement contingent on only three correct responses would squander its potential effectiveness on the first, randomly providing reward for both correct and incorrect instances of it. Conversely, making reinforcement contingent on six correct responses would introduce intermittency of reward that had no bearing on the shaping of the response, because the first two responses would be gone from memory and their control of the delivery of reinforcement would not be reciprocated in the reinforcement's control of them.

This, then, is the strategy of the following experiments: we will change the contingencies for reinforcement over a continuum, and where there is the greatest impact on responding, we will infer that our definition of the (extended) response most closely matches the animal's memory of it (i.e., we choose the model that maximizes the posterior probability of the data given that model). In the process, we will learn the extent of the delay of reinforcement gradient in a context with maximum ecological generality; that is, a context of naturally varying responding that permits various numbers of responses to come under the aegis of reinforcement on each of its occasions.

## 2. The shape of the gradient

### 2.1. Qualitative considerations

Some commonsense analyses will narrow the field of candidates for the form of the delay of reinforcement gradient. We can immediately rule out gradients that

increase with delay (Fig. 1A): although reminiscence may highlight remote events more clearly than recent ones, this occurs largely for rehearsed or otherwise singular or marked events. Our present concern is with a stream of relatively homogeneous responses, where the most recent tend to be the most salient. Similarly, we rule out an unweighted memory of the past (Fig. 1B), as this would grow insensitive with time and would apportion an ever-decreasing share of our attention to recent events.

A moving window on the past (Fig. 1C) has some of the features required for short-term memory (STM) and is, out of convenience, often used as a model of it (e.g., McDowell et al. 1992; Wearden & Clark 1989). But the corners define a concentration of information that requires energy to generate and maintain. In particular, the window in Figure 1C gives equal weight to everything in the epoch immediately preceding reinforcement and no weight to events just before that epoch. To achieve such a sharp discrimination would require remembering all the events up to the edge and then erasing the oldest as the newest is added. Not only is this unrealistically "computation-intensive," but there is also no justification for such absolute discrimination against the slightly older. Surely memory must decay rather than persist inviolate out to a point of total collapse.

Figure 1D gives one picture of decay, an exponential gradient traced by Equation 1:

$$\beta = \lambda e^{-\lambda d}. \quad (1)$$

Equation 1 is a weighting function, so if a response occurred at, say, 2 sec before reinforcement, we would evaluate the equation at  $d = 2$  to find the weight of the impact of reinforcement on that response. If  $\lambda = \frac{1}{2}$ , the event would be strengthened 78% of the amount of an event at  $d = 0$ . To calculate the animal's (or experimenter's) current memory, one would multiply events at each instant before reinforcement by Equation 1 evaluated at that instant and then sum the products. Equation 1 corresponds to one of the simplest electronic memories, an RC (resistance-capacitance) circuit.

A process that automatically accomplishes the same weighting and summing for a sequence of discrete events is the exponentially weighted moving-average (EWMA):

$$M_n = \beta y_n + (1 - \beta)M_{n-1}, \quad 0 \leq \beta \leq 1, \quad (2)$$

where  $M_n$  is the current memory,  $y_n$  the relevant attribute of the current response,  $M_{n-1}$  is the previous memory, and beta ( $\beta$ ) is the *currency parameter*. When  $\beta$  equals 1, all the emphasis is on the most current event and none on prior events. When  $\beta$  is small, most of memory is occupied by older events. Equation 2 may be iterated with each new event, or with each moment of time. In the limit where it is iterated with time and the moments between iterations become arbitrarily small, the process converges on Equation 1, with  $\lambda = -\ln(1 - \beta)$ .

Figure 1F shows how a series of interresponse times (IRTs) are filtered by Equation 2. This process (also called a linear average, or low-pass filter) has several advantages as a model of memory decay. It is very easy to calculate, makes minimal demands on memory (and is thus plausibly attributed to simple neural networks), and makes the most efficient use of information (Davis et al. 1993; Killeen 1981a; 1991; McNamara & Houston 1987).

Another shape that has been proposed for the gradient

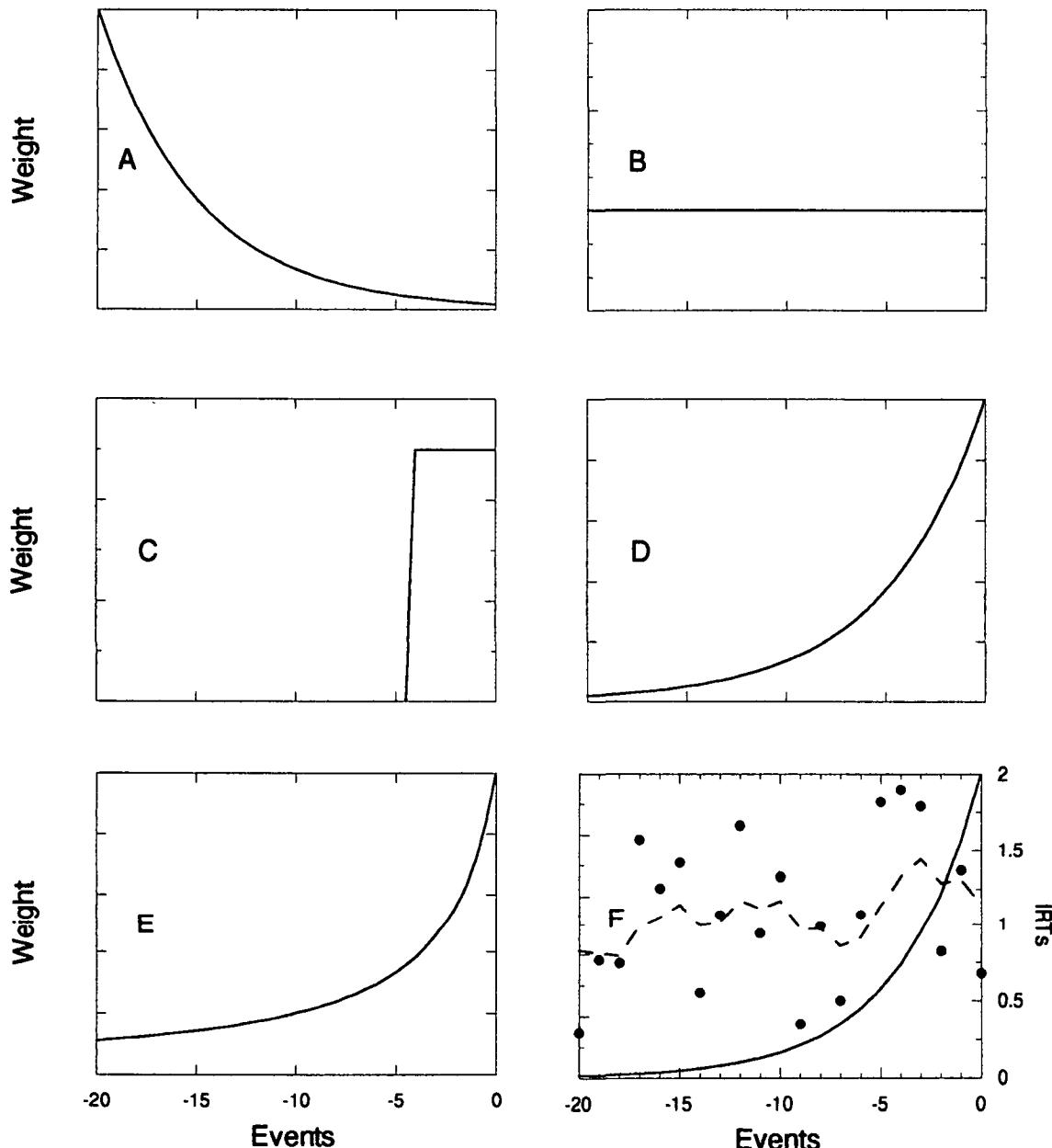


Figure 1. Different ways of weighting the past. The ordinates are the weights of an event in memory, the abscissas the distances from the most current event at  $y = 0$ . The function shown in Panel A may be immediately ruled out, as it weights remote events more heavily than recent ones. The function in Panel B weights all events equally and thus in a long life must give vanishingly small weight to the most recent event. Panel C shows a rectangular window that may serve as a rough approximation to short-term memory, but its calculation is computationally intensive and shows a biologically implausible discontinuity at some point in the past. The function in Panel D is an exponential decay (Equation 1), a continuous approximation to the discrete geometric discounting of Equation 2. The latter is easily calculated and shows the gradual degradation that we believe to be true of short-term memory. Panel E shows a hyperbolic weighting, which decays more quickly than the exponential at first and more slowly later. Panel F demonstrates how Equation 2 weights attributes of recent responses, in particular, interresponse times (IRTs), when  $\beta$  is  $\frac{1}{4}$ .

is the hyperbolic function (Fig. 1E; Mazur 1984; Rachlin et al. 1991). The present experiments are not intended to distinguish between these similar forms and their major conclusions should be robust relative to the ultimate decision between them. The computational convenience, correspondence with simple elements in linear systems theory (McDowell et al. 1992; McDowell & Kessel 1979), and intuitive clarity of the EWMA will make it our process of choice in the present analysis and its continuous realization, the exponential decay, the inferred form of the gradient.

## 2.2. Event driven or time driven?

Time is measured by counting events. In the case of Newtonian time we devise precisely periodic events that continually index counters; behavior's time, however, is driven by those internal pacemakers and external stimuli that capture attention (Killeen 1991). Decay of short-term memory is very slow in an undisturbed environment (Brown 1958) and dismally fast when other events are interpolated (Peterson & Peterson 1959). Analogously, the conditioning of a stimulus or response to a reward is

debased by interposing other stimuli or responses. Such distractors do not so much subvert attention while time elapses but rather by entering memory they move time along, iterating Equation 2 and thereby downweighting prior stimuli and responses. Although the theory is developed as a discrete, event-driven process, we shall later see (sect. 5.3.1) that updating is continuous during a response. For now, however, we think in terms of Equation 2 operating on quantal responses.

### 2.3. Quantitative considerations

It is easy enough to assert that conditioning should be best when our definition of behavior is the same as the organism's, but how do we know the organism's definition? What is the proper value for  $\beta$  in Equation 2? The curve in Figure 1F shows how an organism operating according to Equation 2 with  $\beta = \frac{1}{4}$  weights a sequence of IRTs, and the broken line gives the resulting average. If an experimenter makes reinforcement contingent on only the last IRT, requiring that its value be less than 1, the last response will be judged to have satisfied the criterion. But because many of the recent responses were associated with long IRTs, the subject's memory of its behavior is that of slow responding. Reinforcement at this point will, perversely from the experimenter's viewpoint, move behavior in the wrong direction. We may use this failure to communicate as a technique for inferring the organism's value of  $\beta$ . We may test various candidate values of  $\beta$ , and if one stands out as most powerful in controlling behavior, we infer, *per hypothesis*, that it is also the organism's rate constant. The approach is not unlike that used to determine the natural frequency of systems by applying various driving frequencies and determining where resonance occurs. To see this better, it is useful to simulate the postulated results.

## 3. Experiment 1: Simulation

Is reinforcement most effective when the criterion for its delivery weights the past in the same manner as the subject? In Equation 3:

$$\hat{M}_n = \alpha y_n + (1 - \alpha)\hat{M}_{n-1}, \quad 0 \leq \alpha \leq 1, \quad (3)$$

$\hat{M}_n$  signifies our current definition of the relevant response attribute, based on a linear average of the most recent response and the previous average. We may make reinforcement contingent on a value of  $\hat{M}_n$  that exceeds some criterion value. The only differences between Equation 2 and Equation 3 are the values of the currency parameters and the organisms – subject or experimenter – in which they reside. If our value for  $\alpha \approx 1$ , then we are attending only to the last response and the gradient is very steep. If at the same time the organism's value of  $\beta$  is substantially less than 1, conditioning should be less than optimal, as the experimenter and organism are focusing on different epochs of behavior. The point of this simulation is to determine whether it is true that conditioning proceeds fastest when our definition of the response coincides with the animal's, that is, when  $\alpha = \beta$ .

In these simulations groups of three "stat-rats" were constructed with the same memory parameter  $\beta$  for each animal within a group but with values of 1.0, 0.50, 0.25,

0.12, 0.06, and 0.03 characterizing the different groups. The IRTs of the "rats" were selected from a normal distribution with a mean of  $M_n$  and a standard deviation of 1:  $N(M_n, 1)$ . Within conditions, a value of  $\alpha$  was assigned and held constant for 1,000 responses. After each response Equation 3 was iterated and the value of  $\hat{M}_n$ , the experimenter's average, was tested to see whether it exceeded the value for the last trial. If  $\hat{M}_n$  equalled or exceeded that criterion, the stat-rat was "reinforced" for its behavior. Reinforcement incremented the mean value of the population from which responses are selected to  $M_n$ , the value specified by Equation 2, the organism's picture of what was getting reinforced. At the same time, the criterion for reinforcement was increased to  $\hat{M}_n$ , the value specified by Equation 3, the experimenter's conception of the last response. Thus, for each new reinforcement, the response attribute had to exceed the value that received reinforcement on the last trial. There are two separate definitions of the response, however, each of which involves a weighted memory of the past: behavior changes according to the animal's definition, whereas the contingencies change according to the experimenter's definition. After 1,000 training trials the final value of  $\hat{M}_n$  was recorded, and a new value for  $\alpha$  was explored.

Figure 2 shows the results. The ordinate is the maximum value of  $M$  that was achieved, the abscissas are the values of the experimenter's memory constant  $\alpha$ , and the parameter is the value of the "subjects" memory constant,  $\beta$ . Learning does indeed proceed fastest when the experimenter's weighting of the past, epitomized by  $\alpha$ , equals the animal's, epitomized by  $\beta$ . In the next experiment we will ascertain whether this also holds true for real organisms.

Note also that the overall level of acquisition ( $M$ ) was greatest under the conditions in which  $\beta$  was large. This may seem strange, for should not lengthened memory (corresponding to smaller values of  $\beta$ ) enhance learning? Not necessarily. Decreasing the value of  $\beta$  decreases the variance of  $M_n$  by the factor  $\beta/(2 - \beta)$ , providing fewer instances of extreme scores that the "trainer" could capitalize on to move responding quickly in the preferred direction. When the past weighs heavily in memory, behavior inevitably becomes conservative.

## 4. Experiment 2: Tuning curves for pigeons' short-term memory

### 4.1. Procedure

Four pigeons (*Columba livia*), subjects 14, 37, 40, and 41, with various prior histories of experimentation, were placed in a standard enclosure with the central Gerbrands response key transilluminated with white light. A houselight provided general illumination while white noise masked ambient sounds. The reinforcer was mixed grain, available from a standard feeder for approximately 2.5 seconds. The experiment was controlled by a computer, which generated criteria for reinforcement in a manner similar to that of the previous experiment.

The subjects were trained to respond on a variable-interval (VI) 60-sec schedule, in which reinforcement followed the first response after an interval of time that was selected from an approximately constant-probability distribution (Catania & Reynolds 1968). After several months of pilot experiments we arrived at the following *percentile reinforcement schedule* (Galicka 1988; Platt 1973). The criterion for reinforcement was

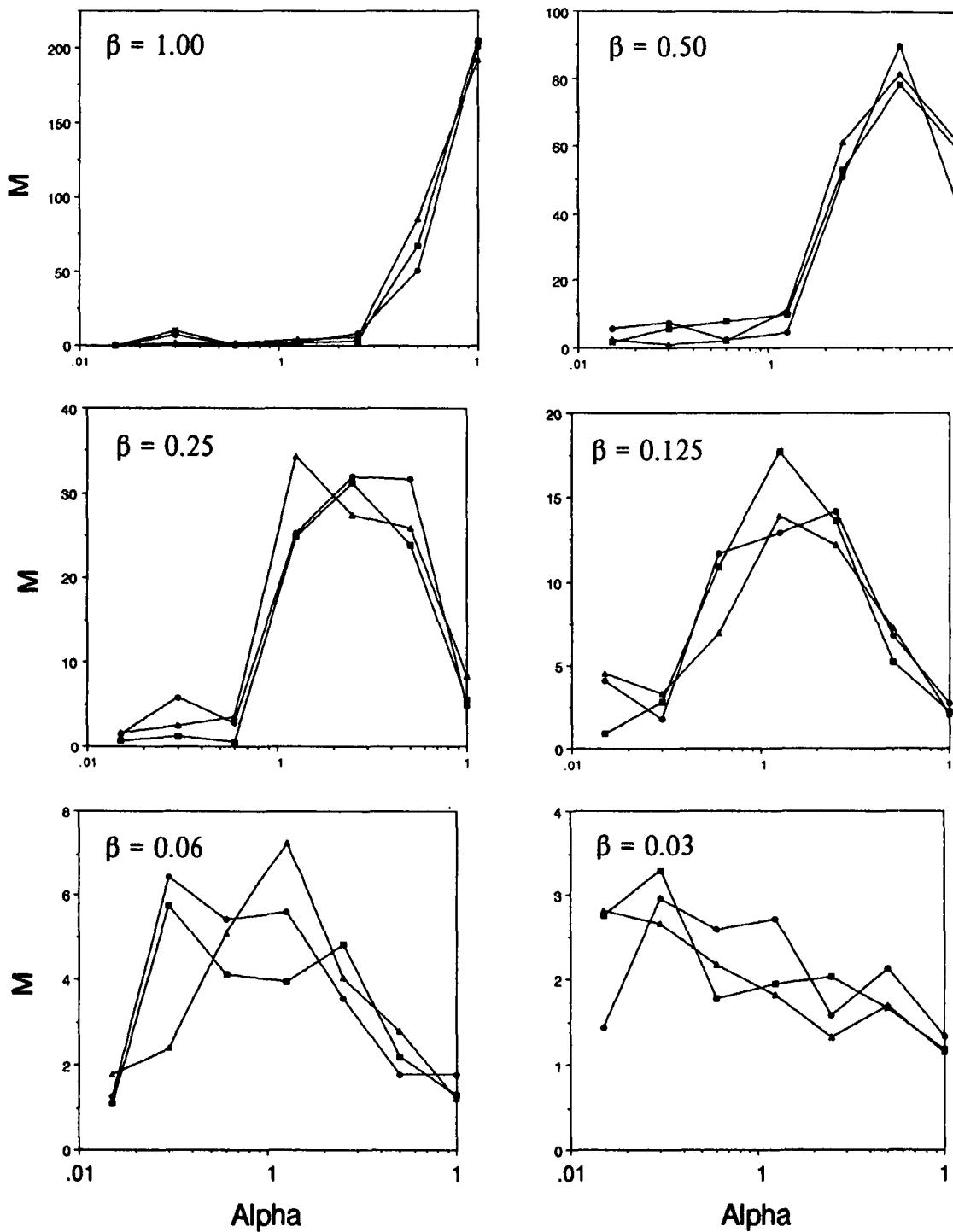


Figure 2. Results of the simulation experiment for several imputed values of the organism's memory decay rate ( $\beta$ ). The ordinate is the distance above the beginning value of zero that the reinforcement contingencies were able to move responding over the course of 1,000 trials. The abscissas are the values of  $\alpha$ , the experimenter's definition of the response and criterion for reinforcement. Notice that the functions reach their maximum when  $\alpha \approx \beta$ .

based on the recent history of responding and adjusted to maintain a constant upward (or downward) force on the length of the times between responses (IRTs). This was accomplished by setting the criterion for reinforcement above the 80th (or below the 20th) percentile of IRTs emitted since the last reinforcement. Reinforcement was delivered whenever the VI 60-sec schedule had primed and the experimenter's measure of the animal's response,  $\hat{M}_n$ , exceeded the appropriate percentile.  $\hat{M}_n$  was calculated from Equation 3.

The pigeons were randomly assigned either to an ascending

(subjects 14 & 41) or descending (subjects 37 & 40) series of values for  $\alpha$  from the set: 1.0, 0.50, 0.25, 0.12, 0.06, and 0.03. Once assigned, the value was maintained for four weeks (five sessions per week, 54 reinforcers per session). During the first and third weeks of each condition the pigeons were reinforced whenever the VI schedule was primed and the value of the experimenter's weighted average of IRTs,  $\hat{M}_n$ , was less than 80% of the recent IRTs (i.e., below their 20th percentile). For the second and fourth weeks,  $\hat{M}_n$  had to be greater than 80% of the recent IRTs (i.e., it had to be above their 80th percentile). This

cycle was continued until the animals had received exposure to all values of  $\alpha$ .

#### 4.2. Results

Response rates changed in the direction in which the percentile schedule forced them, and the force was more effective at some values of  $\alpha$  than at others. Figure 3 shows response rate as a function of sessions for each of the animals. The data from the first and third weeks are averaged, as are those for the second and fourth, so each condition in which a different value for  $\alpha$  was explored is represented by 10 data points. Consider the data for subject 37. The increase in response rate shown by the first 5 data points is consistent with the experimental contingencies during the first and third weeks, which required the last IRT to be shorter than 80% of the recent IRTs. The decreasing trend over the next 5 data points is consistent with the experimental contingencies during the second and fourth weeks. Over all values of  $\alpha$ , the changes appear to be greatest at  $\alpha = 0.5$  for this subject and  $\alpha = 0.25$  for the other subjects. Subjects 14 and 41 began at  $\alpha = 0.03$ , but their data have been plotted from right to left so that the values of  $\alpha$  coincide with those from subjects 37 and 40.

In order to characterize better the efficacy of the reinforcement contingencies, I have fitted linear "learning curves" to these data. Linear regression lines have two parameters, slope

and intercept, either or both of which may be affected by the changes in response rate. To reduce the number of parameters and simplify analysis, I forced the origin of the straight lines through the first datum and let the slopes alone reflect the changes wrought by the contingencies.

Because contingencies can have an important effect within the first session, I based the learning curves on the last condition of the previous session. Thus, for subject 37 condition  $\alpha = 1$ , the first datum on which the descending line is based (at an ordinate just under 1) is the average of the last sessions at the 20th percentile (weeks 1 and 3). The lowest point of the next, ascending, line for  $\alpha = 0.5$  (at an ordinate of 0.7) is based on the average of the last session at  $\alpha = 1$ , 80th percentile (week 4 of that condition), and the last session at  $\alpha = 0.5$ , 80th percentile (week 2 of that condition). The lines are oriented to minimize the sum of squares deviation around the remaining 4 data points. Thus, each line starts on the last datum from the previous condition and runs through the first 4 data points of the condition it represents. The lines organize our appreciation of the data and show that learning was fastest for intermediate values of  $\alpha$ .

This analysis is summarized in Figure 4, where the absolute values of the slopes of the ascending and descending curves for each condition are averaged and plotted as a function of  $\alpha$ . Subjects 14 and 41, with ascending values for  $\alpha$ , showed better overall learning

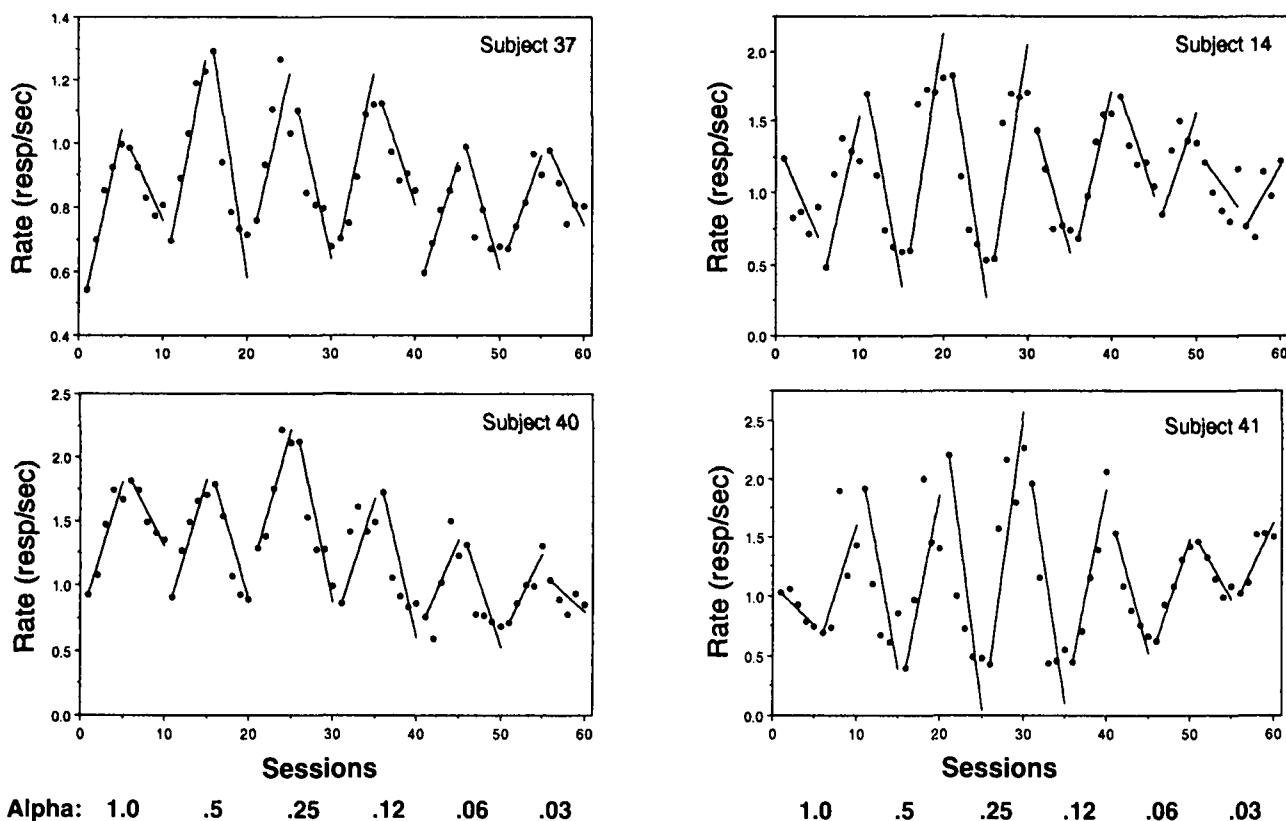


Figure 3. Learning curves for four pigeons whose response rate was alternately driven high and then low by experimental contingencies. The ordinate is response rate in pecks per second. The abscissas are the nominal session numbers (after two blocks of five sessions, the conditions were replicated and are plotted here as averages over the nominal session numbers). In the first block of five sessions for subjects 37 and 40 response rate was driven high by experimental contingencies corresponding to Equation 3, with a value of 1 for  $\alpha$ . For the second block of five sessions rate was driven low by the same experimental contingencies. (For the third and fourth blocks, not explicitly shown, these conditions were replicated and are combined here with the earlier data to yield the figure as plotted.) The value for  $\alpha$  was then decreased to 0.5 and the procedure continued. For the other two pigeons, conditioning started with a value of 0.03 for  $\alpha$ , which was then increased in like fashion. The historical run of the data in this figure goes from right to left for these two subjects. For all animals, each symbol corresponds to the response rate over the course of an experimental session and derives from the average of two sessions spaced two weeks apart. The straight lines are regressions anchored in the last session of the previous condition. Note that the slopes of these lines are greatest at  $\alpha$ 's of 0.5 or 0.25.

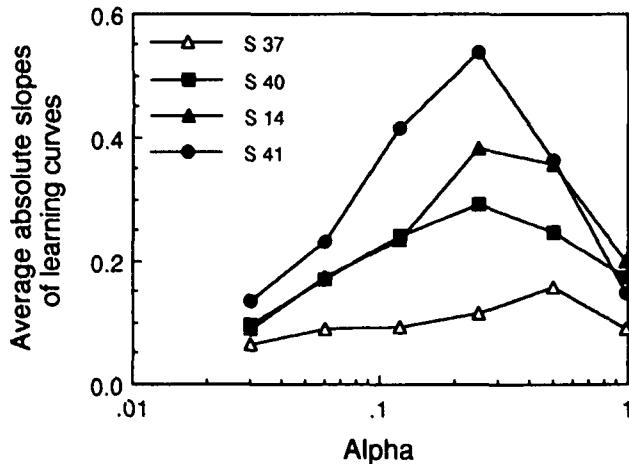


Figure 4. The slopes of the learning curves shown in Figure 3. The ascending slope and the absolute value of the descending slope for each value of  $\alpha$  are averaged and displayed for each subject.

than subjects 37 and 40. Subjects 14, 40, and 41 showed peak learning at  $\alpha = 0.25$ , whereas subject 37 peaked at  $\alpha = 0.50$ .

We may characterize these results by saying that pigeons' response memory is consistent with a linear average having a currency parameter of about  $\frac{1}{4}$ . In Figure 5, I have replotted the data from Figure 4 normalized to set the maximum slope at 1 and added a theoretical curve generated by the assumptions that motivated this experiment – that conditioning depends on the correlation between the experimenter's requirements for reinforcement and the organism's understanding of them (see Appendix A). The curve shows the correlation coefficients for different values of  $\alpha$  when  $\beta = \frac{1}{4}$ .

#### 4.3. Discussion

This experiment showed that we have much better control of organisms' behavior insofar as we construe a

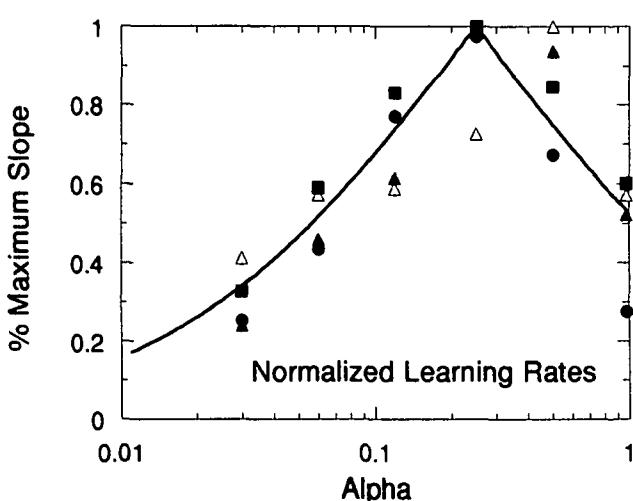


Figure 5. A tuning curve for reinforcement. The data from Figure 4 are normalized to the maximum learning rate. The curve is the predicted locus of the points for  $\beta = 0.25$  (see Appendix A and line 1 of Table 5). The left tail of the curve approaches  $\alpha/\beta$  as  $\alpha \rightarrow 0$ .

response to consist not of a punctate operation of a manipulandum but rather of a fading memory of behavior in which recent events are weighted more heavily than earlier ones and in which the mean age of the memory ( $1/\beta$ ) encompasses about four responses. To get a picture of how the impact of the past decreases with each new response, we may assign the memory  $M_0$  a current value of  $\beta$  and the response attribute  $y_n$  a value of zero for each iteration of Equation 3 thereafter. It is this process that generated the weighting functions shown in Figure 1F. For  $\beta = \frac{1}{4}$ , after 5 responses the weight has dropped to 0.06,  $(\beta(1 - \beta)^N)$ , and after 10 responses to 0.01.

Other researchers have studied the role of responses preceding the one that produces the reinforcer. Catania (1971) executed a particularly neat set of experiments, in one of which he provided reinforcement to pigeons for sequences of responses between keys A and B. Figure 6 shows the proportion of responses on key B for each of the reinforced sequences. Obviously, the greater the number of B responses required, the greater the proportion. But the position of the B requirement in the sequence also had an effect on the proportion of B responses. If we assign a response attribute of 0 for key A responses and 1 for key B responses, by iterating Equation 2 we may predict the animal's memory for B responses (relative to A responses) upon reinforcement under each of the conditions. We set  $\beta = \frac{1}{4}$ , and successively assign values of 1 or 0 to  $y_n$  as appropriate for each of the sequences represented on the x-axis, to obtain the predictions graphed next to the obtained data. Although the predictions underestimate the proportion of responses on the B key, they show a perfect rank-order correlation with the obtained data.<sup>1</sup>

For  $\beta = \frac{1}{4}$ , if we reinforce a single response in isolation we are only utilizing 25% of the potential effectiveness of the reinforcer. This is evident in Catania's (1971) data, where we may infer from the last column that reinforcing an A preceded by three B's generated a probability of A equal to about 30% (the complement of that column), whereas when preceded by three A's the probability of

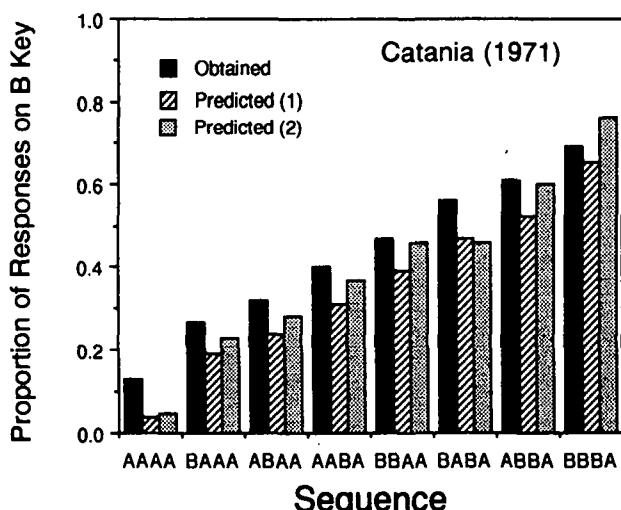


Figure 6. Proportion of responses on the B key when reinforcement was contingent on the sequence of responses shown. The solid bars represent the obtained data and the diagonally striped bar the data predicted from Equation 2 with a value of  $\frac{1}{4}$  for the currency parameter. The second set of predictions is the same as the first but gives 50% less weight to the final response.

responding A increased to almost 90% (the complement of the first column). Extended response requirements increase the coupling between the response and reinforcement; they also strengthen the control by stimuli present during that responding, facilitating discriminations that are otherwise difficult or impossible to establish (Williams 1972).

This importance of responses that precede the last response has been recognized in both the design of reinforcement contingencies, such as the “changeover delay” used in concurrent schedules to insulate responses on one key from reinforcement intended for those on another, and in the design of models of the resulting performance (Davison & Jenkins 1985). The current characterization of the extended response should facilitate those efforts.

Why is response memory so brief? Whatever the total capacity of STM, it must be allocated as shown in the bottom panels of Figure 1. To the extent that a longer view of the past permits one to deal with remote associations, it also undermines one's ability to deal with immediate precursors. Short-term memory is a temporal lens, focused by evolutionary pressures on the epoch that is most likely to have causal relevance to events of biological significance. Its responsiveness to the most immediate input (the derivative of Equation 2 with respect to  $y$ ) is simply  $\beta$ , the currency parameter. As we take the longer view, we necessarily weight the present less heavily; and the longer view, comprising more events, has more inertia. The exponentially decreasing weights of Equations 1 and 2 respect the exponentially ramifying chains of causal relevance, thus appropriately allocating credit for a reinforcer most strongly to those events most proximal to it.

## 5. A general theory of reinforcement

As noted by Zeiler and Buchman (1979), two counterpoised forces affect the control of behavior under schedules of reinforcement: as we increase the rate of reinforcement, we activate more behavior, but at the same time we decrease the number of responses that each reinforcer can influence. This is because reinforcement isolates responses that precede it from other reinforcers that follow it, and thus truncates the reach of reinforcement (Catania et al. 1988; Killeen & Smith 1984; Williams 1978; see also Appendix B). Because of this, continual reinforcement of one response after another can provide only a fraction ( $\beta$ ) of the strengthening that is available when many responses precede the reinforced response.

This insight is developed by identifying three factors that affect the control of behavior by reinforcement; they constitute the fundamental axioms of the theory:

1. *Activation*: an incentive activates  $a$  seconds of responding;
2. *Temporal constraint*: a response requires delta ( $\delta$ ) seconds for its completion;
3. *Coupling*: reinforcement occurs when an incentive enters a memory that also contains a response. The proportion of memory occupied by target responses, and thus the effectiveness of reinforcement, is given by the coupling coefficient, zeta ( $\zeta$ ).

Simple models of these assumptions are developed and justified in Appendix B (activation and constraint) and

Appendix C (coupling). The manner in which reinforcement moves the trajectories of behavior toward the states described by those asymptotic models is discussed in Appendix D (dynamics). The mathematical principles of reinforcement are summarized in the body of the target article, where they are applied to representative data.

### 5.1. Activation and constraint

From Assumptions 1 and 2 it follows that incentives delivered at a rate of  $R$  per second can potentially instigate  $a/\delta$  responses per second. But each response “uses up” some of the time available for making other responses. Staddon (1977) has shown that a simple correction for such temporal constraints yields Equation 4, originally introduced by Herrnstein (1970):

$$\hat{B} = \frac{kR}{R + c}, \quad (4)$$

where  $\hat{B}$  is the measured rate of responding,  $R$  is the rate of reinforcement, and  $k$  and  $c$  are free parameters:  $k$  is the asymptotic rate of responding as  $R \rightarrow \infty$  (Herrnstein 1974);  $c$  is interpreted by Herrnstein as the rate of reinforcement available for other, unrecorded responses. Over a large range, response rates on interval schedules adhere closely to Equation 4 (McDowell 1980), which has been rederived from a variety of assumptions (see, e.g., Williams 1988). Our model of temporal constraint also yields Equation 4,  $k$  also designates the asymptotic response rate, but  $c = 1/a$  (see Appendix B).

Thus far, the development assumes that all the activation is focused onto target responses, but we have seen from Experiment 2 that this is only the case if the contingencies are carefully matched to the animal's memory. This is rarely the case. The rate of responding will be less than that predicted by Equation 4 to the extent that the coupling between incentive and target response is less than perfect. I now apply the activation-constraint model to ratio schedules, where its failure will set the stage for introduction of the coupling coefficient.

### 5.2. Fixed-ratio schedules

To use Equation 4 we must know the rate of reinforcement,  $R$ . On fixed-ratio (FR) schedules, reinforcement is delivered immediately after the  $N$ th response. Rate of reinforcement is therefore proportional to the rate of responding ( $\hat{B}$ ) and inversely proportional to the ratio requirement ( $N$ ); it follows that the *schedule-feedback function* for ratio schedules is:

$$R = \hat{B}/N. \quad (5)$$

Substituting into Equation 4 and rearranging give the rate of responding predicted on the basis of Equation 4 (Pear 1975):

$$\hat{B} = k - cN. \quad (6)$$

Equation 6 predicts that as  $N$  approaches zero, response rate will approach its maximum,  $k$ . It fits *none* of the FR data, some of which are shown in Figures 7 and 8. This is because Equation 4 takes a single operant response as the event that is strengthened and does not address the issue of how reinforcement affects more or less of the

animal's memory of previous responding. But we already know that when  $N = 1$ , reward is only  $\beta$  times as effective as the case when  $N$  is very large. In general, the ceiling  $k$  will depend on the extent to which we match our definition of the reinforced response to the animal's definition. On interval schedules, for instance, any response that takes time moves the animal along toward satisfaction of the schedule, but many of those responses are not measured by the experimenter, thus lowering the apparent value of the ceiling. In the following sections I derive a general solution to this problem by introducing a *coefficient of coupling* as an implicit factor of  $k$ . To aid memory, refer to Table 1 for a synopsis of parameters and their interpretation.

### 5.3. Coupling behavior to incentives

Equation 2 shows that the first response contributes a memory of strength  $\beta$ , and iteration of that equation shows that the second adds another  $\beta$  while the impact of the first has decayed to  $(1 - \beta)\beta$ , and so on. As the

Table 1. Cast of characters, in order of appearance

Symbol	Interpretation
$\beta$	Beta, weight in short-term memory assigned to the most recent response
$M_n$	Contents of the subject's memory after the $n$ th response
$\alpha$	Alpha, weight assigned by an experimenter to most recent response
$\hat{M}_n$	Experimenter's characterization of subject's memory after the $n$ th response
$B$	Rate of responding predicted by the complete model
$\delta$	Delta, minimum interresponse time; reciprocal of maximum response rate
$k$	Constant, approximately equal to asymptotic response rate; reinterpreted as product of coupling constant and maximum response rate
$R$	Rate of reinforcement
$a$	Specific activation; number of seconds of responding elicited by a given incentive under the operative motivational conditions
$N$	Number of responses required to satisfy a ratio schedule
$\lambda$	Lambda, measured rate of decay of short-term memory
$\zeta$	Zeta, coupling coefficient; degree to which memory is filled by target responses
$\rho$	Rho, coupling constant; proportion of target responses in response trajectory. Depends on experimental conditions, specific response, and extent of conditioning; assumes values of around 1 for ratio schedules and $\frac{1}{2}$ for interval schedules under standard conditions
$\lambda'$	Lambda-prime, intrinsic rate of decay of short-term memory; $\lambda' = \lambda/\delta/\rho$
$w_i$	Weight given in memory to the $i$ th response back from an incentive; $w_i = \beta$
$\tau$	Tau, interresponse time; $\tau = 1/B$

number of responses,  $j$ , increases to the number required for reinforcement,  $N$ , iteration of Equation 2 predicts that the organism's response memory,  $M_N$ , will fill according to:

$$M_N = \sum_{j=1}^N \beta(1-\beta)^{j-1}. \quad (7)$$

The sum of this series,  $1-(1-\beta)^{N-1}$ , may be approximated by:

$$M_N = 1 - e^{-\lambda N}, \quad (8)$$

where  $\lambda = -\ln(1 - \beta)$  (that is,  $\lambda$  is just slightly larger than  $\beta$  over its behaviorally relevant range).

Equation 8 states that the contribution of the ratio requirement to the definition of the target response increases as a cumulative exponential function of  $N$ . When  $N$  is small, the number of responses in memory must be small as well; as  $N$  is increased, the response memory,  $M_N$ , approaches its maximum ("saturates") and response rate approaches its ceiling. On FR schedules, all  $N$  target responses must have occurred before reinforcement, so the coupling between incentive and responding,  $\zeta$ , is proportional to the degree of saturation of memory by the target response up to that point,  $\zeta = \rho M_N$ . The constant  $\rho$  will be interpreted in section 6; here it is fixed at 1.

I rewrite Equation 6, making the coupling coefficient explicit and substituting  $1/a$  for  $c$  to obtain the predicted response rate on ratio schedules:

$$B = \frac{\zeta}{\delta} - \frac{N}{a}, \quad N \leq a/\delta; \\ B = 0, \quad N > a/\delta; \quad (9)$$

where

$$\zeta = \rho(1 - e^{-\lambda N}), \text{ and } \delta, a > 0.$$

As the response requirement increases, more and more of the effects of a reinforcer make contact with responses, and move the rate toward its ceiling, as shown by the first term in Equation 9. At the same time, however, the arousal is decreasing linearly with  $N$ , as shown by the second term. Thus, our model of FR performance predicts an inverted U-shaped change in response rate as a function of ratio requirement. Figures 7 and 8 show that Equation 9 provides an accurate picture of changes in response rate for pigeons, rats, and mice.

Notice that the linear descending segment, the pure arousal effect, governs the function after the accumulation of responses in memory has saturated, above FRs of about 25. If we were to project this asymptotic linear segment back to the left, it would intercept the  $y$ -axis at an ordinate of  $1/\delta$ . To the right, the asymptote intercepts the  $x$ -axis at  $a/\delta$  (by which point  $\zeta \approx \rho \approx 1$ ), showing that when the number of responses demanded by the experimenter equals or exceeds the number elicited by the incentive ( $N \geq a/\delta$ ), response rate must fall to zero: the "ratio is strained" to the breaking point and performance extinguishes. Notice that the ratio  $a/\delta$  converts the seconds of responding that may be sustained by the incentive ( $a$ ) into

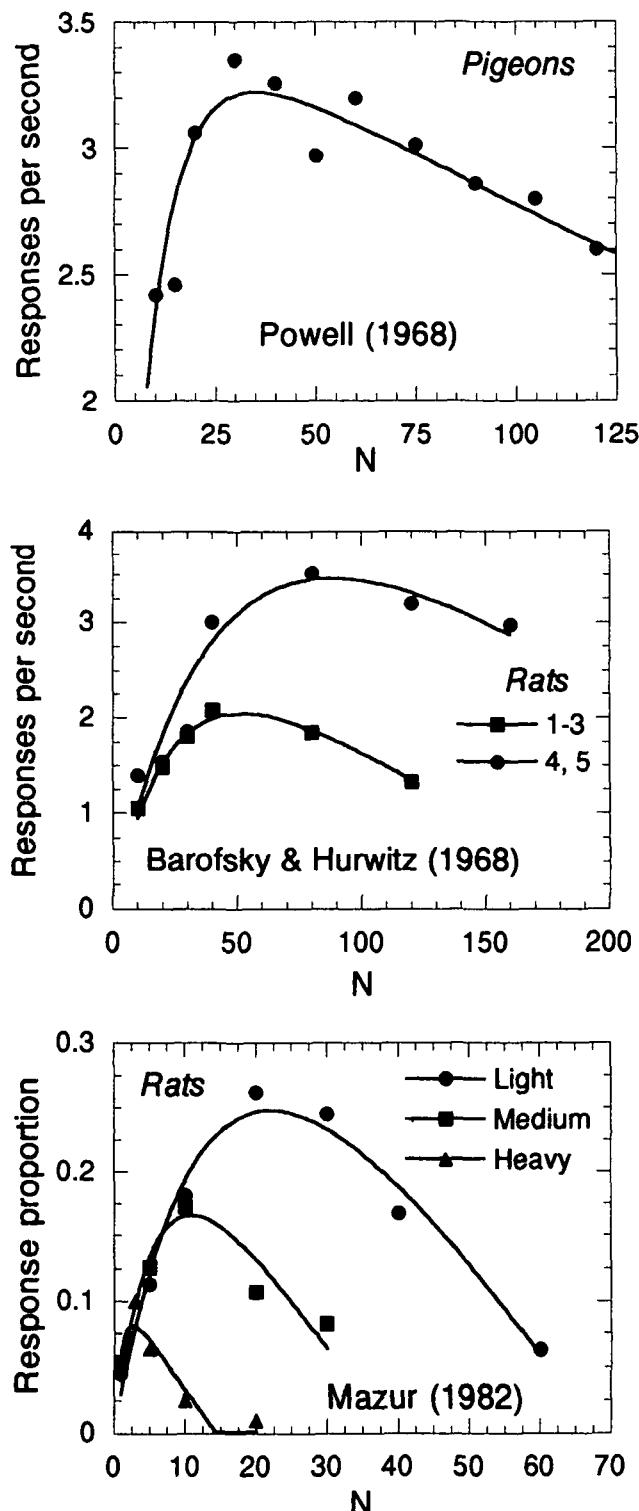


Figure 7. Mean response rates on fixed-ratio schedules as a function of ratio size ( $N$ ). The first panel shows the average “running rate” (i.e., rate excluding the postreinforcement pause) of three pigeons (Powell 1968, as reported in Mazur 1982). The second panel shows the average data from five rats; the third shows the average data from three rats who experienced different lever weights. Equation 9 generated the curves, whose parameters are found in Table 2.

the number of responses of duration  $\delta$  that may be sustained:  $a/\delta$  is the “extinction ratio” hypothesized by Skinner (1938).

The dimensions of these parameters reflect the role

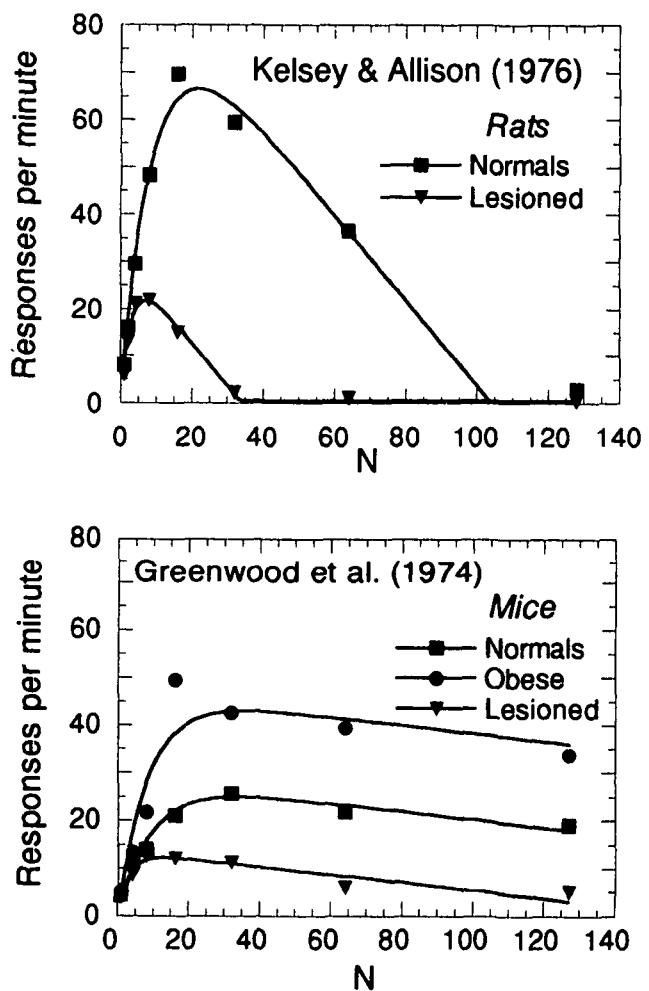


Figure 8. Mean response rates on FR schedules as a function of ratio size. The top panel shows the average data of seven control (sham-lesioned) rats and five rats made obese by surgical lesions of the VMH. The bottom panel shows the data from three groups of mice: normal black mice, normal mice made obese by aurothioglucose lesions, and genetically obese mice (Greenwood et al. 1974). Equation 9 generated the curves, whose parameters are found in Table 2.

they play in the theory: the reciprocal of  $\lambda$  tells us the average number of responses that are coupled to a reinforcer in memory;  $\delta$  tells us the number of seconds it takes to make a response, and thus constitutes both the lower limit of an IRT and the reciprocal of the maximal attainable response rate;  $a$  tells us the number of seconds of responding that can be activated or elicited by each incentive and thus provides a measure of incentive-motivation. Because these three key parameters – each identifying a different causal factor in the control of behavior – may be directly inferred from such graphs (see Appendix C), such mapping of response rates under different FR values provides a very useful diagnostic technique. I shall return to analysis of these figures and their parameters in the next section.

**5.3.1. Uncovering hidden structure.** It turns out that the rate of memory decay ( $\lambda$ ) covaries with the estimated response duration ( $\delta$ ):  $r = 0.81$  for ratio schedules and 0.79 for interval schedules over some two dozen experimental conditions. The longer or more complicated a

response, the greater the proportion of memory it occupies. We incorporate this regularity of nature into the structure of the theory by multiplying the rate of memory decay by  $\delta$ , thus reducing the variance in the decay rate parameter across experiments and schedule types. It is accomplished by writing  $\lambda = \lambda' \delta / \rho$ , where  $\lambda'$  is the intrinsic rate of decay per second of responding. The constant  $\rho$  is formally introduced here for consistency with the theory as developed in section 6, but is fixed at 1. This correction for decay of memory during the act of responding is enforced in all applications of the theory, making it technically a mixed, "metered model" of memory (Clark 1976; Mesterton-Gibbons 1989).

Let us now review the FR data and the parameters of the model that optimize the goodness of fit to them. The "running rates," the response rate once responding has begun, are shown for Powell's (1968) study and are predicted under the assumption that memory decays at the rate of 0.38 per response-second, that the minimum duration of a key-peck is 0.28 sec, and that a single reinforcer under his conditions can sustain 123 seconds of such responding (Fig. 7). Barofsky and Hurwitz (1968) found that their rats formed two groups, a high response rate group that could sustain larger FR values and a lower response rate group. Table 2 shows that the only difference between them was their minimum response durations,  $\delta$ . Mazur (1982) varied both FR value and the force required to operate the lever. Figure 7 shows that his data also took the predicted shape; Table 2 shows us that  $\delta$  increased with the force requirement, as did rate of decay of memory (even with the adjustment noted above), while the motivational parameter  $a$  remained unchanged.

Kelsey and Allison (1976; see also Fig. 8) compared control rats with rats having surgical lesions in the ven-

tromedial hypothalamus (VMH), which caused them to overeat freely available food and become obese but at the same time to be less ready to work for food. Greenwood et al. (1974) chemically lesioned the VMH of mice and also studied a group of genetically obese mice. Figure 8 shows that response rates were depressed in the lesioned groups of both species. Table 2 tells us that in neither case was memory or the specific activation ( $a$ ), affected; all the depressive effects were due to motor impairment.<sup>2</sup> Greenwood et al. concluded that "the naturally obese animals do not display the behavioral patterns associated with rodents made obese by hypothalamic damage" (p. 687); the current model is more specific: genetically obese mice had superior motor facility and enhanced motivation. The model also suggests a more rapid decay of memory for them, although that is driven by their very high rate at a single point ( $N = 16$ ).

Note the important role of  $\delta$  in Equation 9. Many of the effects of operations that have been traditionally interpreted as motivational may actually affect motivation only indirectly by modifying the minimum response duration. Where  $\delta$  is increased, as it may be by lesions, organisms will be unable to sustain high ratio requirements even though they are as motivated as control animals.

#### 5.4. Variable-ratio schedules

Variable-ratio (VR) schedules provide reinforcement after a variable number of responses averaging  $N$ . Their idealized implementations as "random-ratio" schedules provide a constant probability of reinforcement,  $p = 1/N$ , after every response. As I show in Appendix C, the coupling to these schedules increases as a hyperbolic function of their mean requirement and the appropriate model is:

Table 2. Parameter values for ratio schedules

Study	Species	Condition	Parameters <sup>a</sup>		
			$\lambda' \left( \frac{1}{\text{sec}} \right)$	$\delta \left( \frac{\text{sec}}{\text{resp}} \right)$	$a \left( \frac{\text{sec}}{\text{reinf}} \right)$
<i>Schedule: FR (Equation 9)</i>					
Powell (1968)	Pigeons		0.38	0.28	123
Barofsky & Hurwitz (1968)	Rats	# 1,2,3	0.13	0.31	65
		# 4,5	0.13	0.18	65
Mazur (1982) <sup>b</sup>	Rats	Lo Wt	0.12	0.63	44
		Med Wt	0.15	1.12	44
		Hi Wt	0.29	2.60	44
Kelsey & Allison (1976)	Rats	Normal	0.18	0.64	66
		Lesioned	0.18	1.94	66
Greenwood et al. (1974)	Mice	Normal	0.05	2.20	390
		Lesioned	0.05	3.73	390
		Obese	0.09	1.15	860
<i>Schedule: VR (Equation 10)</i>					
Mazur (1982) <sup>b</sup>	Rats		0.18	0.40	90
McDowell & Wixted (1986)	Humans	31 & 32	>>1	0.51	2,400
		36 & 37	0.78	0.34	100

<sup>a</sup>The measured rate of decay of memory ( $\lambda$ ) increases with the duration of the response ( $\delta$ ), so that  $\lambda = \lambda' \delta / \rho$ .  $\rho$  is fixed at 1.

<sup>b</sup>Mazur reported these data as % of time responding; for consistency of units I multiplied his data by 3, a calibration based on analysis of other experiments using similar equipment (Mazur 1983).

$$B = \frac{\zeta}{\delta} - \frac{N}{a}, \quad N \leq (a/\delta - 1/\lambda); \quad (10)$$

and  $B = 0,$  elsewhere;

where  $\zeta = \frac{p\lambda N}{1 + \lambda N},$  and  $p, a, \lambda > 0.$

The exponential approach to saturation of memory found under FR schedules (Equation 8) is thus replaced with a similar but slower hyperbolic approach to saturation under VR schedules; otherwise the models are identical. Figure 9 shows response rates of rats and humans on VR schedules and Table 2 the associated parameter values of Equation 10, which directs the curves through them. The coupling constant  $p$  is fixed at 1.

Although the fits to the rat data in Figure 9 are good, the data from the humans seem to tell a new story. For them there is negligible downturn of responding at high rates of reinforcement (low  $N$ ), which imputes large values to  $\lambda.$  Why should human memory appear to saturate so quickly? In using Equation 7 we assume that at the start of each trial memory contains none of the target responses because it is filled with consummatory responses involv-

ing the previous reinforcer. These reinforcers were points on a display, however, which interfere less with response memory than do primary reinforcers. The quick saturation is due not to a memory that is quickly filled but rather to one that is not greatly disrupted by symbolic reinforcement and which therefore includes responses preceding prior reinforcers. The response rates of subjects 31 and 32 persisted even for very large ratio values, suggesting that although the point displays were brief they were vested with substantial incentive value (Hayes & Hayes 1992).

The contribution of a response sequence to coupling is a concave function of its length (Equation 8), so that the longer sequences of responses on a VR do not compensate for the occasions on which memory is truncated by the short sequences. We therefore predict that FR schedules will command higher running rates than equal-sized VR schedules (except for very large values of  $N$ , where saturation is virtually complete for all sequence lengths), and this is what Mazur (1983) found.

Nowhere does the present theory postulate that ratio schedules reinforce short IRTs, nor that "molar contingencies" favor high response rates on ratio schedules, as do some alternate theories; yet it provides a good account of the data. Where contingencies of reinforcement attend only to responses and not to their temporal attributes, our models of reinforcement also need only attend to responses. Under ratio schedules only the target responses move the animal toward reinforcement, so they differentially fill memory at the time of reinforcement. Interval schedules maintain lower response rates, because any response (including unmeasured "interresponses") will move the animal toward reinforcement and thus debase the coupling to the target response. When the interresponses are experimentally identified (e.g., Catania 1971), we may find a coherent "response unit" (Platt 1971; Platt & Day 1979). Often, however, as in the case of interval schedules, they are not measured; this does not make them go away but rather necessitates their treatment as hypothetical constructs.

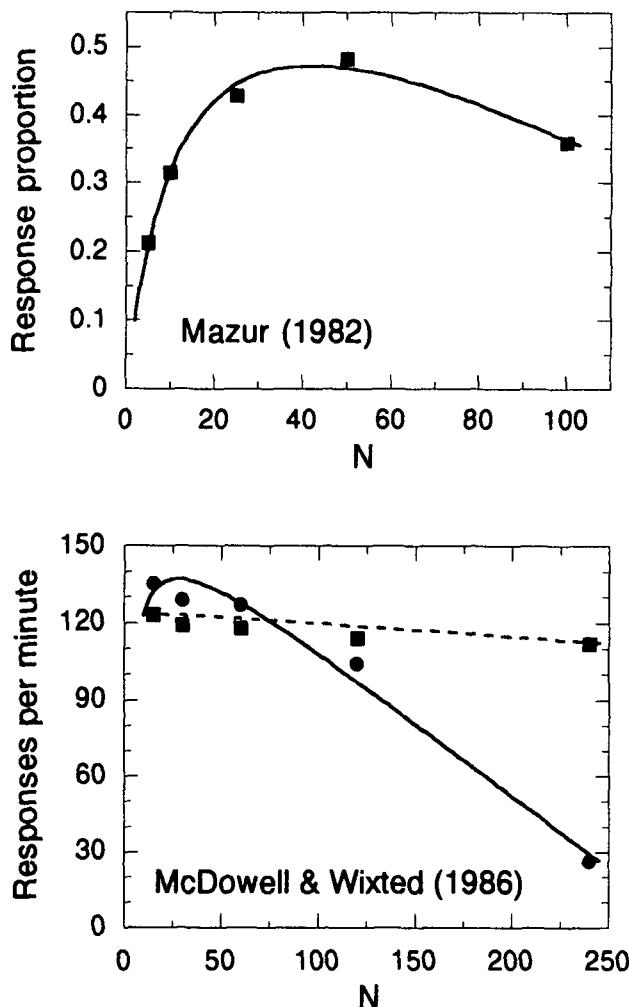


Figure 9. Mean response rates on variable-ratio schedules as a function of mean ratio requirement. The top panel shows the average data for a group of Lashley rats. The bottom panel shows the average data for human subjects, the squares from subjects 31 and 32, and the circles from subjects 36 and 37. Equation 10 generated the curves, whose parameters are found in Table 2.

### 5.5. Interval schedules

Under interval schedules reinforcement is provided for the first response after a fixed (*fixed-interval*, FI) or variable (*variable-interval*, VI) time. Herrnstein's (1970) model, Equation 4, provides a relatively accurate picture of response rates under them but the "maximum" rate,  $k$ , is much lower than expected. The median value of asymptotic response rate calculated by Herrnstein for the VI data of Catania and Reynolds (1968) was  $k = 1.2$  responses per second, only  $\frac{1}{3}$  of that attainable by pigeons under ratio schedules. Thus, when schedules make the time (or responses) between target responses a property upon which reinforcement bears, the models must be expanded to take those events into account (Platt 1979). There are various ways to do this. The present approach treats an interresponse time (IRT) as a sequence of responses ending with the target response. Each response, whether recorded or not, is susceptible to reinforcement, but only the final target response that terminates the IRT gets the full weight of reinforcement,  $\beta.$  Memory is filled with a mixture of target responses and unmeasured responses, so that even when it is saturated after a long string of behavior, only the proportion  $p$  will contribute to

the strengthening of the target response. Interval contingencies debase coupling by strengthening *any* sequence of responses that precede the target response. They do not differentially reinforce long IRTs; they do not reinforce IRTs at all (as Reynolds & McLeod [1970] also concluded; see also Wearden & Clark 1989). They reinforce the terminal response, giving it a weight of  $\beta$ , and they reinforce whatever mixture of target responses and interresponses happen to precede it in memory, giving them a weight of  $1 - \beta$ . Whether or not the mixture contains other target responses (and is thus comprised of "short" IRTs) or not (and is thus part of a long IRT) is irrelevant. Rather than punish premature target responses, these schedules indifferently strengthen them and everything else. It is this slackness that lets coupling drift to less than maximal levels, being stabilized only by the increment of  $\beta$  guaranteed by contiguous reinforcement.

In Appendix C, I show that the coupling coefficient for VI schedules is approximately:

$$\zeta \approx \frac{\rho\lambda B}{\lambda B + R}. \quad (11)$$

To predict behavior, this equation is inserted into the arousal-constraint model, Equations 4 and B7, and solved for  $B$ , yielding:

$$B = \frac{kR}{(R + 1/a)} - \frac{R}{\lambda}, \quad (12)$$

where

$$k = \rho/\delta, \quad \lambda, a, \delta > 0,$$

and

$$R < \lambda k - 1/a.$$

Note that the first term on the right-hand side (*rhs*) of Equation 12 is the familiar hyperbolic function of rate of reinforcement. Throughout most of their range, *response rates inherit the hyperbolic form* of Equation 4, generated by arousal and temporal constraints on responding. Equation 12 differs from the traditional version of Equation 4 in that it includes an inhibitory effect, due to truncation of memory, that is a linear function of reinforcement rate (second term in the *rhs* of Equation 12), and which can lead to an actual downturn in responding at very high rates of reinforcement. At low rates of reinforcement, or where  $\lambda$  is large, the suppressive effect is minimal and the functions similar. Aside from the memorial effect, the slackness in the contingencies will also keep rates below their theoretical maximum. This slackness is realized in a smaller value of  $\rho$ , which has emerged as a factor of  $k$ , and thus lowers ceiling response rates. Because the value of  $\rho$  is generally underconstrained by the data reviewed in this target article, I routinely set it to 1, letting the impact of variation in it be carried by changes in the value of  $\delta$ . Under interval schedules, its true value is on the order of  $\frac{1}{2}$  (see Appendices C and D).

Figure 10 shows the performance of rats under different conditions of deprivation and Figure 11 shows the performance of pigeons and rats of whom different responses topographies or forces were required. The parameters for Equation 12 are displayed in Table 3.

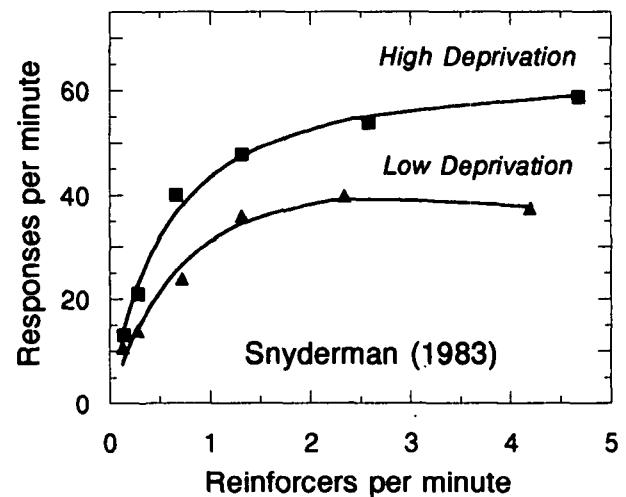
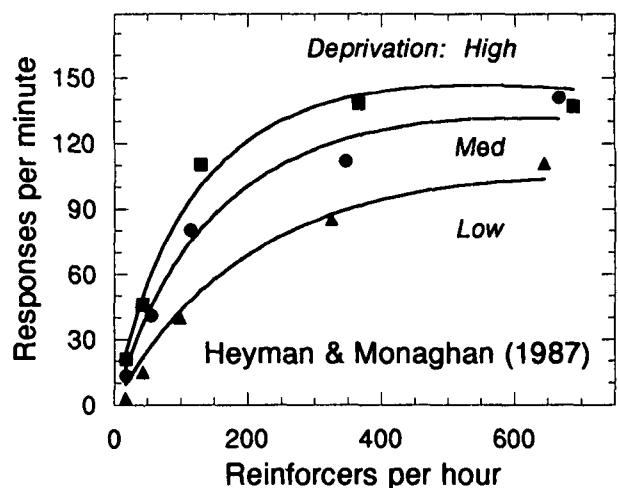


Figure 10. Response rates of rats reinforced on VI schedules under high (squares), medium (circles), and low (triangles) levels of deprivation. The curves are from Equation 12.

**5.5.1. Motivational effects.** Changes in deprivation, quality, and quantity of reinforcer should affect the activation parameter,  $a$ ; we always find this to be the case. Heyman and Monaghan (1987) cite the data from nine experiments in which motivational variables affected the value of  $a$ , including some of their own, shown in the top panel of Figure 10.

Heyman and Monaghan (1987) also argue that motivational variables do not affect the maximum response rate. Attempting to fit the data with just the hyperbolic part of Equation 12 can, however, cause one to infer incorrectly that  $k$  changes with motivation (when motivation is high, the rate curve will rise higher before the descending asymptote carries it down). McDowell and Wood (1984) argued that  $k$  increased with magnitude of reinforcement, but over 99% of the variance in their data may be accounted for by Equation 12 with  $k$  invariant. The data of Hamilton et al. (1985) show a clear downturn under VI 5 schedules, one that was greatest under motivational conditions that encouraged the highest response rates.

**5.5.2. Response topography.** Changes in the form of the response should obviously affect the maximum response

Table 3. *Parameter values for interval schedules*

Study	Species	Condition	Parameters <sup>a</sup>		
			$\lambda' \left( \frac{1}{\text{sec}} \right)$	$\delta \left( \frac{\text{sec}}{\text{resp}} \right)$	$a \left( \frac{\text{sec}}{\text{reinf}} \right)$
<i>Schedule: VI (Equation 12)</i>					
Heyman & Monaghan (1987)	Rats	Lo Dep	0.90	0.25	10
		Med Dep	0.90	0.25	17
		Hi Dep	0.90	0.25	24
		Lo Wt	0.90	0.25	15
		MedLo Wt	0.90	0.54	15
		MedHi Wt	0.90	0.51	15
		Hi Wt	0.90	0.81	15
Snyderman (1983)	Rats	Lo Dep	0.24	0.80	51
		Hi Dep	0.75	0.80	93
McSweeney (1978)	Pigeons	Peck	0.30	0.98	960
		Treadle	>1.0	3.60	960
McDowell & Wood (1984)	Pigeons	Lo Amt	0.90	0.39	15
		Med Amt	0.90	0.39	13
		Hi Amt	0.90	0.39	179
<i>Schedule: FI (Equation 12)</i>					
Allen & Kenshalo (1976)	Monkey	Press	0.44	0.50	13
		Drink	0.20	0.32	13

<sup>a</sup>The measured rate of decay of memory ( $\lambda$ ) increases with the duration of the response ( $\delta$ ), so that  $\lambda = \lambda' \delta / \rho$ .  $\rho$  is fixed at 1.

rate, and Heyman and Monaghan (1987) cite numerous studies that show this to be the case. Figure 11 shows the data from two of those studies, which are well fit by Equation 12, whose parameters tell us that the major effect was on  $\delta$ , as expected.

**5.5.3. Duration of reinforcement.** Manipulating the duration of an incentive affects more than activation level. Brief incentives do not completely reset response memory. The strengthening effect of a consummatory response increases with its duration according to the same cumulative exponential function as the saturation of memory for any other response (Equation 8; Killeen 1985); its ability to erase memory follows the same function (Killeen & Smith 1984). As the incentive's duration is extended, therefore, it both increases  $a$  and more effectively erases memory for the prior target responses and should thus have mixed effects: at short interreinforcement intervals, where truncation of memory keeps coupling below its maximum, increasing duration should degrade whatever residual coupling to prior responses is left and thus work against the increasing activation. At large interreinforcement intervals, however, there is adequate time to saturate memory with target responses so that the arousal effects should dominate the amnesic effects of the reinforcer and response rates should show greater covariation with the magnitude of the incentive. The idiosyncratic properties of very brief incentives, such as electrical stimulation of the brain, may be due, in part, to the interaction of these two aspects of reinforcement – their brevity permits coupling to responses that were emitted before the prior incentive, supporting maximal rates when  $R$  is large; yet because the growth of arousal

during the incentive is so minimal, the resulting small values of  $a$  cannot support responding at lower values of  $R$ .

**5.5.4. Fixed-interval schedules.** The highest rates of reinforcement, where we would expect to find the best evidence for a downturn in responding due to weakened coupling, are obtainable under FI contingencies. Figure 12 shows the rates of lever-pressing for food pellets by a rhesus monkey under high rates of reinforcement (Allen & Kenshalo 1976). Jason's rate of pressing shows a definite downturn as the interval between reinforcers fell below 10 sec; another monkey's responding did not show the downturn. In Appendix C, I show that the approximate rate model for FI schedules is the same as that for VI schedules; I thus use Equation 12 to fit these data.

Under FI schedules animals pause for about one-half the interval length before they begin to respond, whereas under VI schedules animals respond at a relatively uniform rate throughout the interval. The quantitative predictions are not undermined by this qualitative difference but merely require adjustments in the parameter values (see Appendix D).

Periodic schedules elicit "adjunctive" behavior, such as schedule-induced polydipsia, which follows the same principles as operant responses (Appendix C). Figure 12 shows the rate of concurrent drinking by Jason. The downturn in drinking at high reinforcement rates is almost surely due to competition from the terminal response rather than different limits on memory. These data remind us that the theory awaits elaboration for ceilings on response rates lowered by the concurrent emission of other responses.

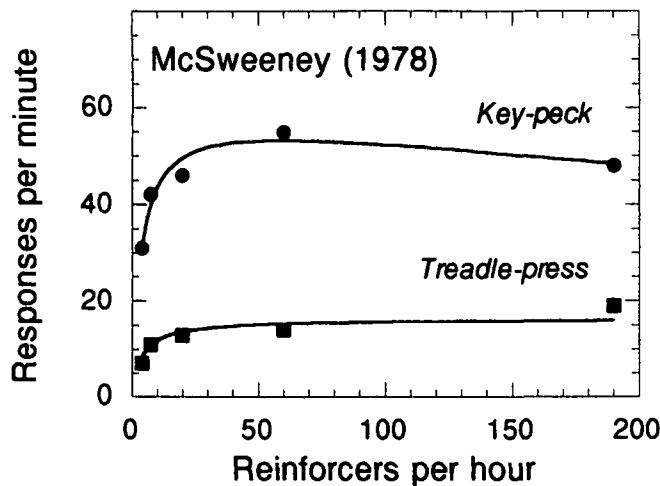


Figure 11. Response rates of pigeons reinforced on variable-interval schedules for key-pecking or treadle-pressing (top panel). Response rates of rats reinforced on VI schedules for depressing levers with different force requirements (bottom panel). The curves are from Equation 12, with parameters listed in Table 3.

### 5.6. Reprise: Schedule effects

The present theory gives us a more precise way of interpreting motivational changes, yet it is not insensitive to the variety of ways in which motivational operations can affect behavior. Motivational variables may affect parameters other than  $a$ : a highly motivated animal may live more in the present, weighting the most recent response more heavily (Killeen 1984; 1985), thus increasing  $\lambda$ . The data of Snyderman (1983; see also Fig. 10) show increases in both  $a$  and  $\lambda$  with deprivation level. Efficient response topographies may be more readily shaped in highly motivated subjects, resulting in smaller values of  $\delta$ . Coupling to food-relevant target responses may be better for highly deprived animals, thus increasing  $p$  (see below). But none of these adjustments are necessary to account for any data I have analyzed with the exception of Snyderman's.

Mapping response rates under interval schedules is not the most efficient way to test this theory or to estimate values for the fundamental parameters. Equation 12 and Figures 10–12 are offered to show that the theory is

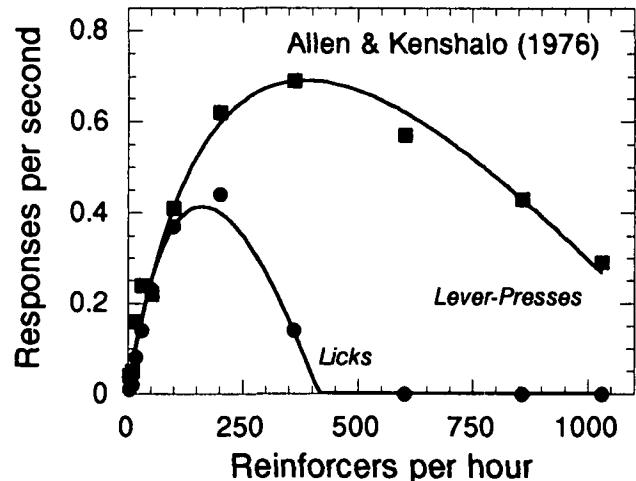


Figure 12. Response rates of a rhesus monkey on fixed-interval schedules. The squares depict the rate of lever-pressing for pellets as a function of rate of reinforcement; the circles depict the rate of schedule-induced water-drinking. The functions are from Equation 12.

consistent with these data and in fact converges on a standard model of them (Equation 4) at low and moderate rates of reinforcement.

Systematic differences in basic phenomena may be found under other experimental paradigms such as the "closed economy," which typically arranges very long duration reinforcers, low motivational levels, and extended sessions (e.g., Collier et al. 1992; Timberlake & Peden 1987; Zeiler 1991). In many cases these effects, as well as systematic within-interval decreases in response rates under more traditional conditions (McSweeney 1992), may be accommodated by letting the value of  $a$  decrease within a session as hunger decreases.

The present approach holds that interval schedules do not differentially reinforce pausing; they nondifferentially reinforce *any* response that comes before a target response and occupies time (whether or not that happens to be another target response). This is a subtle point that bears repeating. It makes sense to say the opposite – that interval schedules reinforce long IRTs – only if the "premature" responses debase the reinforcement of the target response class because they require effort, or because they lower the probability of reinforcement for members of that class. Conversely, the present theory assumes that interval schedules command rates lower than ratio schedules, because they give the memory of *everything* that comes before a reinforced target response (including, of course, other target responses and other stimuli) the substantial weight of  $1 - \beta$ . It is because of this promiscuity of reinforcement that we see lower rates of responding, and the drifting of those rates, on interval schedules. It is now time to examine the nature of the coupling constant,  $p$ , which carries the burden of this distinction between ratio and interval schedules.

### 6. Preasymptotic behavior: Trajectories in response-space

To understand how behavior shifts from the initial acquisition phase to the established patterns of responding

described by the theory, it is necessary to explore the conditioning process in more detail. What follows is one of many possible model systems, introduced as grounding and as a plausibility proof for the asymptotic model, which is the primary focus of this article.

Figure 13 shows a slice through "behavior-space" in which successive instances of the target responses are plotted on the  $y$ -axis as a function of all other responses on the  $x$ -axis. Upon each tick of an internal clock the subject emits some response, which moves along the tip of the trajectory either to the right or up. In this figure the probability of a target response is initially set at 0.1, and the line labeled "0" shows the resulting average trajectory before any conditioning has taken place. As expected, nine unmeasured responses are emitted for each target response we record.

Under FR schedules reinforcement occurs as soon as the trajectory crosses the ordinate corresponding to the ratio value  $N$  (the "consequential line," set at  $N = 16$  in Fig. 13). The learning model invoked here assumes that upon reinforcement the probabilities of all events in the trajectory ( $p_i$ ) are increased some proportion ( $w_j$ ) of the distance to their maximum ( $1 - p_i$ ):

$$p'_i = p_i + w_j (1 - p_i), \quad (13)$$

where  $p'_i$  is the updated probability of emitting a response in position  $i$  of the sequence. The proportion  $w_j$  is simply the weight of the item in memory (Equation 7), which is:

$$w_j = \beta(1 - \beta)^{j-1}. \quad (14)$$

Here,  $j$  is 1 for the last (reinforced) response, 2 for the penultimate, and so on, being indexed by *any* response the animal makes: in the case of a sequence of length  $L$ , its value is  $j = L - i + 1$ .

Equation 13 does not tell us which response occurred and was strengthened in position  $i$  of the trajectory. For simplicity, rather than index that separately I lump all nontarget responses together in one category and write their probability as  $1 - p_i$ . Then Equation 13 may be

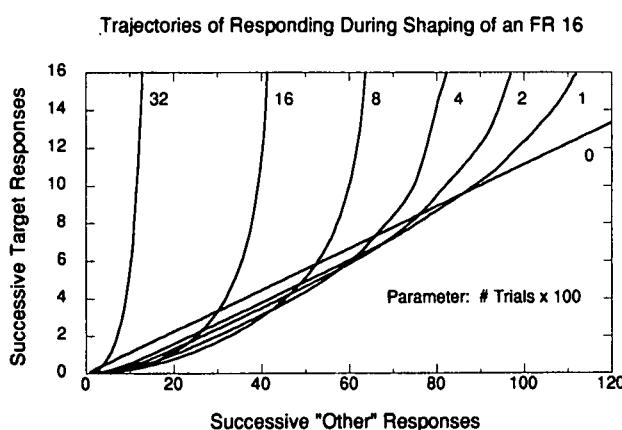


Figure 13. The number of target responses emitted as a function of the number of other responses during conditioning of an FR 16 schedule. The data are from simulations in which the probability of whatever is in memory at the time of reinforcement is increased according to a linear-operator learning model (Equation 15), with the learning rate an exponentially decaying function of the distance of each element from reinforcement (Equation 14). The initial probability of a target response and  $\beta$  were 0.10 in this and in all other simulations.

rewritten to cover both measured and unmeasured responses:

$$p'_i = w_j X_i + (1 - w_j)p_i, \quad (15)$$

where  $p'_i$  is the updated probability of a target response in position  $i$ ;  $w_j$  the weight in memory of any response that is  $j - 1$  elements away from reinforcement;  $X_i$  is 1 for a measured response, 0 for any other response; and  $p_i$  is the prior probability. Note that this model treats measured responses symmetrically with all other responses, for on trials when a nontarget response occurs in position  $i$ , its probability will be increased by exactly the amount that the probability of a target response is decreased. Equation 15 instantiates Equation 2 for each element of the trajectory. It is one of the oldest and most basic mathematical models of the learning process.

What happens when this model is turned on? Intuitions vary: because reinforcement increases whatever sequence of responses the animal emits, and because it is most likely to be doing something other than the target behavior, one might assume that the process would drive the target responding into extinction, but this does not happen. Figure 13 shows the average trajectories after successive blocks of 100 reinforcements. The trajectories walk themselves into a near-vertical position, indicating that most of the behavior now consists of the target response.

Why should this have happened? Whenever a trajectory crosses the consequential line, all responses that comprise it are strengthened (as a decreasing function of their distance from that line). Consider the response that occurs after the fifteenth target response on an FR 16 schedule. If it is also a target response, it is strengthened by the amount  $w_1 = \beta$ ; if it is some other response followed by a target response, it is strengthened by an amount  $w_2 = \beta(1 - \beta)$ , which is less. This marginal bias toward the target response exists at each step back in the sequence. The other response may be followed by a string of other responses before the last target response, further decreasing its weight in memory. To calculate its exact value at any step in the sequence we must sum over all possible histories between that point and completion of the ratio. In Appendix D, I demonstrate how this is done, showing that the response space is a field with gradients at each point oriented toward the consequential line. It is a dynamic field, because the strengthening given to any element both depends upon and determines the probability of making a target response. Under FR contingencies the trajectory evolves toward the leftmost curve in Figure 13. If the probability of making a target response at every opportunity to do so were 1 then the final state would be a vertical line.

For purposes of analysis we have treated each element in the trajectory as a separate response. Our equipment typically counts all target responses as members of the same class wherever they occur in the trajectory; the coupling associated with that class is their weighted sum, with the last element receiving a weight of  $\beta$ , the penultimate one a weight of  $\beta(1 - \beta)$ , and so on. The sum of those weights is the sum of Equation 14 up to  $N$ , that is,  $1 - (1 - \beta)^{N-1} \approx 1 - e^{-\lambda N} = \zeta$ . Thus, the limiting model for ratio schedules (the vertical trajectory) is one toward which the reinforcement process naturally attracts behavior and it is the one given by Equation 8 in the text (with  $\rho = 1$ ).

It is of course unrealistic to think that each element in a trajectory constitutes a unique response that has its own memory register. Therefore, in the present simulations, after each reinforcement the probability at position  $i$  is assigned the average value for it and the ones immediately before and after it. This averaging provides a minimal kind of "coarse graining" that smooths the results of the finite-elements analysis. The lumping of behavior in real organisms is substantial, with bouts of responding that are self-sustaining (i.e., the probability of continuing in some class of responses is greater than that of leaving it). But such verisimilitude was sacrificed for simplicity; insofar as the present model can reflect the trapping effects of reinforcement and give a correct ordering of coupling constants for the various schedules and for parts of them, more realistic models involving additional parameters can do so a fortiori.

When the FR is large, the consequential line is so much higher than the origin that reinforcement cannot reach all the way back to attract the earliest elements but leaves them to drift around their baseline level. The smoothing causes food-oriented responses during the prior reinforcer (at  $i = 0$ ) to propagate forward in the sequence, causing the bowing of the trajectories below the original baseline probabilities (see Fig. 13). The coupling on the simulated schedules is the proportion of responses in the average trajectory that are target responses. Figure 14 shows this proportion for the trajectory as a whole and for the last half of it. The curves through the points are the locus of Equation 8, multiplied by the coupling constant ( $\rho$ ), which tells us how far the trajectory has rotated to the vertical. For the rate during the last half of the schedule  $\rho = 0.9$ , but for the trajectory as a whole  $\rho$  takes the lower value of 0.6. Note that all increases in the proportion of target responses in the stream of behavior will be reflected in changes of  $\rho$ , which therefore provides a measure of learning. Predictions with  $\rho$  fixed at 1 are as

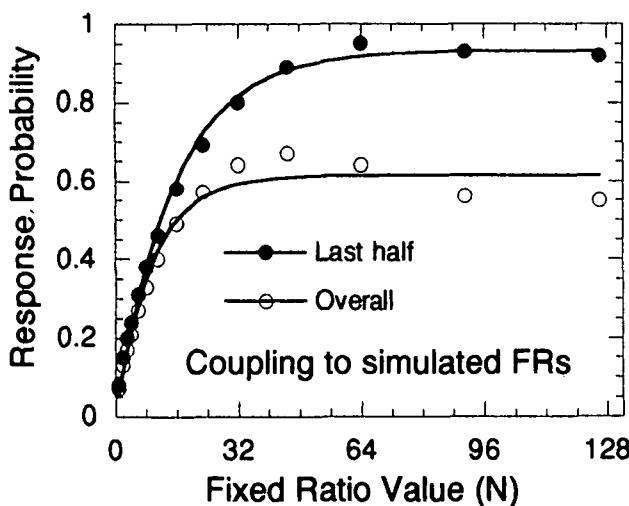


Figure 14. The probability of making a target response on FR schedules as a function of their length, based on the simulations. It is calculated as the ordinate at which the asymptotic trajectory crosses the consequential line divided by the length of the trajectory (ordinate plus abscissa; see Fig. 13) or, for running rates, from the appropriate segments in the last half of the trajectory.  $\rho$  is the asymptote of the functions drawn through the data as  $N \rightarrow \infty$  (line 2 of Table 5).

accurate as those using the precise value, because  $\rho$  is a factor of  $\delta$ ; the only effect of freezing  $\rho$  at 1 is to inflate the recovered value of  $\delta$ . Further discussion of the nature of  $\rho$  and the dynamics of the other basic schedules is found in Appendix D.

In summary,  $\rho$  is the asymptotic proportion of target responses in the trajectory as  $N \rightarrow \infty$ ; ratio schedules drive its value toward 1, interval schedules to lower values. In general, however, it may be fixed at 1, letting subunitary values due to schedule type or incomplete conditioning be accommodated by larger values of  $\delta$ .

## 7. From interval to ratio: Concatenating zetas

The reinforcement of interresponses may be tempered by adding ratio contingencies that require a certain number of target responses to be made after some interval has elapsed. As  $N$  increases, the contingencies increasingly strengthen target responses rather than interresponses, driving rate toward its maximum. Figure 15 shows the predictions for  $N$  ranging from 1 (i.e., a simple VI) to 12, added to a variable-time (VT) schedule. The data are from Catania (1971) and the curve from a concatenation of the coupling coefficients that constitute these contingencies:

$$\zeta = \zeta_{\text{FR}=1}^N + \zeta_{\text{VI}=N}^{BT} \quad (16)$$

in particular:

$$\zeta = (1 - e^{-\lambda N}) + e^{-\lambda N} \left( \frac{\rho \lambda B}{\lambda B + R} \right). \quad (17)$$

Now for the first time we may no longer absorb  $\rho$  into the maximum response rate; the data require that it take a value of  $\rho = 0.34$  for the interval constant. (The other parameters are  $\lambda' = 0.25$ ,  $\delta = 0.38$ , and  $a = 450$ , all in line with the values assumed in other experiments.) The point is not that a curve can be found to fit the data – which could be done more simply than this – but rather to demonstrate how a coherent theory of schedule control may account for transitions between different types of schedules in a principled fashion.

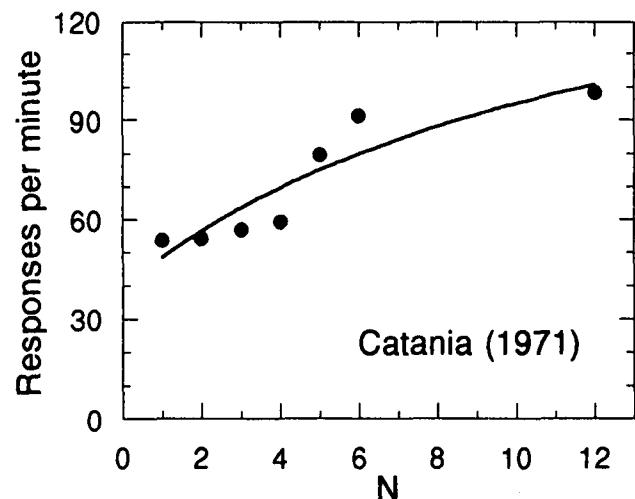


Figure 15. Terminal response rate as a function of the FR component of a tandem reinforcement schedule. The curve is from Equation 17.

We may debase the contingency by going in the other direction, following a ratio requirement with a delay. As the delay is lengthened, concatenation of the appropriate coupling coefficients predicts a smooth decrease in response rates, as was observed by Catania and Keller (1981).

The success of Equation 16 is promising. Properly written coupling coefficients may be concatenated in ways that correspond to the contingencies of reinforcement: averaged for mixed schedules and sequenced, as in Equation 16, for tandem (series) schedules. The maximum of two coefficients would be appropriate for "conjoint" (parallel) schedules, as the contingencies most tightly coupled to behavior would drive responding in the direction in which they would be even more effective. Estimating the coupling on multiple schedules requires considering the reinforcing strength of noncontingent component changes and how they couple to both responses and interresponses, as well as local changes in the value of  $a$ . Similar issues arise for concurrent schedules (parallel contingencies for different responses), which also require a generalized model of temporal constraint.

The development of techniques for combining coefficients may lead us to a unified system of models that is as useful in describing the consequences of reinforcement schedules as the Laplace transform is in describing the arrangements of linear systems. Conversely, the theory can guide us in designing contingencies that are resonant with the operating system of the organism, replacing the old designations of schedules with ones that are based on the design principles of the organism.

## 8. Curtailing the response

The minimum IRT for both pigeons and rats is around one-quarter second; pigeons often make double pecks, functionally increasing the duration of a "response" to about 0.6 sec (Arbuckle & Lattal 1988). Modern computer technology makes it possible to reinforce animals within a few milliseconds of their first contact with the operandum, and thus to intercept their stream of behavior before a response has fully run its course – before the nascent response has fully entered memory. To the extent that the experimenter's definition of the response is instantaneous, so that it is reinforced in the first instant of its occurrence, coupling approaches zero (see Equations 18 and A5). It should not be surprising then that schedules that delay reinforcement by one-half sec from the onset of a response generate higher rates of responding (Arbuckle & Lattal 1988). If reinforcement is precipitate, it will not couple to behavior. In conditioning, as in humor, timing is of the essence.

Rather than treat the weight given to a response as constant, we should take:

$$\beta = w_1 = 1 - e^{-\lambda' \delta / \rho}, \quad \rho > 0. \quad (18)$$

In the first two experiments we essentially used Equation 18 with  $\delta/\rho$  set equal to 1; that is, we treated all responses alike (i.e., as quanta). Subsequent analysis belied this quantal assumption, because the rate of memory decay was highly correlated with  $\delta$  and with schedule type, leading us to assert  $\lambda = \lambda' \delta / \rho$ . Whereas  $\delta$  tells us the minimal IRT,  $\rho$  tells us how many interresponses come

along for the ride (viz.,  $1/\rho - 1$ ), thus functionally increasing the duration of the target response (or, equivalently, increasing the rate of memory decay when measured in terms of target responses alone). The appearance of  $\rho$  in this equation captures our assumption that all responses, not just target responses, index memory. Catania's (1971) data were consistent with a decreased weight for the last response, as shown in Figure 6.

Traditional contingencies of reinforcement thus not only truncate the memory of responding, they may also curtail the manifestation of the target response, in both ways undermining their own control of behavior. Skinner (1938, p. 40) thought that our definition of the unit of behavior was correct when it yielded "orderly dynamic changes" in behavior. But this is not enough. All of the curves in Figure 3 are relatively "orderly." We must look for a *resonance* between our definition of the response and the subject's, as indicated by optimal control and maximal rate of learning, as seen in the tuning curves in Figure 5. The current analysis suggests that our criteria for reinforcement should take into account an exponentially weighted history of responding that epitomizes the animal's memory for its own behavior; it should follow the onset of the most recent response no sooner than  $\delta$  sec. Performance on the various schedules of reinforcement is a straightforward consequence of the extent to which our contingencies of reinforcement respect the animal's memory and manifestation of its behavior.

## 9. General discussion

### 9.1. Synopsis

The basic principles underlying this calculus of contingencies are represented in Table 4. It will be useful to review them before a more general discussion.

Line 1 gives the linear average that governs the decay of memory from one response to the next. The currency parameter,  $\beta$ , characterizes the weight given to the modal target response. Memory does not decay when the organism is not responding.

Line 2 tells us the rate at which responses (both target and other responses) are elicited by incentives occurring at a rate of  $R$ . This instigating force for behavior is proportional to the historical rate of delivery of incentives in the experimental context (Killeen et al. 1978). One incentive under the specified conditions of motivation will support  $a$  seconds of behavior –  $a/\delta$  responses if each requires  $\delta$  seconds for its completion. The tilde over the measure of behavior indicates that it is a propensity to respond, not yet a prediction about observed behavior.

Line 3 formalizes the observation that responses take time, which detracts from the time available for the emission of other responses. A blocked-counter model (Bharucha-Reid 1960) provides the basic compensation for this depression of observed rates below their "theoretical" level, while line 2 provides that theoretical level. Appendix B gives the derivation plus supportive data. The cap over the measure of behavior reminds us that we have now moved to an empirically measurable response rate, but we have not yet integrated these first two principles along with the third to yield the final prediction.

Line 4 derives from the combination of lines 2 and 3,

Table 4. Fundamental equations

Construct	Equation		Equation no. in text
1. Memory	$M_n = \beta y_n + (1-\beta)M_{n-1}$		(2)
2. Activation	$\tilde{B} = \frac{a}{\delta} R$	$\delta > 0$	(B1)
3. Ceiling constraints	$\hat{B} = \tilde{B}/(1 + \delta\tilde{B})$		(B3)
4. Behavior	$B = \frac{\zeta R/\delta}{R + 1/a},$	$\delta, a > 0$	(4; B7; C19)
5. Currency parameter	$\beta = 1 - e^{-\lambda'\delta/\rho}$	$\delta, \rho > 0$	(18)

and adds the coupling coefficient zeta ( $\zeta$ ) to complete the model of the contingent control of responding. This is the "canonical" equation: to predict response rates one specifies how a particular schedule couples incentives to the organism's memory of its behavior, and inserts the appropriate expression for  $\zeta$  into Equation 4. Prototypical expressions for  $\zeta$  are given in Table 5.

Where rate of reinforcement ( $R$ ) varies substantially with rate of responding, this must be taken into account by using an appropriate schedule-feedback function. In this target article I have assumed that reinforcement rate is proportional to response rate under ratio schedules and independent of it under interval schedules. It is, of course, always possible to measure the obtained rate of reinforcement and use Equation 4 directly, as is necessary when the schedule-feedback functions are obscure. This is possible because nowhere in this theory do I assume that the schedule-feedback function controls behavior directly (i.e., this is not a "molar" optimization theory); it only does so indirectly, through the rate of reinforcement it sustains.

Line 5 gives the currency parameter a derivative position in the theory: it is the cumulative weight of reinforcement that is brought to bear on a response of duration  $\delta$ , when the measured target response comprises the pro-

portion  $\rho$  of the total behavior stream and the rate of decay of memory while the animal is responding is  $\lambda'$ .

Prototypical coupling coefficients are derived in Appendix C and displayed in Table 5. These are inserted into the canonical activation-constraint equation to predict behavior. Line 1 (and Fig. 5) gives the coupling coefficient for the weighted reinforcement of IRTs in Experiment 2. The remaining lines give the coefficients for responding under some traditional reinforcement schedules. Because  $B$  appears on both sides of the equations for interval schedules, these equations must be iterated until the predictions converge. However, for an approximate solution we may use a truncated power series to represent the exponential term, yielding a general model for interval schedules, line 8. This permits approximate solutions to the basic equations that are much simpler and whose average predictions are usually within 1% of those afforded by the precise equations.

For most of the traditional types of scheduling arrangements, the response-interresponse proportion,  $\rho$ , is too strongly conflated with the minimum response time,  $\delta$ , to be independently estimated. Making a virtue of this necessity I set  $\rho$  to 1, letting  $\delta$  absorb the residual variations in its value. This maneuver permits predictions for all basic schedules based on the same three parame-

Table 5. Coupling coefficients for representative schedules

Contingency	Coefficient <sup>a</sup>		Equation no. in text
1. Expts. 1 & 2	$\zeta = 1 - \left( \frac{\alpha}{\beta} \right)^{\frac{\beta}{\beta-\alpha}} + \left( \frac{\alpha}{\beta} \right)^{\frac{\beta}{\beta-\alpha}}$	$\alpha > \beta > 0$	(A5)
2. FR $N$	$\zeta = \rho(1 - e^{-\lambda N})$		(8; C6)
3. VR $N$	$\zeta = \frac{\rho\lambda N}{1 + \lambda N}$		(C17)
4. FT $T$	$\zeta = \rho(1 - e^{-\lambda BT})$		(C21)
5. FI $T$	$\zeta = \beta + \rho(1 - \beta - e^{-\lambda BT})$		(C24)
6. VT $T$	$\zeta = \frac{\rho\lambda BT}{1 + \lambda BT}$		(C29)
7. VI $T$	$\zeta = \beta + (1 - \beta) \left( \frac{\rho\lambda BT}{1 + \lambda BT} \right)$		(C31)
8. All interval schedules	$\zeta \approx \frac{\rho\lambda BT}{1 + \lambda BT}$		(11; C22)
9. Inhibit (DRL) $\tau$	$\zeta = \beta(\beta + e^{-\lambda\tau})$	$\tau > \delta > 0$	(C35)

<sup>a</sup> $\beta = 1 - e^{-\lambda}; \lambda = \lambda'\delta/\rho$ .

ters. The resulting general forms for the various schedules are presented in Table 6.

Under a few scheduling arrangements we are permitted or forced to assign a real value to  $\rho$ , as in the last study analyzed in this target article. Then it takes a value of about  $\frac{1}{3}$ , which is consistent with independent estimates of the efficacy of ratio schedules relative to interval schedules (Zuriff [1970] estimates that each reinforcer on variable-ratio schedules will generate  $2\frac{1}{2}$  to 3 times the control of measured responding as on variable-interval schedules).

The cumulative memory of a sequence of incentives generates a heightened state of arousal whose asymptote is given by line 2 of Table 4. This arousal can become associated with the target response, with other responses, with stimuli, and with the experimental context; the degree of these associations is given by the appropriate coupling coefficients, yielding summative models such as those shown in Table 6. The role of such activation – the memory of incentives past – in instigating behavior was introduced as a fundamental assumption of the present theory. On returning to it we now see it as another manifestation of the operation of short-term memory, obeying the same rules of saturation and displacement but operating on emotionally salient stimuli. The passage of incentives through short-term memory increments a longer-term state of activation that elicits diffuse behavior; their cooccurrence in memory with other stimuli and responses directs the association of that arousal, a process we call *conditioning*.

Because the constituent processes involve changes in behavior under the force of incentives, I refer to this ensemble of principles as a *mechanics* of behavior, with the asymptotic performance models of section 5 providing the *statics*, and the learning models of section 6 and Appendix D the *dynamics*.

## 9.2. Degrees of freedom

There are three central assumptions of this theory: the activation of behavior by incentives, constraints on responding, and memory as the mediator of reinforcement. An important ancillary assumption is the continual indexing of memory by responses, including observing and consummatory responses, which displace the memory of

target responses. These assumptions may be captured by models in different ways; this article displays but one of many possible instrumentations. I have omitted certain considerations to spare the reader too heavy a burden of detail (e.g., the amount of behavior supported by an incentive, line 2 of Table 4, should also take into account the activation allocated to unmeasured responses by including  $\rho$  as a divisor of  $a$  wherever it occurs; however, the only place such detail would have an important impact is on the reader's patience). Further application of the theory to data will reduce the field of candidate implementations, modify and expand the contents of Tables 2–6, and refine some of the implementations chosen here. There is a vast amount of data available for this task. The present theory motivates well-defined alternative models and their selection by the likelihood ratios of their predictions given that data (Sakamoto et al. 1986).

Important ways in which the theory must be developed include: specification of the coupling coefficients for concurrent and multiple schedules; predictions of the variance in behavior associated with nonlinearities in the governing equations; integration with research on short-term memory for sequences of stimuli; contact with the extensive literature on associative conditioning; extension to memory of sequences of heterogeneous stimuli and responses; treatment of the conditioning of stimuli in a consistent manner; and grounding the constructs in the pharmacological and biobehavioral literature.

## 9.3. Relation to other theories and phenomena

**9.3.1. Conditioning at long delays.** Conditioning is best with short delays between a response and reinforcer, yet an impressive amount of conditioning can occur at long delays. According to our event-indexed account of memory, the passage of time between a response and a reinforcer is irrelevant; it is the occurrence of other responses that block control by filling memory with unrecorded behavior. Lattal and Gleeson (1990) showed conditioning with delays of 10 to 30 seconds in rats and pigeons and suggested that "in relative terms, this interval may be quite immediate" (p. 38; see also Wilkenfield et al. 1992). According to our account, it is relative – relative to the total amount of other behavior that fills memory. And that is driven by overall rate of incitement, causing memory to

Table 6. Exemplary models

Schedule	Predicted response rate	Equation no. in text
Ratio schedules	$B = \frac{\xi}{\delta} - \frac{N}{a},$ $B = 0$	$N < a/\delta; \delta, a > 0$ ( $\xi$ given by lines 2 & 3 in Table 5)
Interval schedules	$B = \frac{kR}{(R + 1/a)} - \frac{R}{\lambda},$ $B = 0$	$k = \rho/\delta$ $\lambda, a, \delta > 0$ $R < \lambda k - 1/a$
		(12) elsewhere

"lengthen" in unstimulating environments (when measured in temporal units; Killeen 1984; 1991; Williams 1978). Ours is an "ordinal contiguity" account, in that it is the number of times that memory is indexed that discounts a reinforcer's control of a response, and that number depends on the richness of the environment and the activity of the organism.

Some types of conditioning such as taste aversion may occur over a span of hours, not fractions of a minute. Revusky and Garcia (1970) explained these effects in terms of an interference theory:

On a practical level, as the time between a stimulus and a consequence is increased, the probability of a learned association between them will be reduced. . . . On a theoretical level, however, it is quite likely that it is not the increased time itself which interferes with the learned association but the fact that an increased duration of time is likely to contain an increased number of interfering events. (p. 41)

Events relevant to the consequence interfere the most: Revusky and Garcia cite hours-long delay gradients for operant responses when these are "protected" from interference by a change of context, such as removal from the experimental chamber. Responses that are "marked" by stimulus change (Lieberman et al. 1985) or conditioned reinforcers also become more memorable.

This target article addresses only distance along the dimension of homogeneous operant responses. Similar processes of coupling may occur simultaneously along all dimensions of the organism's psychological space. I have argued in general that the force of an incentive decreases exponentially along those dimensions (Killeen 1992). The present paper constitutes a working out of the implications of that theory for one exemplary continuum.

**9.3.2. Response selection.** The relative excitability of various actions, as reflected in their coupling constants, governs their prevalence in the repertoire and thus their availability to memory. Contingent reinforcement adds a guaranteed weight of  $\beta$  to the target response, but this may not be adequate to compete with a more "prepared" response having an intrinsically larger value of  $\rho$ . The availability of responses with larger coupling constants may lead to "instinctive drift," with the topography of the target behavior degenerating toward that of a more easily excited form, even when that is less effective in obtaining reinforcement (Breland & Breland 1961). The gradients that attract the target response are often quite attenuated and may provide weak competition for paths through instinctively prepared responses.

When reinforcement contingencies may be satisfied by a variety of discrete acts (Reed et al. 1991) or topographical variants (Stokes & Balsam 1991), the history and genetic predispositions of the organism will favor some variants; their presence in memory at the time of reinforcement will further strengthen them and lead them to dominate the repertoire even after the experimental constraints are relaxed (Davis & Platt 1983). "If certain species-specific behaviors happened to occur prior to the arbitrary response, they would become more strongly reinforced than the arbitrary response although the delay of reinforcement was longer" (Revusky & Garcia 1970, p. 49; see also Timberlake 1983). These ideas were formalized by Staddon and Zhang (1989), whose model of re-

sponse selection was "based on assumptions [of]: arousal and adaptation, the idea that reinforcement transiently energizes a range of activities . . . ; strength and competition, the idea that each activity has a certain tendency to occur and that the strongest will win; and variability, the notion of a repertoire of activities" (p. 190). They invoked an exponentially weighted moving-average of differences in short-term memory for different responses (with a parameter akin to  $\beta$ ) and differences in excitability of various responses (with a parameter akin to  $\rho$ ).

Inferences concerning associative ability depend on how well contingencies of reinforcement respect species differences in depth of short-term memory and response excitability. Specification of the optimal value of  $\alpha$  must be a precondition for comparisons of abilities across species – that is, a precondition for a comparative psychology of learning [see also Macphail: "The Comparative Psychology of Intelligence" *BBS* 10 (4) 1987].

Skinner (1938), among others, noted that "when a reinforcement depends upon the completion of a number of similar acts, the whole group tends to acquire the status of a single response" (p. 300). The present theory is a latter-day response-unit hypothesis, with exponentially decaying weights for the constituent acts. Our model does not require that the acts be similar, however, only that they be in memory at the time of reinforcement. Contingencies that provision memory with a mixture of acts – some measured and others merely filling time – will strengthen that mixture; but only a fraction ( $\rho$ ) of those acts will be measured and contribute to the coupling of the target response to reinforcement.

**9.3.3. The correlation-based law of effect.** Learning theorists often counterpose contiguity (closeness of two events) and contingency (relative frequencies of one event in the context of another, often treated as a correlation). Thus Baum (1973) has argued for a "correlation-based law of effect" in which the units of analysis are reinforcement rates and response rates whose correlation over stretches of time constitutes the law of effect; in his theory, correlation is more important than contiguity. Williams (1976) challenged that theory by finding substantial decreases in response rates in schedules that maintained a correlation over delays of 3 to 15 sec while disrupting contiguity.

The present theory cuts across this distinction, as zeta ( $\zeta$ ) is a contiguity-weighted correlation coefficient (see Figs. 1F and A1) and thus shares features of both accounts. But it correlates the contents of memory on the occasion of reinforcement and the experimenter's requirements, not rates of responding and rates of reinforcement. The present theory is response-indexed, Baum's (1973) time-indexed. There is less emphasis here on schedule-feedback functions as their only role is to control rates of reinforcement and contents of memory; in particular, it is not assumed that organisms respond faster on ratio schedules *because* they are correlated with higher rates of reinforcement. These important but technical differences aside, the present work may be seen as a realization of Baum's early vision.

Williams (1976) noted that one of the reasons for the correlational approach had been that "As of now, the reduction of free operant behavior to molecular principles has been notably unsuccessful" (p. 442; see also Papini &

Bitterman 1990; Thomas 1983). The present account is just such a reduction – albeit one in which correlation plays a key role.

**9.3.4. Foraging.** The exponentially weighted moving-average model has also appeared in the behavioral-ecology literature on optimal foraging (for a review, see Kacelnik et al. 1987). In some experiments, however, there is little or no evidence for the control of behavior by any of the responses preceding the most recent (e.g., Cuthill et al. 1990). The “responses” in question often comprise extended episodes of travel, however, and thus, according to our present analysis (e.g., Equation 18) the most recent may have washed the penultimate target responses out of short-term memory.

McNamara and Houston (1987) have shown that to forage optimally, the currency parameter for an animal’s memory of reinforcement should vary with the rate of reinforcement in the environment as a whole, giving more weight to the most recent events in rich environments. [See also Houston & McNamara: “A Framework for the Functional Analysis of Behavior” *BBS* 11(2) 1988.] There is good evidence that this happens. Yet we have not explicitly specified such a dependency of  $\lambda$  on rate of reinforcement in our theory. Does this make it both nonoptimal and inconsistent with the data? No: the present theory entails an implicit dependency of rate of memory decay on rate of reinforcement because memory is indexed by responses and their rate depends directly on the rate of reinforcement (see Table 4). Because this dependency is nonlinear, curved by ceiling constraints, it has been easier to develop the theory in the response domain rather than the time domain. Furthermore, there are times when the animal is not responding; we expect no changes in memory then. Nonetheless, it is clear that as rate of incitement increases, so also will rate of responding, causing the depth of memory measured in seconds to shorten – as it must, to satisfy considerations of optimality in both foraging decisions and causality detection.

**9.3.5. Memory.** No one has argued more strongly than Shimp (1976a; 1976b) for a reevaluation of our treatment of behavior in terms of its structure, an extension of our definition of the response to sequences of acts, and a recognition that this is tantamount to introducing “memory” into the experimental analysis of behavior (e.g., “it makes sense to ask what relations obtain between what a subject can remember of its own recent behavior in short-term memory experiments . . . and what behavioral patterns can be established as behavioral units when those patterns systematically precede a reinforcer,” Shimp 1976a, pp. 125–26; cf. Branch 1977). The present theory suggests that the relation is one of identity. Responses not remembered – whether because of displacement or inattention – are not reinforceable.

The present theory has much in common with the “standard models” of human short-term memory (Atkinson & Wickens 1971; Deutsch & Deutsch 1975; Norman 1970); it is an interference theory, as many of those are. In both, the weighted contents of short-term memory are transferred to a long-term store (Equation 15). There is even a close similarity between the mathematical models derived from both traditions.

The exponentially decreasing memory is consistent with the demands of a temporal lens focused on the most likely causal candidates preceding an important event. Potential candidates ramify exponentially as time passes and their a priori weight in memory must decrease in like manner. The present models are simple because they take homogeneous response sequences as their domain. Qualitatively different stimuli and responses will be weighted and averaged into their own long-term stores, perhaps in a manner similar to the ordinarily tagged responses sketched in Equation 15. Distinctive stimuli may differentiate otherwise similar responses: animals may remember the last responses they made in a white alley without interference from similar responses in the home cage or in dark alleys and may cumulate these tagged memories over substantial temporal interludes (Capaldi 1992). This gives rise to a discrete-trial version of the bitonic function (e.g., Fig. 7) known as the “partial reinforcement acquisition effect.”

The present theory treats short-term memory as a low-pass filter between responses (and stimuli) as input and behavior (and long-term memory) as output. It may be instantiated as a neural field in which an input attracts a proportion  $\beta$  of the available resources. If the activity on the field is normalized by lateral inhibition, then the attention allocated to the new input is found by decrementing all the other representations by  $1 - \beta$ . If the decrement is random, those with the largest representation will be taxed in proportion to the resources they command. This yields the necessary geometric/exponentially weighted decay of memory. When an incentive enters memory, its contents are averaged into their appropriate long-term stores – including the memory of the incentive, which engenders a cumulating increase in activation level. The activation may involve a different neural locus (e.g., the ventral striatum) from the short-term memory of other stimuli and responses (e.g., the dorsal striatum; Robbins & Everitt 1992) and may thus have different rate-constants associated with it. Rawlins (1985) characterizes the hippocampus as a temporary memory store with properties similar to those inferred here. Deadwyler (1985) enlisted electrophysiological data that showed that the dentate gyrus functions as an updateable buffer with the weighting of items decreasing as an exponential function of the number of more recent items in the buffer. Jarrad (1985) argued that it was the number of interpolated items, rather than the passage of time, that differentially affected animals with hippocampal lesions. These notes should by now sound familiar; the mathematical principles of reinforcement developed here may provide both a useful tool for physiological analysis and a functional module for the representation of short-term memory in more complete neural networks, such as the Sutton-Barto-Desmond model (e.g., Moore 1991), Donahoe and Palmer’s (1993; Donahoe et al. 1982), and others (e.g., Grossberg 1975; Levine 1991).

## 10. Conclusion

We reinforce the animal’s representation of its behavior, not the behavior itself. This representation is a weighted memory of past responses. The present theory places the representation into a dynamic system of equations for

predicting the effects of reinforcement contingencies. It is a nonmetaphorical theory, in that each of its key terms is well defined in ordinary language. It is constructive, in that it permits explicit predictions, just as it provides elbow room for reconstruction of the particular instantiating models in the light of data. It posits no mental algebras, only excitement, the diffuse activation of behavior, and the focusing of that incitement on specific responses by the correlation between experimental demands and the organism's memory at the time of those demands. It is a theory of how incentives fuel behavior, time constrains it, and contingencies direct it. It is not a "cognitive" theory, as we have come to know them; rather it brings us a step closer to understanding cognitive processes such as short-term memory in behavioral terms.

#### 11. APPENDIX A: The correlation between the experimenter's and the organism's definitions of a response

In this appendix I derive the smooth curve through the data shown in Figure 5. I call the ordinates of this curve the *coupling* between the reinforcer and behavior, whose value I designate as zeta ( $\zeta$ ). Zeta is the proportion of the weight of incentives that is brought to bear on target responses by the contingencies of reinforcement.

Assume that we use a definition of the response that decays more quickly than the animal's memory of its behavior (as is the case if we count only the last response or interresponse time, or in general when  $\alpha > \beta$  in Equations 2 and 3). Then we have a situation such as that pictured in Figure A1. Up to the point that the curves cross, the limiting factor on conditioning is the animal's memory for its behavior, which is spread over a long

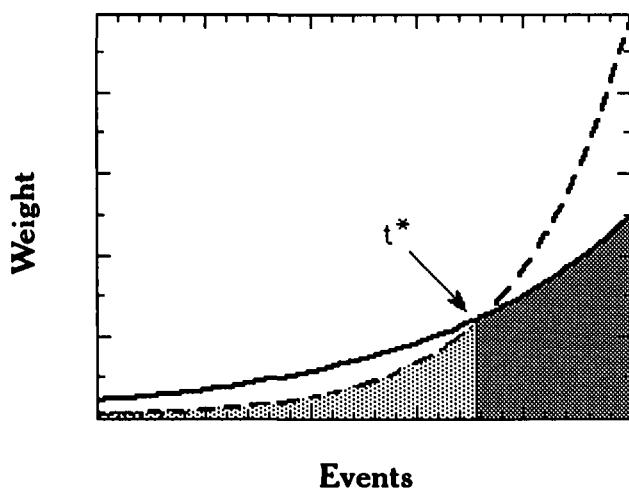


Figure A1. The experimenter's definition of the response (dashed curve) versus the organism's (continuous curve). The gray area to the right of the crossover point,  $t^*$ , represents the proportion of the animal's recent response memory that is constrained by the experimenter's definition; in this case, the constraint is complete. However, the experimental control that was overallocated to the recent past (the area between the curves) is unavailable to constrain behavior more remote in time or action. This results in underconstraint to the left of  $t^*$ , as indicated by the deficit of the experimental definition below the level of the animal's memory. The intersection of the area under these two curves, as indicated by the shaded areas in the figure, gives the correlation between the two definitions (Equation A5). In this case, its value is  $\zeta = 0.75$ .

time and not so concentrated in the recent past. After the curves cross, the animal remains more sensitive to older behavior than the experimenter. Thus, from this point back the limiting factor is the lower tail on the contingency curve. To ascertain the impact that a contingency characterized by a particular  $\alpha$  will have on an animal whose characteristic memory decay rate is  $\beta$ , we must calculate the shaded area in this figure.

The first step is to calculate that value of  $t$  at which point the curves cross,  $t^*$ . That happens when:

$$\alpha e^{-\alpha t} = \beta e^{-\beta t}, \quad (A1)$$

which occurs when:

$$t^* = \frac{\ln(\alpha/\beta)}{\alpha - \beta} \quad \alpha \neq \beta. \quad (A2)$$

Now we can integrate the shallower curve, the animal's memory, from 0 to  $t^*$ , and the steeper curve, the experimenter's definition, from  $t^*$  to  $\infty$ :

$$\zeta = \int_0^{t^*} \beta e^{-\beta t} dt + \int_{t^*}^{\infty} \alpha e^{-\alpha t} dt \quad (A3)$$

This gives:

$$\zeta = 1 - e^{-\beta t^*} + e^{-\alpha t^*}. \quad (A4)$$

Next substitute in the values for  $t^*$ , and simplify to obtain:

$$\zeta = 1 - \left( \frac{\alpha}{\beta} \right) \frac{\beta}{\beta - \alpha} + \left( \frac{\alpha}{\beta} \right) \frac{\alpha}{\beta - \alpha} \quad \alpha > \beta > 0. \quad (A5)$$

If  $\alpha \gg \beta$  the middle term goes to 1 and  $\zeta \sim \beta/\alpha$ .

When the experimenter's definition extends into the past farther than the animal's memory, the rate constants are exchanged to give:

$$\zeta = 1 + \left( \frac{\alpha}{\beta} \right) \frac{\beta}{\beta - \alpha} - \left( \frac{\alpha}{\beta} \right) \frac{\alpha}{\beta - \alpha} \quad \beta > \alpha > 0. \quad (A6)$$

If  $\beta \gg \alpha$ , the last term goes to 1 and  $\zeta \sim \alpha/\beta$ .

When the two processes are perfectly matched, the coupling is given by the area under either curve, corresponding to perfect coupling:

$$\zeta = 1 \quad \alpha = \beta. \quad (A7)$$

It is these equations that describe the coupling of incentives to behavior and generate the smooth curve through the data in Figure 5.

For mnemonic convenience, I have employed here the same Greek letters as used in the discrete Equations 2 and 3, but this is not quite correct. If  $\beta$  in Equation 2 goes to 1, the rate constant in the continuous process must go to infinity. I use the parameter  $\lambda$  for the continuous version of  $\beta$ , and remind you that  $\beta = 1 - e^{-\lambda}$ . The curve through the data in Figure 5 was generated using a value of 0.28 for the animal's memory rate in Equations A5 and A6, and this converts to an equivalent discrete  $\beta$  of 0.25. The relation between  $\beta$  and  $\lambda$  reaches its definitive form in Equation 18, where the memory decay rate is expressed in terms of its intrinsic rate, the duration of the response, and the proportion of the total behavior the target response constitutes:  $\lambda = \lambda' \delta / \rho$ .

The experimenter's definition of the response in Experiment 2 was an exponential decay function, as shown in Figure A1. The close map between the predicted coupling values and the data shown in Figure 5 provides important evidence that pigeons' short-term memory is also an exponential decay function.

## 12. APPENDIX B: The elicitation and competition models

### 12.1. Arousal

Incentives are the prime movers of organisms. Repeated delivery of incentives has a cumulative effect on the arousal level of the organism, motivating a rate of responding that is proportional to the rate of reinforcement. Although not unreasonable as an axiom, this assertion may be derived from other considerations. Staddon (1977) assumed that animals will not respond below a threshold probability of reinforcement; on schedules that deliver reinforcements randomly in time, this threshold is exceeded in proportion to the rate of reinforcement, yielding the required proportionality.

Killeen et al. (1978; Killeen 1982) showed that each delivery of an incentive excites an organism, with excitation decaying exponentially. Repeated delivery of incentives cumulates the activation, which becomes associated with the context. After many incitements the level of activation will rise in proportion to the rate of incitement,  $A = aR$ , with the constant  $a$  equal to the area under the decay curve from a single incentive.

To get from activation level ( $A$ ) to the theoretical rate of responses of duration  $\delta$ ,  $(\hat{B})$ , we divide by  $\delta$ :<sup>3</sup>

$$\hat{B} = aR/\delta. \quad (B1)$$

Throughout this target article I have used the term *incentive* to refer to a behaviorally potent stimulus that arouses the organism through the elicitation of consummatory responses, and *reinforcement* to refer to the conjunction of those consummatory responses (or the stimuli that elicit them) with the target responses or stimuli in short-term memory. Thus Equation B1 carries the burden of incentive motivation, whereas  $\zeta$  carries the burden of reinforcement. An aroused animal will be active, but without a consistent predictor of the incentive the activation will become diffusely associated with stimuli in the environment or with other responses from the animal's repertoire. Both activation and coupling must be above zero for instrumental "conditioning" to occur.

### 12.2. Response ceilings

Responses are not instantaneous. We may think of the experimental organism as a Geiger counter that receives impulses to respond at a rate of  $\hat{B}$ , but if a response is already in progress, the incipient response is blocked; this is a blocked-counter model. If we observe an empirical rate of  $\hat{B}$  responses per second for responses of  $\delta$  sec duration, then out of every second  $\delta\hat{B}$  sec are blocked. We may correct our measured rate by dividing it by the residual time available for responding,  $1 - \delta\hat{B}$ , to infer the theoretical rate that engendered it:

$$\hat{B} = \hat{B}/(1 - \delta\hat{B}). \quad (B2)$$

We generally wish to go in the other direction, predicting the measured rate from the theoretical, and so rearrange Equation B2 to give:

$$\hat{B} = \hat{B}/(1 + \delta\hat{B}). \quad (B3)$$

Note that the model is developed for the case in which the measured response is the dominant occupation of the animal. If there are substantial ongoing rates of other responses, such as adjunctive responses or concurrent operants, the denominator of Equation B2 must be expanded to take them into account.

So far the model involves no motivational assumptions. We add them by predicting  $\hat{B}$  from Equation B1 and rearranging to discover the traditional hyperbolic form:

$$\hat{B} = \frac{R/\delta}{R + 1/a}, \quad a, \delta > 0. \quad (B4)$$

This is Equation 4 in the text, with  $k = 1/\delta$  and  $c = 1/a$ . In the case of two competing responses it becomes:

$$\hat{B}_1 = R(1 - \delta B_2)/\delta(R + 1/a).$$

This correction for ceilings on response rate assumes that responses occur randomly in time, which is approximately true (see, e.g., Kintsch 1965), and that they are associated with a fixed refractory period,  $\delta$ , which is also approximately true. Palya (1992) successfully simulated a large database of pigeon IRTs with Bernoulli trials every  $\delta = 0.35$  sec (plus a small normal error), with a probability of a peck of 0.65 at each trial. This approximates the random responding assumed by the blocked-counter model and is consistent with the simulations in this paper.

There are other derivations of Equation B4 that do not involve response-rate ceilings. However, we know that there must be a ceiling on response rate, and correction for that should occur prior to implementation of other hypotheses. Furthermore, this derivation carries with it a testable prediction. Bharucha-Reid (1960) gives the variance of the number of counts in a time period  $t$  for this process as:

$$\sigma^2 = \hat{B}t(1 + \hat{B}\delta)^{-3}. \quad (B5)$$

From this we may infer that as rate of responding increases toward its ceiling, the relative variability of responding – the coefficient of variation – should decrease according to:

$$\frac{\sigma}{\hat{B}} = \frac{1 - \hat{B}\delta}{\sqrt{\hat{B}}}. \quad (B6)$$

McSweeney (1974) manipulated response rates by varying rate of reinforcement and deprivation level. Figure B1 shows that the predicted decrease in relative variability of responding is observed. The curve is Equation B6, with  $\delta = 0.30$  sec. None of the assumptions concerning motivation or coupling were involved in deriving Equation B6, and so its approximate correctness gives us some confidence in using the blocked-counter model as a building block of the theory.

Equation B4 is a model of activation-elicitation under response ceilings, not of contingent control. To be complete it requires that we introduce the coefficient of coupling that is

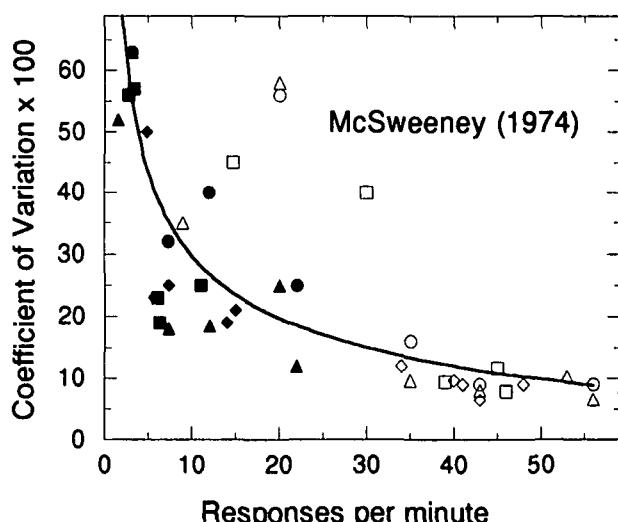


Figure B1. Relative variability of response rate (standard deviation divided by the mean) as a function of response rate. The filled symbols come from a VI 4-min schedule, the unfilled symbols from a VI 1-min schedule. Response rates were also varied by manipulating the percentage of free feeding weight between 80% and 110%. The smooth curve is from Equation B6, with  $\delta = 0.30$  sec. The data are from four pigeons, reported by McSweeney (1974).

characteristic of the reinforcement schedule in question. These coefficients are derived in Appendix C. If we represent the coefficient as  $\zeta$ , and write  $k = \zeta/\delta$ , then Equation B4 finally becomes:

$$B = \frac{kR}{R + 1/a}, \quad k = \zeta/\delta; a, \delta > 0. \quad (\text{B7})$$

This is the *canonical equation* for the control of behavior by incentives. We need merely specify the coupling coefficient,  $\zeta$ , appropriate to the contingencies we impose on organisms to derive specific predictions of response rates.

Catania et al. (1988) assumed a time-based exponential decay of memory and the zeroing of memory by reinforcement to derive Equation B4. In contrast, the present treatment derives the hyperbola from activation and constraints on response rate; the zeroing of memory determines the coupling coefficient, which modulates the hyperbola but does not generate it.

Note that the coupling acts on the realizable response rate (Equation B4), not on the activation coefficient,  $a$ . It is behavior that is coupled to incentives, not potential behavior.

The blocked-counter model is not the only possible way to treat response ceilings. Instead we could assume that impulses to respond that occur while a response is in progress disrupt the ongoing behavior, causing the animal to abort the response. Such "stuttering" is unlikely in a ballistic movement such as a key-peck (or may be hidden in the  $\frac{1}{4}$ -sec band of "reflexive" responses), but might be apparent in extended sequences such as grooming or speech under high activation. Just such a hypothesis has been proposed by Robbins and associates (e.g., Robbins & Sahakian 1983) for the behavioral effects of large doses of psychomotor stimulant drugs. This behavior may be captured by a Type II, paralyzable counter model (e.g., Ross 1970), in which case the canonical equation becomes:

$$B = kaRe^{-aR}, \quad k = \zeta/\delta, \delta > 0. \quad (\text{B8})$$

### 13. APPENDIX C: The contingency models

Incentives fuel behavior (Equation B1), time constrains it (Equation B3), and contingencies direct it. They do so by making responses in memory at the time of reinforcement predictors of an incentive. The essence of contingencies is the correlation they effect between an incentive and an organism's memory of its behavior; organisms survive by confusing correlation with causality, acting as though repetition of the behavioral "causes" would bring about repetition of the reinforcing "effect." Evolution has smiled on this paralogism, populating the world with organisms that strive to refill memory with the apparent causes of the incentive, as best they can remember them.

Paradoxically, reinforcement itself may impair that memory. "A response-reinforcer pairing may be assumed to be the occasion for retroactive interference: Short-term memory for a response may be *reduced* by a reinforcer intervening between the response and subsequent recall of that response" (Shimp 1976c, p. 493). Killeen and Smith (1984) showed that animals' accuracy in identifying the response that caused reinforcement decreases to chance as the duration of the reinforcement increases to 4 sec. An incentive fills memory with the act of consuming it, which progressively displaces the memory of the responses that brought it about. This displacement follows the same time course as the decrease in marginal utility of an incentive, because the latter moments of extended incentives are more tightly coupled to the preceding consumption of the incentive than to the earlier instrumental responses. At short interreinforcement intervals, incentives may also be coupled to the consumption of the prior incentive and reinforce the premature responses that have been called "goal tracking." A thorough treatment would involve the convolution of responses with incentives over their mutual residence in memory. A simpler analysis follows.

### 13.1. Schedule-feedback functions

To use Equation B7 or B8 we must know the rate of reinforcement,  $R$ . On fixed-ratio (FR) schedules, reinforcement is delivered after the  $N$ th response. Rate of reinforcement is therefore proportional to the rate of responding ( $\dot{B}$ ) and inversely proportional to the ratio requirement ( $N$ ); the *schedule-feedback function for ratio schedules* is thus:

$$R = B/N. \quad (\text{C1})$$

On "constant-probability" variable-interval (VI) schedules with mean  $T$ , the probability of reinforcement after an IRT of length  $\tau$  is:

$$p(S^R|t=\tau) = 1 - e^{-\tau/T}.$$

The average rate of reinforcement – the schedule-feedback function for VI schedules – equals that probability divided by the average IRT,  $\tau$ :

$$R = (1 - e^{-\tau/T})/\tau. \quad (\text{C2})$$

See Baum (1992) for other candidate functions. To learn how much changes in rate of responding can change rate of reinforcement, we replace  $\tau$  with  $1/B$  and take the derivative with respect to  $B$ . For animals responding once per second on a VI 60-sec schedule, the slope is 0.033, showing that response rate has virtually no leverage on reinforcement rate and permitting us to treat the two as independent (see also Ettinger et al. 1987). L'Hospital's rule for limits shows that when  $\tau \ll T$  (as is usually the case) then  $R \approx 1/T$ . Therefore we write the approximate schedule-feedback function for interval schedules as:

$$R = 1/T, \quad (\text{C2}')$$

where  $T$  is the programmed minimum average time between reinforcers. Where this approximation fails (as  $\tau \rightarrow T$ ), Equation C2 may be used, or the prevailing rates may be substituted into the canonical equation. This does not undermine prediction, because no direct causal role is assigned to the schedule-feedback function. That function acts indirectly to determine activation level, which is the prime causal agent.

### 13.2. Ratio schedules

Substituting Equation C1 into the canonical equation B7 and rearranging gives:

$$B = \frac{\zeta}{\delta} - \frac{N}{a}, \quad \delta, a > 0; N \leq a/\delta, \quad (\text{C3})$$

and otherwise 0. This is the basis of Equations 9 and 10 in the text, where  $N$  is understood to be the average ratio requirement. Now we require the expression for zeta ( $\zeta$ ), the coupling coefficient.

**13.2.1. Fixed-ratio (FR) schedules.** What is the coupling coefficient for FR schedules? The first response contributes  $\beta$  toward memory of that class of responses, the second adds another  $\beta$ , with the impact of the first decaying to  $(1 - \beta)\beta$ , and so on. As the number of responses,  $j$ , increases to the number required for reinforcement,  $N$ , iteration increases the proportion of memory that is filled,  $M_N$ , to:

$$M_N = \sum_{j=1}^N \beta(1 - \beta)^{j-1} = 1 - (1 - \beta)^{N-1}. \quad (\text{C4})$$

This may be approximated by:

$$M_N \approx 1 - e^{-\lambda N}, \quad (\text{C5})$$

where  $\lambda = -\ln(1 - \beta)$ ; for small  $\beta$ ,  $\lambda \approx \beta$ . Only target responses move the animal toward reinforcement. If these constitute a proportion  $\rho$  of the total behavior, then  $\zeta = \rho M_N$ :

$$\zeta = \rho(1 - e^{-\lambda N}). \quad (\text{C6})$$

The theory presented in the main text of this article concerns primarily the changes in coupling as a function of  $N$ . In section 6 and in Appendix D, I discuss how the contingencies of reinforcement determine  $\rho$ . For ratio schedules we may set  $\rho$  equal to 1, although in predicting overall rates on FR schedules this will result in an overestimation of  $\delta$ .

We expand Equation C3 by inserting the coupling coefficient given by Equation C6 to achieve the complete model of FR performance. When  $N$  is small, so is  $\zeta$ ; as  $N$  increases, the coupling to behavior approaches its maximum ("saturates"), and response rate approaches the linearly decreasing asymptote  $1/\delta - N/a$ . We may extrapolate the asymptote back to the  $y$ -axis (where  $N = 0$ ) to estimate the ceiling on response rate,  $1/\delta$ , and we may extend it forward to the  $x$ -axis (where  $B = 0$ ) to estimate<sup>4</sup> the extinction ratio  $a/\delta$ . These predictions hold equally well for response rate after the pause ("running rate") and for overall rate, but in the latter estimates of  $\delta$  will be inflated.

**13.2.2. Variable-ratio (VR) schedules.** On VR schedules reinforcement occurs after a variable number of responses have been emitted. The random-ratio schedule is an "idealized" variable-ratio schedule in which there is a constant probability of reinforcement after each response. The probability of reinforcement after a single response,  $p$ , is  $1/N$ . The probability of reinforcement after two responses is:

$$p_2 = (1 - p)p, \quad (\text{C7})$$

and after  $n$  responses,

$$p_n = p(1 - p)^{n-1} \quad (\text{C8})$$

Assuming continuity, we may write this as:

$$p_n = p e^{-np}. \quad (\text{C9})$$

The probability that the process will stop with reinforcement when  $n = n'$  is:

$$p(\text{reinf}|n=n') = \int_0^{n'} pe^{-np}dn \quad (\text{C10})$$

$$p(\text{reinf}|n=n') = 1 - e^{-n'p} \quad (\text{C11})$$

The probability that the process will *not* have ended before this point is thus simply:

$$p(\overline{\text{reinf}}|n=n') = 1 - (1 - e^{-n'p}) = e^{-n'p}. \quad (\text{C12})$$

The responses up to  $n'$  are not equally weighted in memory. The most recent is given a weight of  $\beta$ , the penultimate response a weight of  $\beta(1 - \beta)$ , and in general  $w_n = \lambda e^{-\lambda n}$ , where, as before, we have replaced the parameter  $\beta$  with its continuous version,  $\lambda$ . The contribution of responses to memory is the integral over the decay process, weighted by the probability that the schedule will not have terminated by that point:

$$M = \int_0^{\infty} e^{-n'p} \lambda e^{-\lambda n} dn' \quad (\text{C13})$$

$$M = \lambda \int_0^{\infty} e^{-n'(p+\lambda)} dn' \quad (\text{C14})$$

$$M = \frac{\lambda}{p + \lambda}. \quad (\text{C15})$$

Remembering that  $p = 1/N$ , and that only the proportion  $\rho$  behavior are target responses, we may finally write the coupling coefficient for VR schedules as:

$$\zeta = \frac{\rho \lambda N}{1 + \lambda N}. \quad (\text{C16})$$

This makes sense. As the ratio requirement increases, the coupling improves, as the organism will be making longer sequences of responses before reinforcement truncates the process. The rate of approach to saturation is governed by the rate constant; where  $\lambda$  is large, saturation is quick. This hyperbolic equation is shallower (less concave) than the cumulative-exponential saturation on FR schedules; response rates increase more slowly and peak slightly later on VR than on equal-valued FR schedules. This is because it is the shortest component ratios that limit the amount of coupling; because they increase as a fraction of  $N$ , there will continue to be marginal improvements in coupling even for large  $N$ .

Inserting the ratio schedule-feedback function into the elicitation-constraint model (Equation B7) gave us Equation C3 for FR schedules. Equation C3 also predicts rates on VR schedules, but with  $\zeta$  now given by Equation C16.

### 13.3. Interval schedules

Substituting the approximate schedule-feedback function for interval schedules (Equation C2') into the canonical equation (B7), rearranging, and remembering that  $k = \zeta/\delta$ , gives the rate of responding predicted by the activation-constraint model. This is simply another form of the canonical equation when  $R = 1/T$ :

$$B = \frac{\zeta a}{\delta(T + a)}. \quad (\text{C17})$$

**13.3.1. Fixed-time (FT) schedules.** To calculate  $\zeta$  for interval schedules we start with a model of periodic incitement with no response requirement: an FT schedule in which the organism must merely wait  $T$  sec for an incentive to be delivered. To do this I reinterpret temporal requirements as response requirements. This gambit permits a coherent development of the theory for various types of schedules and is a critical assumption about the events (responses, not unfilled time) that index memory. As was the case for the FR model, we integrate the contents of memory from one reinforcement to the next, when the number of responses emitted will equal the response rate ( $B$ ) times time ( $T$ ):  $N = BT$ , which forms the upper limit of the integral:

$$M_N = \lambda \int_0^N e^{-\lambda n} dn. \quad (\text{C18})$$

The exponent is the memory-decay rate ( $\lambda$ ) multiplied by the total number of responses,  $n$ . Equation C18 represents the saturation of memory by *all* responses elicited in the context. The target response may be a major component of that ensemble, or a minor one; coupling will equal Equation C18 times the coupling constant  $\rho$ . Solving Equation C18 gives the *coupling coefficient for FT schedules*:

$$\zeta = \rho(1 - e^{-\lambda BT}), \quad (\text{C19})$$

where the factor  $\rho$  represents the proportion of the coupling specific to the target response. Compare this with Equation C6. The critical difference in these schedules is that the ratio contingencies attract  $\rho$  toward 1, whereas noncontingent incitement does not.

It is useful to approximate Equation C19 by expanding the exponential in a power series, and truncating it after the first two terms. This yields:

$$\zeta \approx \frac{\rho \lambda BT}{1 + \lambda BT}. \quad (\text{C20})$$

Equation C20 provides an important general model for the *adventitious coupling* of responses to an incentive.

Note that response rate,  $B$ , enters the expression for the coupling coefficient in such a way as to generate positive feedback: if a behavior occurs at a high rate, it will more fully saturate memory and increase coupling; if rate is lowered, coupling weakens and rate lowers still more. Thus we predict serious instability in schedules of noncontingent reinforcement.

**13.3.2. Fixed-interval (FI) schedules.** Under FI schedules reinforcement is delivered immediately after a target response, so we are guaranteed at least the strengthening available to a single, contiguous response,  $\beta = w_1$ ; that is, equivalent to the coupling coefficient under FR 1 schedules:

$$w_1 = \beta = 1 - e^{-\lambda}. \quad (\text{C21})$$

The remainder of the coupling is due to the memory for behavior preceding the contingent response; that is, Equation C19 with the lower limit of integration raised from zero to the penultimate response. Adding these partitions of the animal's history gives us the *coupling coefficient for FI schedules*:

$$\zeta = \beta + \rho(1 - \beta - e^{-\lambda BT}). \quad (\text{C22})$$

**13.3.3. Rho.** In an earlier draft of the theory I assumed that the total amount of behavior is constant, with the measured behavior constituting a progressively larger share as its rate increases, so that  $\rho = \delta/\tau = \delta B$ . This was also Herrnstein's (1979) assumption in his derivation of Equation 4. But this assumption provided an inferior account of the data. Furthermore, it is inconsistent with our fundamental assumption that rate of emission of behavior is proportional to the rate of incitement (Equation B1). The present approach takes the initial value of  $\rho$  for some particular response  $j$ ,  $\rho_j$ , as equal to the extent to which it is activated by the context and incentive,  $a_j$ , relative to all the other available responses:

$$\rho_j = \frac{a_j}{\sum_i a_i}. \quad (\text{C23})$$

Responses that are readily elicited in the context have large initial values for the coupling constant  $\rho$ , and thus depend less on contiguity with reinforcement for their instigation, whereas for "unprepared" responses  $\rho$  will be small and most of the coupling will be due to contiguity with the incentive ( $\beta$ ; see Equation C22).

Ratio schedules exert strong selection pressure on the target response, driving its  $\rho$  close to 1. Interval schedules select the target response with the weight  $\beta$ , and all other behaviors elicited as a function of their availability in memory (Equations C20 and C23). The value of  $\rho$  tells us what proportion of this mixture strengthens the target response. Whenever experiments "tune" the experimental environment, they are adjusting the value of  $\rho$ , hopefully increasing it to maximize the coupling of the incentive to the target response.

Because all responses (including unmeasured responses) index memory, and do so as a function of their duration, memory will appear to saturate more quickly under interval schedules. We take these other responses into account by writing  $\lambda = \lambda'/\delta/\rho$ . In Tables 2 and 3 I report values for  $\lambda'$ , fixing  $\rho$  at 1, while recognizing that  $\rho$  actually takes lower values when predicting overall rates on FR schedules, and still lower values for interval schedules. Greater understanding of  $\rho$  requires analysis of the dynamic changes in behavior during conditioning (Appendix D).

Equation C22 is relatively intractable: upon insertion in the canonical equation, the system must be evaluated numerically, allowing successive iterations for the predictions to settle down. Alternatively, we may exploit Equation C20 as an approximate coupling coefficient for interval schedules, with the schedules having a response requirement (e.g., FI) taking larger values of  $\rho$  than those without (e.g., FT). These approximations make

possible a remarkably general formula for predicting responses rates on interval schedules, as will be seen.

**13.3.4. Variable-time (VT) schedules.** How do we calculate  $\zeta$  for VT schedules? A "constant probability" variable-interval schedule operates somewhat like a random-ratio schedule, in that the reinforcers are "set up" according to a similar Poisson process, but one driven continuously in time rather than continually by responses. We may therefore use the same logic as that employed in Equations C7 through C12, with  $t$  substituted for  $n$ , to infer:

$$p(\overline{\text{reinf}}|t=t') = 1 - (1 - e^{-t'/T}) = e^{-t'/T}, \quad (\text{C24})$$

the probability of the schedule not "stopping" with delivery of reinforcement before time  $t'$  is a simple exponential function of that time, with a time constant,  $T$ , the average time between reinforcements.

Following the above logic, we must integrate the probability of the processes not having ended, weighted by the contents of memory, over all time. But our variable of integration in Equation C13 was responses, and Equation C24 gives us a probability in terms of time. The solution is to discretize time in the development of Equation C24, letting  $t' = n'\tau$ , where  $\tau$  is a small unit of time such as that corresponding to an IRT. Then we may write the probability of reinforcement coupling to some behavior as:

$$M = \lambda \int_0^\infty e^{-n'\tau/T} e^{-\lambda n'} dn', \quad (\text{C25})$$

where the first exponential term represents the reinforcement distribution, and the second the weighting of it in memory. Solving Equation C25 tells us the extent to which an interval schedule of mean duration  $T$  can fill memory:

$$M = \frac{\lambda BT}{1 + \lambda BT}. \quad (\text{C26})$$

However, only the targeted instrumental responses will contribute to the coupling between measured behavior and reinforcement. Nontarget interresponses (all of those that fill the IRT such as off-key-pecking, preening, etc.) will occupy memory and will satisfy the temporal contingencies, but not contribute to the coupling of the target response. Multiplying Equation C26 by  $\rho$  gives us the *coupling coefficient for VT schedules*:

$$\zeta = \frac{\rho \lambda BT}{1 + \lambda BT}. \quad (\text{C27})$$

Note that this is the same as Equation C20, but whereas the former was an approximate solution for FT schedules, this is an exact solution for VT schedules. It is similar to the coefficient for VR schedules (Equation C16), with coupling increasing hyperbolically with  $N$  in the former case, and with  $BT$  in the latter.

There is strong positive feedback from behavior under these contingencies, as reflected by the presence of response rate,  $B$ , in the equation. Faster responding leads to better coupling, which leads to faster responding, and so on. This autocatalytic property of the coupling coefficient captures the "superstitious" nature of responding under such schedules, where the adventitious emission of a response permits it to be captured by the reinforcement contingencies. If  $B$  fluctuates low, coupling weakens, and behavior may extinguish. If  $B$  fluctuates high, coupling improves, and response rate will increase toward its ceiling. It is the guaranteed contiguity with reinforcement on regular VI schedules,  $\beta$ , that buffers behavior against such extreme autocatalysis.

**13.3.5. Variable-interval (VI) schedules.** Under VI schedules reinforcement is delivered immediately after a target response, hence we are guaranteed a minimal coupling of  $\beta$ . The remainder of the coupling is due to the memory for behavior preceding the contingent response, and so we add to  $\beta$  a version of Equation C18:

$$\zeta = \beta + \rho \lambda \int_1^\infty e^{-n' \tau / T} e^{-\lambda n'} dn'. \quad (\text{C28})$$

Solving Equation C28 leads to the coupling coefficient for VI schedules:

$$\zeta = \beta + (1 - \beta) \left( \frac{\rho \lambda B T}{1 + \lambda B T} \right). \quad (\text{C29})$$

Equation C29 makes sense. As response rate ( $B$ ) increases, reinforcement contacts more of the targeted responses, and  $\zeta$  increases. This leads to a further increase in response rate, as seen in the initial acquisition of behavior under these schedules. Eventually, ceilings on response rate cause a leveling off of measured behavior.

Despite the stabilizing effect of  $\beta$ , a significant proportion of the impact of reinforcement ( $1 - \beta$ ) is available to strengthen responses indifferently. This is a serious source of instability and leads to cyclicity in response rates on interval schedules, which has been observed and called "metastability."

For routine analysis of interval data, Equation C29 can be further simplified by setting  $\beta$  to zero, yielding Equation C20/27 (and a slightly inflated value for  $\rho$ ). Inserting into C17 and writing  $T$  as  $1/R$  gives:

$$B = \frac{kR}{(R + 1/a)} - \frac{R}{\lambda}, \quad (\text{C30})$$

where

$$k = \frac{\rho}{\delta},$$

and

$$\lambda, a, \delta > 0, \quad R < \lambda \rho / \delta - 1/a.$$

This is Equation 12 in the text. It is the basic hyperbolic model minus a linear function of  $R$  that undermines coupling because of the truncation of memory by reinforcement. (This appendix is like Conway's computer game "Life," in which an original figure is lost after a couple of iterations, only to reemerge later in multiple incarnations. The original assumption of limits to memory has been transmogrified by the mathematical derivations, only to reemerge in Equation C30 in a particularly simple and explicit paraphrase of the originating assumption.)

We have the option of either letting  $\rho$  remain a free parameter, or setting it to 1 and absorbing de facto differences in  $\rho$  into  $\delta$ . There are only three situations in which a precise value can be assigned to  $\rho$ : where explicit empirical estimates of  $\delta$  are obtained; where the exact model (Equation C29) is utilized; and in schedules that combine different types of contingencies (sect. 13.4). Equation C30 generates predictions very close to the exact solution (Equation C29 inserted into Equation C17 and iterated), and has the advantage of mapping experimental data within the limits of empirical reliability without requiring iteration and without invoking  $\rho$  as a free parameter.

Equation C27 tells us that coupling will have fallen to half its maximum value ( $\rho$ ) when  $R = \lambda B$ . Consider a situation in which  $\lambda = 0.1$  and  $B = 2$ . Coupling will fall to half its maximal value when  $R = 0.2$ ; that is, for a VI 5-sec schedule. This is consistent with the data reported by Preston and Fantino (1991), in which pigeons' preference for a VI schedule *decreased* when the average delay to it decreased below VI 10 sec. Reinforcement rate was sufficiently close to its ceiling that further decreases in

the VI schedule hurt the coupling more than it helped the activation, thus yielding this otherwise surprising result (one predicted for other reasons by Fantino & Abarca 1985, p. 357). We see that the present functional approach to the definition of the response has pervasive ramifications, not all of which are immediately obvious (Palmer & Donahoe 1992).

### 13.4. Sequential coupling

We may generalize the logic to develop models for any schedule. Consider, for instance, the continuum of contingencies from intervallike to ratiolike that is generated when the terminal response requirement on an FI schedule is increased from one response to  $N$  responses. Coupling under these "tandem" schedules is calculated by adding the coefficient for ratio schedules of length  $N$  (Equation C6) to that for interval schedules starting from  $N$  (Equations C18 and C19 with the lower bound of integration increased to  $N$ ):

$$\begin{aligned} \zeta &= \zeta_{\text{FR}=1}^N + \zeta_{\text{FI}=N}^{BT}; \\ \zeta &= (1 - e^{-\lambda N}) + \rho(e^{-\lambda N} - e^{-\lambda B T}). \end{aligned} \quad (\text{C31})$$

A similar compound model for FR-VI was presented as Equations 16 and 17, and plotted in Figure 15.

The models in this article concern response rate as the dependent variable, not the interresponse time (IRT). Important progress in disentangling the effects of activation-elicitation from contingency on IRTs has been achieved by Platt (1979; Galbicka & Platt 1986), following the pioneering work of Anger (1956) and Shimp (1973; for a recent review, see Galbicka 1988). In the present theory the treatment of IRTs is straightforward. For an organism to emit an IRT of  $\tau$  sec requires that it *not* emit a response for  $\tau$  sec, and then respond.<sup>5</sup> Target responses during the interval do not enhance the coupling; they undermine it by eliminating the opportunity for the completion of that IRT. The coupling of an incentive to an IRT therefore consists in the coupling to the response that terminates it,  $\zeta_R = \beta$ , in concert with the noncoupling to that response during the previous  $\tau$  sec:  $1 - \zeta_\tau = 1 - [(1 - \beta) - e^{-\lambda \tau}]$ . Because both are required, the expected coupling is their product:

$$\zeta_{\text{IRT}=\tau} = \beta(\beta + e^{-\lambda \tau}), \quad \tau \geq \delta. \quad (\text{C32})$$

For very short IRTs ( $\tau \approx \delta$ ) the parenthetical factor goes to 1 and the coupling becomes  $\beta = w_1$ , the coupling to the response that terminates the IRT. As  $\tau$  increases, coupling falls off exponentially with  $\tau$  to a floor of  $\beta^2$ . As before,  $\zeta$  must be embedded into the canonical equation with an appropriate schedule-feedback function and point-spread function to represent the scalar error in timing. This model will permit analyses complementary to those of section 5, and may also be applied to contingencies that require inhibition of a particular pattern of responses (e.g., Neuringer 1992).

### 13.5. Caveats

In applying these formulas attention must be paid to the units of measurement. It is convenient to measure all events varying with time in units of seconds. I have done this in Tables 2 and 3, but the figures retain the units in which the data were reported.

The exact models for interval schedules require iteration of the canonical equation and coupling coefficient to derive successive approximations of the value for  $B$ . The system converges most quickly if the first estimates for  $B$  are derived from the approximate models.

## 14. APPENDIX D: Behavior as a trajectory

To understand how behavior is drawn from its initial desultory condition to the coherent patterns described by the asymptotic models, we must recognize that behavior constitutes a

trajectory through a field of attractors. Figure 13 sliced through this space, plotting successive instances of the target response on the  $y$ -axis as a function of all other responses on the  $x$ -axis. Upon each tick of an internal clock the trajectory moved ahead, with reinforcement occurring when the trajectory crossed a consequential line. In this appendix I (1) sketch the forces associated with basic schedules; (2) show why they attract behavior; (3) describe the metric of the resulting space; (4) show that the projection of them onto the dimensions measured by the experimenter yields the asymptotic coupling coefficients; and (5) simulate responding as I did in section 6 to show how the coupling constant  $\rho$  takes different values under different schedules.

#### 14.1. Consequential lines and their attractive force

**14.1.1. Fixed-ratio (FR) schedules.** The consequential line of an FR  $N$  schedule extends horizontally from an ordinate of  $N$ . Why should it "attract" behavior? Assume a uniform probability of making the target response  $p_i = \rho$ , and consider the response that occurs after the fifteenth such response on an FR 16 schedule. If it also a target response, it is strengthened by the amount  $w_1 = \beta$ , so that its expected coupling is  $\zeta_{\text{Target}} = \rho\beta$ ; if it is some other response followed by a target response, which occurs with probability  $(1 - \rho)\rho$ , it is strengthened by the amount  $w_2 = \beta(1 - \beta)$ ; if it is two other responses followed by a target response, which occurs with the probability  $(1 - \rho)^2\rho$ , it is strengthened by the amount  $w_3 = \beta(1 - \beta)^2$ . The sum of this infinite series gives an expected

$$\zeta_{\text{Other}} = \frac{\rho\beta(1 - \rho)(1 - \beta)}{1 - (1 - \rho)(1 - \beta)}.$$

Note that if  $\rho$  is close to 1, the target response receives the preponderance of coupling. If  $\rho$  is close to 0, a target response is unlikely to occur; but if it does, it will receive a strength of  $\beta$ . This will feed back into the calculations by increasing the value of  $\rho$ , carrying behavior over the consequential line, and thereby increasing the likelihood of a target response on the next trial. We may calculate the advantage to the target response as  $\zeta_{\text{Target}}/\zeta_{\text{Other}}$ . The derivative of this ratio with respect to  $\rho$  is  $1/[(1 - \rho)^2(1 - \beta)]$ : Whenever  $\rho$  gets larger the advantage gets larger, and the gains are a strongly increasing function of  $\rho$ .

The target response may be unlikely and hard to get going, but once it "catches," it is self-sustaining. A trainer may expedite the process by shaping the organism to make the first response, but as long as  $\rho$  is nonzero, the process will eventually catch by itself. This attraction of behavior toward the target response propagates each step back in the sequence, but with diminishing effectiveness. Too large a requirement of low-probability responses affords many junctures at which the trajectory is strained to the breaking point. As in all of the laws expressed in this article, response probabilities or rates appear on both sides of the equations, causing the metastability that is an inherent part of adaptive behavior.

The attraction exerted by an incentive on behavior is not that of a magnet to a metal. Rather, it is the increased likelihood, because of the conditioning process, of the trajectory rotating into a more vertical position after each reinforcement. Figure D1 shows a schematic view of a small part of the field after some conditioning. The vectors are the gradients, the resultant force at each step of the trajectory. The vertical component of each vector is  $\zeta_{\text{Target}}$ , and the horizontal component  $\zeta_{\text{Other}}$ .

**14.1.2. Variable-ratio (VR) schedules.** Consider a VR schedule with two components,  $N = 8$  and  $N = 24$ . These entail horizontal consequential lines at those ordinates, crossing either of which is reinforced only half the time. Although this reduces the attraction to the consequential lines, it spreads it out more effectively through the behavior space. For example, the attrac-

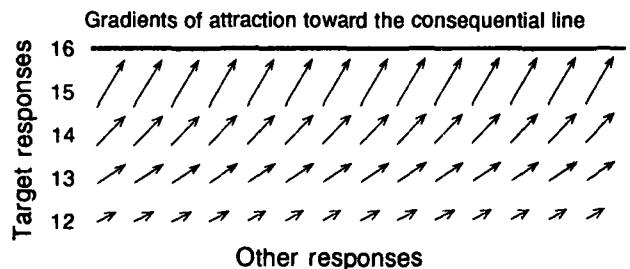


Figure D1. Gradients of attraction to the consequential line on an FR schedule. The length of the vectors gives the weight in memory available to a response. Their direction gives the expected coupling to a target response ( $y$  component) relative to other responses ( $x$  component). This is a dynamic field in the sense that as it drives behavior in the direction of the gradients, the gradients themselves change direction.

tion of the shortest component of the VR on the first element of the trajectory,  $0.5 \times e^{-\lambda 8}$ , is greater than for a comparable FR 16,  $1 \times e^{-\lambda 16}$ . Carrying this logic to its limit, we see that the most effective scheduling reinforces each response with probability  $1/N$ , which is the random ratio version of a VR schedule. Its consequential lines are faint horizontal rules covering the space, at distances corresponding to the durations of the response units. Figure D2 shows the effects of such contingencies on the probability of a target response for various values of  $N$ . The same conditioning model described in section 6 was used for this and all other simulations. The logic of the dynamic field is the same as that for FR schedules, driving the trajectory to an increasingly vertical orientation as  $N$  increases. The curve is from the asymptotic model (line 3 of Table 5), with a recovered value for  $\rho$  of 0.97.

If we look at the contingencies under these schedules backwards, it is possible to come to the wrong conclusions. We might think that since each target response is independently reinforced with probability  $1/N$ , it does not matter what the animal had done before the terminal response, and therefore there should be no attraction to large values of  $\rho$ . But it is a Bayesian fallacy to assume the final reinforced response as a given and look backward; that is because not all paths to the terminal response have equal expected coupling. What controls behavior

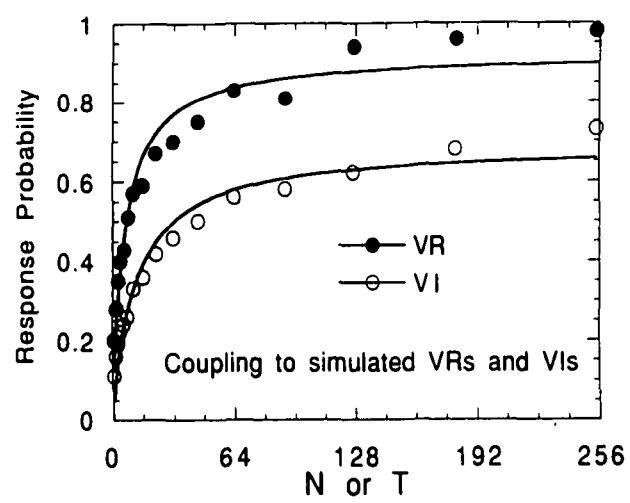


Figure D2. The probability of making a target response on VR and VI schedules as a function of their length, based on the stimulations. The curve through the ratio data is given by line 3 of Table 5 and that through the interval data by line 8 of Table 5, with  $B$  fixed at 1. The asymptote ( $\rho$ ) of the former is 0.97 and that of the latter 0.69.

at each choice point along the trajectory is the probability of reinforcement weighted by its delay. The logic of the previous section applies: in looking forward to the next response, the expected coupling to a target response is  $\rho\beta/N$ , whereas the expected coupling to a nontarget response is at most  $(1 - \beta)$  times that value.

**14.1.3. Fixed-time (FT) schedules and the metric of behavior space.** The consequential lines for an FT schedule are shown in the top panel of Figure D3. As soon as  $T$  seconds has elapsed the incentive will be delivered, and any trajectory that carries the animal  $BT$  responses out from the origin satisfies that criterion. The average length of the trajectory is  $BT$  responses, where  $B$  is the response rate. Whatever the organism was doing at the time of reinforcement is attracted to the consequential line on the next trial, giving rise to the fixation of "superstitious" behavior.

Responses are often exclusive: at any one time an animal may be either grooming or leverpressing, but not both. This means

that it is generally impossible to travel through behavior space on a diagonal. The trajectory moves in quantal steps up or to the right. Because of this, its metric is not Euclidean but rather "city-block." Thus, the distance from the origin to the consequential line is the same whatever sequence of target and other responses is emitted, and all are equally attracted.

For "integral" responses that are continuous blends, however, the metric is Euclidean, and the shortest trajectory from the origin forms a right angle with the consequential line. This is important under ratio schedules, where the continuous components of a response are shaped toward their most efficient form, causing the value of  $\delta$  to approach its true minimum (Williams & Johnston 1992).

**14.1.4. Fixed-interval (FI) schedules.** The consequential lines of these concatenations of FT and FR 1 schedules are drawn in the bottom panel of Figure D3. The satisfaction of the temporal portion may leave the organism anywhere along the diagonal boundary of the interior, requiring one additional target response after each of the possible prior trajectories.

Remember that the consequential lines intersect the axes at  $BT$  responses. Because there is variability in response rate, different parts of the trajectory will cross the line from one trial to the next, spreading the effect of reinforcement over a length of a trajectory. Simulation of these contingencies using the standard program yields saturation curves shown in Figure D4. The data are fit with Equation 11 with  $B$  fixed at 1. Asymptotic coupling ( $\rho$ ) is 0.43 for the trajectory as a whole, 0.56 for the last half of it.

Because there is variability in the rate of emission of responses, the portion of the trajectory that intersects a consequential line is a Gaussian function of its length. Consistent with this, response rates in these simulations are found to be a bell-shaped function of time. Because the rate of responding is proportional to the rate of reinforcement, the standard deviation of this Gaussian function is linear with  $T$ . The model thus provides a mechanism for scalar timing (Gibbon 1977).

**14.1.5. Variable-interval (VI) schedules.** The consequential lines for variable-time schedules are nested diagonals cutting the space along each unit of the dimensions, the crossing of which is probabilistically reinforced. Compound this with the FR 1 requirements and we see a picture similar to the bottom panel of Figure D3, but replicated over the whole space. These contingencies give rise to the couplings shown in Figure D2, and fit by

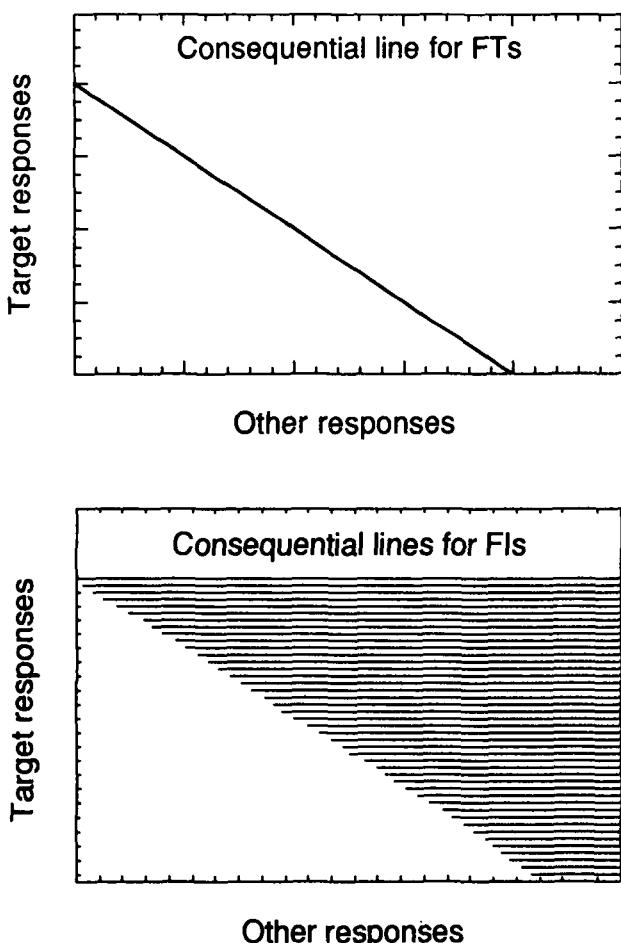


Figure D3. *Top:* The consequential line for periodic noncontingent reinforcement (an FT schedule). Any combination of responses that cross the line will be strengthened. Because the emission of behavior is probabilistic, different portions of the trajectory are cut from one trial to the next, so the "line" actually represents the locus of the center of a Gaussian density projected through the response space. *Bottom:* The consequential lines for FI schedules. Upon completion of the temporal requirements, one additional step is needed for reinforcement. Because of the variance in length of the trajectory, this strengthening of the target response will be distributed over elements of the trajectory as a Gaussian function of their position. The dependency of response rate on reinforcement rate gives this distribution a standard deviation proportional to the time between reinforcers.

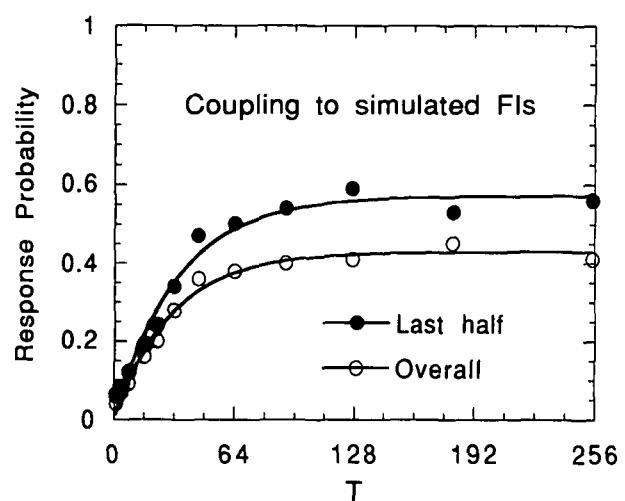


Figure D4. The probability of making a target response on FI schedules as a function of their length, based on the simulations, both for the trajectory as a whole and for its last half. The curves through the data are directed by line 5 of Table 5, with  $B$  fixed at 1.

the asymptotic model (Equation 11, with  $B = 1$ ), with a recovered value of  $\rho = 0.71$ .

#### 14.2. Rho

The coupling constant,  $\rho$ , is in the curious position of playing a central role in the theory, yet being frozen out of most of the asymptotic models, where it is set to 1 and absorbed into the value of  $\delta$ . Rho is the proportion of target responses in the trajectory; the length of the ordinate divided by the length of the trajectory. In Euclidean space, this is the cosine of the angle  $\theta$  between the ordinate and the trajectory, which measures the correlation between the directions of those vectors; in city-block metric it is the ordinate divided by the ordinate plus abscissa. Since the units of the axes are  $\delta$  seconds, the projection of the ordinate onto the trajectory is  $\rho N/\delta$ . The coupling coefficient weights and sums these  $N$  elements, with memory decay occurring at the constant rate of  $\lambda'$  along the trajectory being hastened by the factor  $\delta/\rho$  when measured along the projection onto the target dimension. When we count all target responses alike, then  $\zeta$  is the sum of the elementary coefficients. It is an increasing function of both the (average) length of the ordinate, which is determined by the response requirement ( $N$ ) scaled to the duration of the responses ( $\delta$ ), and  $\rho$ , the projection of the target dimension onto behavior's trajectory.

In calculating the coupling coefficient we said that reinforcement zeros memory. But this is merely a convenient way of dissecting the trajectory into more homogenous parts. During reinforcement the organism is occupied with the "other" consummatory response. The projection of this activity on the target dimension is zero. Starting our summations after reinforcement is equivalent to starting them in the remote past, but assigning zero projection of the target responses onto the trajectory until the completion of the most recent reinforcement. The zeroing of memory by reinforcement is thus seen as a computational approximation rather than an ad hoc theoretical commitment. More accuracy in predicting the effects of frequent brief incentives will be gained by integrating over the incentives and what preceded them (and over subsequent events as well!).

Depiction of reinforcement by a consequential line is also a computational approximation. Extended incentives strengthen contingent responses even while they are displacing them from memory. Thus, we should speak of consequential *regions* in space that attract first instrumental responses, and then consummatory responses. It is not surprising that instrumental responses often share some properties of the consummatory responses toward which they are directed, and that coupling is greatest in experimental arrangements that engineer a high correlation between instrumental and consummatory response elements.

#### ACKNOWLEDGMENTS

This project was supported in part by grants BRSG 2 SO7 RR07112 from NIH, 1 R01 MH48359 from NIMH, and BNS 9021562 from NSF. I thank Heather Cate, who programmed Experiment 2, Ron Epperlein, who checked the math, and Anthony Dickinson, who kept me from assuming that which I should demonstrate.

#### NOTES

1. I seeded  $M_0$  with the obtained proportion of  $B$  key responses; however, the correlation remains perfect even if  $M_0$  is seeded with an arbitrary value such as 0.5.

2. Parameters are considered unchanged when liberating them from a common value improves the percentage of variance accounted for by less than one point.

3. Alternatively, it may take a minimal amount of activation ( $k$ ) to initiate a response, which would entail a relation such as  $\tilde{B} = aR/(k + \delta)$ ; energetic considerations may come into play, and so on.

4. To estimate the three key parameters, first calculate  $\delta$  as the reciprocal of the estimated  $y$ -intercept. Then calculate  $a$  by multiplying the  $x$ -intercept by  $\delta$ . Next select a small value of  $N^*$  that gives rise to responding,  $B^*$ , at about one-half the maximal rate ( $1/\delta$ ). Then  $\lambda = -\ln[1 - \delta(B^* + N^*/a)]/N^*$ . This technique requires data at adequately large values of  $N$  to provide good extrapolations for the first two parameters. The peak of the function occurs at  $N^* = \ln(\lambda a/\delta)/\lambda$ . A similar technique works for VR schedules, although the memorial effects take longer to saturate, thus making the extrapolation less accurate. Throughout, it is assumed that  $\rho = 1$ .

5. Although called an *interresponse* time, most researchers have measured it from the beginning of one response to the beginning of the next, as though  $\delta = 0$ . In this paper I retain the convention of measuring IRTs as cycle times from one response to the next, but remember that that time includes the response time  $\delta$ , which is the minimum IRT.

## Open Peer Commentary

*Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.*

### The response problem

Robert C. Bolles

*Department of Psychology, University of Washington, Seattle, WA 98195*

When learning theorists finally abandoned the stimulus-response (S-R) association as the basic unit of analysis, they could go on to enjoy a variety of new explanatory devices. They also encountered a new problem, however; how to explain the occurrence of the learned response. The old S-R association provided a more or less direct account of the response, because the response was elicited by the stimulus. The Hullian might insert some motivational terms and some inhibition, but the response was inevitably selected from the animal's repertory by the S-R association. When they gave up S-R associations, learning theorists needed some new device to account for the occurrence of the response in a learning situation.

A stream of responses is highly correlated with the stream of reinforcers it is programmed to produce and perhaps that provides a handle on the problem. One possibility is that a reinforcer can reach back in time to increase the probability of the last response, or the last few responses, that just occurred. It is not clear how such backward action might work, however, nor is it clear how the reinforcer could establish stimuli in the environment as discriminative stimuli. How could that be done nonassociatively? I do not think we want reinforcers "reaching back." Another possibility is that through learning the animal comes to expect that a particular response will produce a reinforcer, or incentive, or  $S^*$ . I find this to be a relatively attractive concept, but it does not solve the response problem because it is not clear why the expectation of some desired outcome, such as food, should generate a response that has been correlated with it. The response problem is still there: Where does the response come from? How does it arise?

Killeen offers a new possibility. Perhaps, he seems to suggest, it is not our problem but the animal's. Put the response, or some representation of it, into the animal's short-term memory and let

the animal worry about how to get it transposed into overt behavior. And how do we deal with the correlation of response and reinforcer? Put the reinforcer, or incentive, or  $S^+$ , or some representation of it, into short-term memory also. The reinforcer does not have to reach back to the response, nor does the animal have to expect the response to produce it. The critical terms merely have to reside in short-term memory at the same time. So "coupled," the representation of the incentive leads to the overt response. This is good stuff; I like it.

One reason I like this target article is that it makes the animal responsible for the determination of its behavior. Ever since the early days of Thorndike a major thrust of behaviorism has been to get all the determinants out in the environment. The  $S$  was out there (it is perhaps a light coming on); the  $R$  was out there (it is perhaps a switch closure); and so too the incentive motivator and reinforcer comes out of the environment. There was nothing inside except a bundle of  $S$ - $R$  associations. I want to see all of these things given back to the animal. Put them back inside there where we can do some psychology with them. Put them where they belong.

A second reason I like this model is that Killeen has not been intimidated by the problem of objectivity. Switch closures are objective because we can observe them; short-term memory cannot be observed directly, so there is a problem. Undaunted, Killeen arms himself with a simple two-parameter memory-decay function and proceeds to measure this unobservable thing! There seems to be no problem of objectivity.

A third reason is Killeen's insistence that we seek the animal's definition of the response, how it is coded or represented in memory, rather than relying on our switch closures to define the response. After a rat has been leaning on a lever hundreds of times a day for a few months, it probably has a response definition very much like the experimenter's. Earlier in the course of learning, however, there is likely to be considerable slippage. We think the rat is turning left in the maze, but it thinks it is going West. My students once devised a lever that could be lifted up as well as pushed down, so we were all set to do some response discrimination studies. What we found to our dismay was that animals could not work with this great device because they insisted on defining the different responses in terms of body parts. They would press with their paws and lift with their noses, or vice versa (Bolles 1989). So we were looking at response discrimination but our rats had turned it into a stimulus discrimination problem.

Rats undoubtedly code responses in terms of the stimulus feedback that the response produces. We once had rats working on Sidman avoidance that were pressing a telegraph key type of lever that provided minimal feedback. They could not learn it. Another group had the same silent lever, but they had their cumulative recorders placed on top of their chambers to provide a dull thump for each response. They learned it. So rats seem to require some form of response feedback to define a response, but the programmed incentive does not suffice. A rat can press one lever to gain food pellets and a second lever to get saccharin. It learns the response discrimination with little trouble (when the alternatives are brought under stimulus control). But the discrimination is greatly facilitated by presenting a brief stimulus after one of the responses. That brief stimulus should be totally overshadowed by the incentive that followed every response, but no, the brief stimulus is clearly important in establishing the discrimination. My guess at this point is that responses are defined for the animal by immediate stimulus consequences *other than* the incentive it produces. Response-defining feedback stimuli may have to be hedonically neutral.

## **Validation of behavioural equations: Can neurobiology help?**

**C. M. Bradshaw**

*Department of Psychiatry, University of Manchester, Manchester M13 9PT, United Kingdom*

Killeen's target article is a remarkable achievement. By drawing together aspects of operant behaviour that have usually been treated separately by psychologists (memory, contingency, arousal) and constraining them within a parsimonious mathematical model, Killeen has taken an important step forward towards the elusive goal of a coherent theory of reinforcement schedules. In this commentary I would like, initially, to cavil at some minor discrepancies (as I see them) in the empirical foundations of Equation 12 and then to raise a more general point about the validation of this and other quantitative models of behaviour.

The hyperbolic form of the relation between reinforcement rate and response rate under variable-interval (VI) schedules is well established (see Bradshaw & Szabadi 1988; 1989; de Villiers 1977; Herrnstein 1970; Heyman & Monaghan 1987). Killeen's model predicts this relation and in addition accommodates a departure from the hyperbolic form that is occasionally evident in experimental data, namely, a downturn of response rate at very high reinforcement frequencies. The cost of accommodating this "paradoxical" effect of high reinforcement frequencies is, of course, the inclusion of another free parameter,  $\lambda$ , into the equation. Although  $\lambda$  is a crucial theoretical entity in the model, one might question whether its presence in Equation 12 is fully justified by extant data. The downturn of response rate under VI schedules, when it occurs, is generally slight (e.g., Snyderman 1983). Moreover, technical factors, such as within-session changes in deprivation level and/or brief postreinforcement pauses (PRPs) occasioned by unrecorded postprandial behaviour (which presumably contaminate response rate to an increasing degree at high reinforcement rates) may provide plausible alternative explanations for the effect. Some circumstantial evidence for the contaminating effect of PRPs on response rates in interval schedules is provided by Chung's unpublished data (described by Herrnstein 1970) from tandem fixed-ratio 1 fixed-interval  $t$  schedules, from which PRPs had been eliminated; no downturn of response rate is apparent, even with reinforcement rates of over 2,000 reinforcers per hour.

A second quibble concerns the sensitivity of  $k$  and  $a$  to different types of experimental manipulation. As Killeen rightly points out, we know that  $k$  is sensitive to changes in response requirement and  $a$  to changes in "motivational" variables (reinforcement quantity or quality, deprivation level, etc.). In Killeen's model, these effects are deemed to be specific, in that  $k$  should not be affected by motivational manipulations nor  $a$  by changes in response requirement (this is also posited by Herrnstein's [1970; 1974] account). There seems to be general agreement that a change in  $a$  cannot be achieved by manipulating response topography or effort requirement (see Bradshaw & Szabadi 1988; 1989; Heyman & Monaghan 1987), and a change in this parameter may therefore be used as a "marker" for the motivational impact of interventions such as brain lesions or drug treatments (e.g., Hamilton et al. 1985; Heyman 1983; Morley et al. 1988; Wogar et al. 1991). However, there is a small but annoyingly persistent set of observations that indicate that  $k$  may be sensitive to some motivational manipulations (e.g., Bradshaw et al. 1978; Shah et al. 1991; see also de Villiers 1977 for earlier references). Thus, if an intervention alters  $k$  without affecting  $a$ , one may have confidence that it has influenced the mechanisms of response production (e.g., Heyman et al. 1987); however, if  $a$  is also altered, one cannot exclude the possibility that the change in  $k$  reflects a motivational effect of the intervention. Killeen has convincingly argued that the asymptote of the response-strength function should be seen as a conglomeration

of several parameters that reflect the operation of diverse processes; perhaps there is a motivational parameter hidden in there somewhere that will be revealed by future theoretical analysis.

At a more general level, Killeen's exposition is not only bristling with testable predictions for behavioural experiments; it also points the way towards novel *post hoc* interpretations of neurobehavioural experiments. I would like to go a step further and suggest that the relationship between the model and neurobehavioural data is a reciprocal one, in that such data may provide a vital means of validating Killeen's model and indeed any other mathematically sophisticated model of behaviour. As the modelling of behavioural phenomena becomes increasingly elaborate, there is an inevitable tendency for free parameters to proliferate, rendering the equations ever more versatile. But behavioural data being as they are, it becomes increasingly unrealistic to expect the results of behavioural experiments alone to settle disputes between rival theories. The problem is not, of course, unique to behaviour analysis; A. J. Clark (1933), introducing his now-classic treatment of pharmacological dose-response relations, lamented the fact that empirical dose-response data could be equally well described by a host of different equations. Clark argued that the pharmacologist should be less concerned with the question of which equation best fits the data than with the question of which equation models a process that might reasonably be attributed to an underlying physicochemical mechanism. In the event, Clark's choice of the hyperbolic form that is implied by the mass action law of reversible chemical reactions has been essentially vindicated by subsequent discoveries about the nature of pharmacological receptors, even though its fit to experimental data is seldom ideal due to the operation of mechanisms ("spare" receptors, "transducer" processes, etc.) that were unknown in Clark's lifetime.

Following Skinner, behaviour analysts have tended to look askance at such reductionist approaches to model building, arguing that valid functional relations will remain valid whatever the internal workings of the organism. Skinner's eschewal of neurobiological explanations of behaviour, however, was in part pragmatic; neuroscience was not ready, in the 1940s and 1950s, to provide a satisfactory account of the mediation of behaviour-environment interactions. Such accounts are now conceivable, and it may behoove us to place increasing reliance on physiological plausibility as a criterion for selecting equations to model operant behaviour.

A practical implication of this approach is that a model of behavioural processes that purports to be physiologically realistic is amenable to testing and refutation by neurobiological experiment. For example, neuropharmacological interventions may be expected to exert coherent effects on the value of a parameter in whatever context that parameter appears. Killeen has shown how his model can be applied to transitional and steady-state behaviour in a broad range of reinforcement schedules, thus inviting the challenge of such neurobehavioural testing.

## What defines a legitimate issue for Skinnerian psychology: Philosophy or technology?

Hank Davis

Department of Psychology, University of Guelph, Guelph, Ontario, Canada N1G 2W1

Electronic mail: [psyhank@vm.uoguelph.ca](mailto:psyhank@vm.uoguelph.ca)

Without wishing to sound unduly glib, my immediate response to seeing this target article in *BBS* was to wonder "What's a nice kid like you doing in a place like this?" I have read *BBS* for over

ten years and do not recall seeing detailed comparisons between interval and ratio schedules of reinforcement, or phrases like "concatenating zetas" in these pages before. What has caused this sudden and bold fraternizing across enemy lines?

A closer inspection of Killeen's paper reveals the answer. He purports to address the domain of cognitive concerns that are a central agenda of *BBS*, while remaining true to the radical behaviorist tradition within which he has worked for more than a decade. That is quite an accomplishment.

In the process of addressing short-term memory in strictly behavioral terms, Killeen has also come to grips with a knotty philosophical issue: How can experimenters be sure that their definition of a response (or a stimulus, for that matter) is equivalent to the organism's? Although no one wants to open this can of worms here, the equivalence between subject and experimenter's perception is rarely a given even when humans are involved in both roles. Because we share language and sensory capacities with human subjects, the issue rarely surfaces as a serious consideration (cf. Lubinsky & Thompson 1993). However, when one is dealing with an interspecies situation, as Killeen is, the question involves more than philosophical pedantry. It can be a major practical concern.

Although I have a number of specific questions about Killeen's position, the major focus of my commentary is a more fundamental issue that transcends the discussion of beta weights and coupling coefficients. As one who was also trained as a radical behaviorist, I find it fascinating that a Skinnerian such as Killeen is willing to address questions about matching an animal's perceptions with our own. I recall attending a debate about 30 years ago between Skinner and a physicist (whose name escapes me). Their topic was the nature of the stimulus. Skinner argued that specifying the physical properties of the stimulus in great detail was a pointless endeavor because ultimately it was unknowable to which aspects of the stimulus the rat or pigeon was responding. It was for this reason, Skinner argued, that he was hesitant to use reaction-time data in conditioning experiments. Time between presentation of a discriminative stimulus and emission of the response was only useful if we could be certain that the animal's perception of the cue began the moment it was presented. In short, getting "into the animal's head" landed us solidly in the realm of metaphor.

Thirty years later, a productive scientist from the same behaviorist tradition as Skinner is willing to address, in no uncertain terms, the question of how closely the experimenter's definition of a response matches that of his animal subject. I hasten to add that few will perceive Killeen as "selling out" in raising this question. As he claims in the target article, this "is a non-metaphorical theory," and I doubt that many commentators will be refuting that judgment.

How has Killeen achieved this alchemy? Why has this matter become, barely three decades later, a legitimate question? The answer seems self-evident. Killeen has evolved a technology within the framework of behaviorism that allows him to address the question. This may not seem much of a revelation, but it has a troubling side. It suggests that issues that were earlier denounced as irrelevant to our knowledge base or better left to misguided cognitive psychologists and philosophers suddenly become relevant when one of our own is clever enough to bring them under scrutiny in a manner publishable in the *Journal of the Experimental Analysis of Behavior (JEAB)*.

This is my biggest concern with the target article. To be fair, it is not a concern with Killeen per se but rather with the consistency of his philosophical heritage. The most cynical reading I can give it suggests that philosophical legitimacy is not based on logic but rather on technology.

I will raise one other concern with Killeen's position. For the first 25 pages or so, there was no doubt in my mind that this was a theory about *operant* conditioning, emitted behavior, consequential control. There was nothing in the language or the data cited in support of his theory to suggest otherwise. Indeed,

more than half (48 out of 94) of the journal articles cited by Killeen are drawn from *JEAB*, the ultimate bastion of Skinnerian psychology.

Why, then, in section 9.3.1, para. 2, does Killeen refer to taste-aversion learning as part of his general case? I believe, as do most psychologists, that taste-aversion learning is a special or "prepared" subset of Pavlovian conditioning (e.g., Seligman 1970). Although it is true that such learning involves long delays, which are important to Killeen's theory, it is also true that these delays are between conditioned and unconditioned stimuli rather than between response and reinforcer. Killeen has already made it clear that time in short-term memory is event driven. When these events are overt, skeletally induced behaviors, as they are in most of Killeen's examples, the picture is fairly clear. In bringing taste-aversion learning into the discussion, however, Killeen risks drawing a chain of covert, visceral events into the marking of time. Worse yet, the possibility exists that time may be marked by a mixed sequence of these two classes of events. This is hardly the "dimension of homogeneous operant responses" to which the author wishes to confine his theory.

Does Killeen wish us to assume that his theory is general to both operant and Pavlovian conditioning? If he does, I would prefer to see this rather important extension made explicit. If he does not, as I believe he implies in section 9.3.1, then why import such data in support of his position? While hinting at broader generality, this reference detracts from an otherwise rigorous argument.

## **What do reinforcers strengthen? The unit of selection**

John W. Donahoe

*Department of Psychology, Program in Neuroscience and Behavior,  
University of Massachusetts at Amherst, Amherst, MA 01003  
Electronic mail: burgos@psych.umass.edu*

Evolutionary theory makes a crucial distinction between the actual *focus* of the process of natural selection by the environment and the *products* of that process. This distinction is sometimes expressed as one between selection *for* and selection *of* a characteristic (Sober 1984). For example, a certain constellation of genes might increase in frequency in a population, not because a given individual possessing those genes would necessarily have an increased reproductive fitness, but because the survival of members of a set of related individuals possessing those genes was favored. Thus the focus of natural selection is *for* increased gene frequency in the population but the resulting product of selection may be a decreased probability of survival for some individuals possessing those genes (cf. the analysis of altruism as an instance of inclusive fitness; Hamilton 1964). The true focus of natural selection is the gene, not the individual. Or, in other terms, the unit of selection is the gene as expressed in that environment and not the individual possessing those genes – although the occurrence of the two entities is often highly correlated.

The distinction between selection *for* and selection *of* applies with equal force and benefit to the understanding of reinforcement. Reinforcers are conventionally said to "increase the frequency of the responses they follow" (Reynolds 1968, p. 6), which suggests that the focus of the selecting effect of reinforcement is the response itself. This view of reinforcement is respected throughout much of the target article but is inconsistent with Skinner's initial use of the term. [See *BBS* special issue on the work of B. F. Skinner: "Canonical Papers of B. F. Skinner" *BBS* 7(4) 1984.] To wit: "Conditioning is . . . the study . . . of the operation of reinforcement and its effect upon reflex strength" (emphasis added; Skinner 1938, p. 62). In

operant procedures, which are the subject of the target article's theoretical analyses, "the reinforcing stimulus is correlated with a response . . . [but] the effect is a change in [s. $R^o$ ]" (Skinner 1938, p. 66), where [s. $R^o$ ] symbolizes a relation between a stimulus and a response that Skinner termed a "reflex." In short, for Skinner the unit of selection was not a response but a *relation* between a response and its guiding stimulus. What was affected by reinforcement was an *operant* – an environment-behavior relation – not merely a response. (It is true, however, that – under the influence of positivistic philosophy – Skinner subsequently diminished the role of the guiding stimulus because of difficulties in its specification for many operants; cf. Coleman 1984. To that extent, Skinner conflated a pragmatic point with an epistemological one.)

If what is selected by reinforcement is an environment-behavior relation, then response frequency is an inappropriate measure of selection unless the stimuli guiding the various occurrences of the measured response are the same. If a response of a given criterion topography is guided by different stimuli – which is the case except under the most highly controlled conditions – then counting all instances of the same response topography as a measure of the strength of one operant invites confusion.

This is not a trivial or purely technical matter. Defining the subject matter of behavioral science on the basis of response topography alone is largely responsible for difficulties in disciplines such as linguistics. For example, when considered outside of the circumstances in which it occurred, the sentence "The shooting of the policeman was awful" is ambiguous and invites unfortunate attempts at disambiguation through appeals to syntactical rules. However, when uttered at a shooting range or in the presence of a wounded policeman, the sentence is not ambiguous. As another example of the inadequacy of response topography as the sole measure of the effects of selection, the neuropsychological patient who upon hearing the phrase "Roses are red and violets are blue" continues with "sugar is sweet and so are you" may be unable to say "sugar" in response to the question, "What is the sweet substance that you put in coffee?" Responses of the same topography selected in different circumstances are members of different (albeit complex) operants. To measure the strength of the "sugar" response by counting its overall frequency is not a useful enterprise.

As applied to the target article, the acceptance of all occurrences of responses of the criterion topography as instances of the same operant encourages the introduction of various parameters to accommodate the fact that when operants are so defined they bear no simple functional relation to reinforcement. Moreover, this practice discourages the search for stimuli that do guide the diverse operants. Space permits consideration of only one such parameter,  $M_n$ , the "contents of the subject's memory after the  $n^{th}$  response." The issue is not whether a reinforcer may have different effects on a response as a function of the environmental and neural context in which that response occurs – this is clearly the case – but whether understanding such differences is advanced by postulating a "memory" in place of identifying that context. First, experimental analyses at both the behavioral and neural levels (e.g., Stein & Belluzzi 1989) indicate that the action of reinforcers is restricted to what Skinner called "the moment of reinforcement" (e.g., Ferster & Skinner 1957, p. 3). Introducing a "memory" for past responses rather than seeking to identify the different contexts in which responses of the same topography may occur is inconsistent with the momentary action of reinforcers. (Again, the issue is not whether the action of reinforcement may appear to extend backward over appreciable time periods – selection *of* – but whether the reinforcer does, in fact, select temporally remote responses – selection *for*. The latter is a proposition for which experimental analysis provides no direct support.) The fact that the cumulative effect of the momentary action of reinforcement yields longer-term regularities that are captured by the notion of

"memory" does not force the postulation of "memory." To do so is to risk confusing products of selection with their causes. Second, the discouraging effect of "memory" on the search for guiding variables is illustrated by the conceptualization of the response requirement ( $N$ ) in ratio schedules as affecting the number of responses residing in  $M_n$ . Since with small ratios the value in "memory" may be less than the maximum value, the effect of a reinforcer on response frequency is held to be less than if the contents of "memory" were at the maximum value more likely present with larger ratios. This account does not exploit the fact that at small ratios the response preceding the reinforcer occurs shortly after the preceding reinforcer and therefore that such responses may be partially guided by the discriminative effects of the reinforcing stimulus. Thus, successive responses within ratio schedules having different values of  $N$  are almost certainly not members of the same operant class and should not be regarded as different strengths of the same response.

There is a great deal that is good in Killeen's thoughtful and beautifully written paper, but the proposition that "the present models are simple because they take as their domain homogeneous response sequences" may be claiming a domain populated by the behavior of "stat-rats" rather than organisms.

## Moving beyond schedules and rate: A new trajectory?

Gregory Galbicka

*Department of Medical Neurosciences, Walter Reed Army Institute of Research, Washington, DC 20307-5100  
Electronic mail: gag@wrair-emh1.army.mil*

Killeen's axioms of activation, temporal constraint, and coupling provide the basis of an effort laudable for its attempt to derive laws about steady-state as well as transition-state behavior from a simple set of fundamental axioms, not assumptions, about how organisms function as "organic mechanisms." No appeals to optimization, maximization, or matching of reinforcers, commodities, or resources are needed. They are replaced by three observations: "incentives fuel behavior . . . , time constrains it . . . , and contingencies direct it." Organisms are viewed as neither rational nor irrational, neither economists nor gamblers, they are simply biological motors revved by incentives, incapable of distinguishing correlation from causality, making responses that previously occurred in conjunction with reinforcers (or at least ones remembered as having occurred contemporaneously). An over-simplified view, perhaps, but a virtuous simplicity in the present context of initial modeling.

As modeling is not my forte, I will leave it to others to criticize the formal properties of the model in detail. Superficially, it appears firmly grounded, synthesizes a large range of extant research, raises many questions each of which is eminently testable, and as such it fulfills my simple criteria for a valid contribution to the field. I would like to address two aspects of Killeen's target article that emphasize what I see as severe shortcomings in current analyses of behavior.

First, I was surprised by the layout of the paper. The majority of the effort details coupling coefficients for traditional schedules of reinforcement, whereas the excellent exposition and interesting implications are left buried in a series of appendixes. Were I presented the text of the paper without the subdivisions, I would easily reverse the order, presenting the theory development from the axioms and dynamics as presented in the appendixes, and tacking on the derivation of the coupling coefficients for the various schedules in a series of afterthought appendixes.

The editorial contingencies shaping Killeen's exposition in its present form constitute my first complaint – schedules of reinforcement have so dominated the analysis of operant behavior that they all but preempt analyses of basic process. Clearly,

more direct experimental questions can be posed to help validate the model than those dealing with steady-state schedule-controlled responding; unfortunately, only the latter will contact the vast majority of research conducted on operant behavior in the last few decades. For example, Killeen's suggestion that IRTs (interresponse times) are epiphenomena of reinforcement for target and nontarget responses is directly supported by Richardson and Loughead's (1974) observation that physical restraint shortened IRTs and increased response rates under  $IRT > t$  schedules. Restraint should decrease the probability of a nontarget response occurring within a reinforced sequence, increasing  $\rho$  despite the overall decrease in reinforcement rate this entails under  $IRT > t$  schedules. Sex differences observed under these schedules (e.g., van Hest et al. 1987) also indirectly support this notion, in that females are more active in an open field and make fewer unreinforced responses under  $IRT > t$  schedules than do males, and adding objects promoting collateral (nontarget) behaviors facilitates acquisition of long IRTs in males but not in females. These results all speak to the present analysis but get lost in Killeen's heroic efforts to accommodate various schedule effects.

As I and others have noted before, simple schedules are only simple to program, not to analyze, and "to the extent ease of programming is correlated with the number of variables left uncontrolled, a continued reliance on simple schedules as analytic devices ensures continued ignorance of the variables controlling behavior" (Galbicka & Platt 1986, p. 379). This is not to argue that theories of operant behavior need not ultimately accommodate schedule-controlled behavior. Events must be scheduled, and scheduling generates clear effects. Killeen's model, however, makes explicit that the fundamental processes and not the schedules themselves must be the subject of the model. That is, variables affecting response duration, such as effort, might be expressed differently under different schedules; however, only response duration can be a unitary parameter of the model. Schedules are derived, not fundamental, determinants of behavior. As such they should play a supporting and not major role in the development of behavioral models.

The second comment is prompted by Killeen's explicit accounting for response duration via the parameter  $\delta$ . The analysis of behavior has been conducted primarily with instantaneous responses, making viable the notion of a free operant, with rate as its primary measure. But as behaviors become temporally expansive, more questions arise about the importance of an operant's freedom and the use of rate. It is particularly peculiar that we have adopted this strategy because everyday life is not full of free operants. There exist a multitude of constraints due to the presentation of establishing, discriminative, and consequent stimuli as well as to the behavior itself that curtail a response's occurrence. Writing, cooking, playing an instrument, and participating in your favorite sport are all clearly operants, and yet are clearly not all free operants. Furthermore, the rate at which I write a paragraph is much less important than other dimensions along which that particular behavior has been shaped. It is a trivialization of the concept of rate to subdivide occasions into "good" and "bad" paragraphs and to attempt to rescue the analysis by counting instances of each. Rate is only one dimension of behavior modified by its consequences – intensity, direction, quality, proficiency, and variability also qualify as potential operant dimensions.

A mathematical model of reinforcement must show how these other dimensions are shaped by contingencies in the same way it accounts for rate. For example, my recent research (e.g., Galbicka et al. 1993) has involved shaping rats to emit a series of left-lever presses (a "run") followed by a single right-lever press. Runs more closely corresponding to some target value (e.g., 12) are reinforced. Responding rapidly changes along the dimension of run length, but changes less consistently along other dimensions, such as rate. As such, mathematical models of reinforcement accounting only for rate remain severely limited

in their application to more general questions of behavior change. Although Killeen limits himself primarily to discussion of rate in the target article, the notion of behavior as a trajectory that can easily be expanded to include warping of other dimensions of behavior-space by reinforcement demonstrates his sensitivity to this limitation.

Not only does behavior change along dimensions other than rate, reinforcement can create response units that previously did not exist. In the example cited above, food pellets probably initially reinforce lever pressing, but it is possible that reinforcement or other stimuli generate a type of "behavioral chunking" (e.g., Fountain 1990; Terrace 1991) whereby the entire sequence functions as a unit. One way to test this might be to compare the amount of responding activated by a single feeding in groups of subjects trained to different target values. If reinforcement activates a fixed amount of behavior, it should be simple to distinguish whether the number of lever presses or sequences remains constant across groups. Changes in how behavior is chunked in memory would of course drastically alter predictions made by Killeen's model.

On the lighter side, Killeen has done us all a great service by quantifying an age-old adage. To wit, according to the model, memory decays as a function of responding, not time. Responding is proportional to the rate of incentive delivery. Incentive delivery is itself associated with positive affect. From this, one can easily see that time does seem to fly only when you are having fun. Of course, Killeen would probably express this as  $\text{fun} = \text{flies}^{\alpha t}$ , generating a heated debate about the form of the function. If nothing else, Killeen has provided substantial fodder for argument with his present model. That alone is worth the price of admission.

## **Has learning been shown to be attractor modification within reinforcement modelling?**

Robert A. M. Gregson

Department of Psychology, Australian National University, Canberra ACT 0200, Australia

Electronic mail: [rag655@cscgpo.anu.edu.au](mailto:rag655@cscgpo.anu.edu.au)

Killeen provides us with an extensive and meticulous coverage of the use of linear systems theory in operant behaviour analyses. I can see (almost always) *what* mathematics he is using but I find it difficult to see *why*, because he shifts his ground during the argument. His paper (Killeen 1992) on "mechanics of the animate" gave me more certain intuitive insight into his style of theorising, however Appendix D, which comes close to nonlinear dynamics, lacks precision both in the use of terminology and in the use of the ideas he employs. In Table 4 only Construct 1 is a trajectory equation, and the representation of the dynamics as a path between a family of attractors onto a point attractor, which is implicit in his remarks on Figure D1, is nowhere formalised, just hinted at.

There are three interwoven strands to Killeen's argument: calculus, time series, and attractor dynamics. The first is fine as deterministic theory, though it does not seem to match data in Figures 5, 14, and B1 very closely and the graphs tell us nothing about goodness of fit per se. I am not asking for significance testing, because I do not espouse Neyman-Pearson inference, but, given that there is a range of competing models and approximations to them (e.g., Killeen's truncated Taylor series) and even AIC (Akaike Information Criterion) theory gets its appropriate citation, we might be favoured with at least a likelihood ratio now and then.

What I question fundamentally are two things:

1. The treatment of the coupling between two time series, one within the organism and the other set by the experimenter (perhaps contingently upon the first). I ran some simulations as

soon as I got Killeen's target article, but instead treated the internal behaviour, that is, Killeen's core Equation 2, as an AR (autoregressive) process, and reconsidered the problem as one of seeking optimum transfer functions between an AR of a priori indeterminate order and a fixed order AR process used as a forcing function by the experimenter. It became apparent that one needs more assumptions to run such a model and that the variance estimates Killeen sometimes gives without derivation are not sustained (see Lyttkann 1963). Agreed, one gets wide changes in the shape of the linear transfer coefficients, as in his simulations, with the choice of  $\alpha$  and  $\beta$  and with the sequence of  $Y_n$  values. But Killeen jumps between considering a time series paced by an internal clock in almost continuous time and one paced by events and not times, in a sort of discrete series. I think he should combine his ideas by superposing the events on the clock and asking about the parametrization of the interevent interval distribution where events effectively become times, and times become events. There is now sufficient literature on the theory of irregular time series and on series which jump between states as a consequence of interventions at points in time for this to be considered. If one does so, that fine-grained structure of behaviour may be more readily described.

2. The use of AR descriptions instead of MA (moving average) (as Killeen prefers) seems imperative if we want to use the attractor theory he mentions in Appendix D. If Killeen regards the attractor idea as basic and is not using it just as a metaphor then he should have begun with it. One cannot say from what he reports (because his calculus is about terminal states and not trajectories) whether the behaviour dynamics are (1) on one complicated attractor such as the well-known Rössler attractor (see e.g., Fischer & Smith 1985), with the forcing function inducing jumping between lobes in the phase space, or (2) he wants as many attractors as there are potential responses with just one designated as the target, or (3) he wants to have one attractor but to push it about in the phase space, thus changing its limiting behaviour, or (4) whether he wants to remove noise progressively from one single attractor (Kapitaniak 1990) because this choice of variants does not have simple consequences.

All this may sound esoteric, but if Killeen is serious about using modern physics and not the eighteenth-century kind that handles Table 5 quite well, then we want to know what the trajectories are before the system gets to a steady state. Killeen does not seem to be aware of this when he refers to the difficult problem of estimating how long it takes to learn under a schedule.

All the equations he derives in Table 4 for the different schedules of reinforcement do constitute a coherent set but they are about terminal relationships and not about the dynamics of getting to what he calls the consequential lines. The paths to limit points in Figure D1 are not themselves the attractor equations. I ask whether the behaviours Killeen reinforces are in the limit perfectly stereotyped and invariant (as I gather some pigeon pecking is), because unless they are, we do not want a point attractor representation for them. Killeen's references to metrics in section 14.1.3 are unconvincing; the properties he describes do not necessarily or sufficiently imply trajectories embedded in metric space. After all, dynamics are often in  $D_2$  space where  $D_2$  is not an integer. The reference to Bayes' theorem in section 14.1.2 is equally unclear; perhaps Killeen means that putting a demonstrably erroneous expression for a conditional probability into Bayes' equation would lead to a wrong calculation of an inverse probability. That is surely something with which any Bayesian would agree. Also, the assumption that an additional Gaussian process creates and imposes the uncertainty on when the behaviour crosses a boundary (Killeen's comments on Figure D3) is not necessary as soon as one uses attractor theory, because the uncertainty is already built into the dynamics. Perhaps I should observe that Gaussian noise can be the consequence of an infinite dimensional attractor and that both psychological and physiological data now show

that such dimensionality is excessive to model behaviour and its substrate processes.

One can write attractor theory that brings together learning, trajectories, and neural networks (Gregson 1993), and we can expect in the future to have more useful and subtle examples of this approach if only as a currently fashionable paradigm, but one cannot do it by starting at the output end and only using asymptotic stabilities. I think Killeen is to be thanked and applauded for making it possible for us to see the difficulties in using traditional calculus and the resultant differential equations to pin down the vagaries of behaviour dynamics.

## Reinforcement without representation

Stephen José Hanson

*Siemens Corporate Research, 755 College Rd. East, Princeton, NJ 08540;  
Cognitive Science Laboratory, Princeton University, Princeton, NJ 08544  
Electronic mail: jose@clarity.princeton.edu*

**Pandora's box.** Mathematical reinforcement (learning) theory has had a checkered past. There was once a romantic era when to be a learning theorist meant that one had a school of thought, an entire system or framework for learning, and most especially a mathematical model of reinforcement. Sadly, by the mid-1960s the cognitivist revolution swept away much of the interest in learning and with it the myriad rich mathematical formalisms that preoccupied psychologists for nearly 50 years. The mathematical models of Hull, Skinner, Estes, and many others in this epoch are considered a mere historical curiosity now. The mainstream cognitive psychologist has disdain for such "naïve" models of learning, and the typical animal learning psychologist, although he may in principle be sympathetic to the underlying concepts of the modeling, will probably argue (in commentaries accompanying this one, I suspect) that such modeling is unnecessary, premature, and has had little success to speak of.

I believe that Killeen has opened up a Pandora's box with his target article, one that simultaneously reaches back to early twentieth-century mathematical learning theory and into the present revolution of connectionist (neural network) modeling. My focus in this commentary is on representational aspects of the learning models. I will argue that learning models from the past avoided, sidestepped, hid, and ignored the representational problem that cognitivists eventually marshalled to destroy the credibility of behaviorist and animal learning theory and research.

Unfortunately, at least on this point, the cognitivist-behaviorist polarization was quite unnecessary. As connectionist models have shown, it was quite possible to incorporate representational issues (e.g., Hanson & Burr 1990) in the context of *standard* mathematical reinforcement learning models. It was paradoxically cognitive kinds of arguments that led (accidentally) to the recent explosion of such learning models. It is hence possible to argue that connectionist learning models such as back-propagation form a "land bridge" (albeit a narrow one) between the mathematical learning models of the past and the cognitivist research. Viewed in this light, Killeen's target article provides the necessary land mass at the thinnest part of the isthmus near the behaviorist forefront.

**The Art Deco of learning theory.** To clarify this point we have to return to the more stylized, grander time of mathematical reinforcement theories, what one might call the Art Deco period of mathematical psychology. It is instructive to examine the formal aspects of several of the early mathematical learning models and compare them with one of the simplest versions of a connectionist learning rule: the delta rule.

The next three models have basically the same form. Learning accumulates in strength by some difference between present strength and some asymptotic strength available in the system. For Hull (1943) the variable was "habit" strength:

$$\frac{d_s H_R}{dt} = i(M - s H_R) \quad (1)$$

$M$  is an upper limit of habit strength dependent on the magnitude and delay of reward. The parameter  $i$  controls the proportion of growth of habit strength on each trial  $n$ .

For Estes (1950) a probabilistic formulation was proposed:

$$\Delta p_n = \Theta(1 - p_n) \quad (2)$$

Instead of conditioning strength, Estes introduced the idea that responses were conditioned to various fractions of the stimulus environment: the stimulus environment controlled the elicitation of the response by the amount of prior conditioning to the stimulus environment. During reinforcement, stochastic sampling of the environment by the organism's sensors results in the binding of some further fraction of the stimulus environment to the response. Over  $n$  trials, the response tends toward an asymptotic probability of 1.0. Estes appropriately named this learning model stimulus sampling theory.

Both these types of models are generic enough to apply to either instrumental or Pavlovian conditioning. In the early 1970s there was a proposal by Rescorla and Wagner (1972) for a general Pavlovian model that incorporated stimulus compounds. This model was very general in form and could account qualitatively for many kinds of Pavlovian conditioning effects arising from interactions between stimuli. For Rescorla and Wagner, their variable of interest was associative strength:

$$\Delta V_n^i = \alpha \beta (\lambda - \sum_{i=1}^n V_n^i) \quad (3)$$

Although the form of this model is identical to the ones above (Equations 1 and 2), there are two important differences: one difference appears in the asymptotic strength that is now a parameter; the second includes the computation of the present associative strength of a given stimulus that depends upon the sum of associative strengths from other stimuli. This will be a key point below. So although the form is identical to the learning models above, Rescorla and Wagner introduced a factor in the growth of the learning curve that could mediate the change from the environment. Note that depending on the number and kinds of stimuli and their controlling laws, this could allow the learning curve to change in very complex ways. Unfortunately, Rescorla and Wagner focused on simple second-order interactions, leaving more complex possibilities to the imagination.

**Whining cognitivists.** It is interesting that all three formulas above had at least a decade of dominance in mathematical learning theory and behavioral psychology. Each receded from interest when cognitivists complained that such models had no relevance to "real" learning. "What's happening during learning?" they would ask. "It can't just be a smooth change in behavior to some asymptotic performance." Even the myriad impressive statistical predictions derived from stimulus sampling left one a bit cold: in theory, after all, learning curves weren't even smooth!

Unfortunately, the whining cognitivists were right. There was no attempt to deal with what might be happening "behind" the learning curve. Each model was attempting to deal with a mapping between the correlation of the reinforcement and the change in behavior. Each in their own way left the door open for representational descriptions – habit strength, stimulus sampling, associative strength – but each scrupulously avoided being pinned down. Where in the above equations could representation appear?

**Back to the future.** Well, no one really asked that question. Representation was studied, but without interest in learning or especially in mathematical learning theory. Incredibly, for several decades it was as if learning just did not happen or did not need to be studied. By the early 1980s, psychology was primed

for learning research and through some sort of sudden release from a two-decade-old inhibition, mathematical psychologists (e.g., Rumelhart et al. 1986) were able to burst in with a simple mathematical learning model not too different from the ones above. Indeed, it looked identical, which was true horror for antibehaviorists or antiassociationists (e.g., Fodor 1989) who thought they had gotten rid of this stuff years ago. There was one not so little difference, however: *The model could arbitrarily represent (approximate) almost any function.*

In contrast, the simple prior models represent single values along a one-dimensional curve. Why does this matter? To end the story we have to look at one last equation: consider the delta rule,

$$\Delta w_n = \alpha(t - f(x, w_n)) \frac{\delta f}{\delta w_n}$$

If you squint you can see the family resemblance to the other equation, and in the second term (in the difference) of the equation there is a general function  $f$  of some input vector  $x$ , with  $t$ , as usual, the asymptotic learning value. This is where the trick occurs. If it were a neural network, many interesting things would then happen:

1. The input dimensionality can increase dramatically, which is much more compatible with the requirements and experimental work of cognitive psychologists.
2. Learning is a function of underlying constraints on the complexity of the stimulus space.
3. Arbitrary transformations, mappings, and representations can form in the service of the task the system is set to learn.

Surprisingly, through one relatively small expansion of a term in a 50-year-old equation, learning and representation becomes coupled. The results for psychology and AI have been nothing less than a major paradigm shift.

**Coupling coefficients.** From this perspective, we see Killeen's contribution in a slightly different light. In the tradition of the Art Deco period, Killeen has introduced a framework for reinforcement theory and a system for characterizing behavior under various reinforcement schedules. He has carefully introduced a representational scheme that controls and delimits specific functions representing the vagaries of short-term memory in organisms constrained to learn arbitrary mappings between their behavior and memory for temporal and spatial constraints of reinforcing events. I believe that Killeen's coupling coefficient, a seemingly benign mathematical device like the delta rule, opens a Pandora's box of models, each of which adds further stones to the land bridge between the behaviorists and cognitivists.

## Integration and specificity of retrieval in a memory-based model of reinforcement

Marvin D. Krank

Department of Psychology, Mount Allison University, Sackville, New Brunswick, Canada E0A 3C0  
Electronic mail: mkrank@mta.ca

Killeen's demonstration that a mathematical model of response and incentive memories can account for behavior on several classic schedules of reinforcement represents an important step in the development of a "cognitive" theory of reinforcement. Particularly noteworthy is the suggestion that short-term or working memory limitations determine the organism's definition of a response and that the simultaneous presence of response and incentive memories generates responding and governs response rates. However, the model does not live up to its claim as a general theory of reinforcement. Features such as stimulus control and operant-Pavlovian interactions, which are fundamental to understanding motivational effects of reinforce-

ment, are not discussed. More importantly, the nature of these omissions points to a general problem: Killeen's memory model does not include a sufficient role for retrieval processes.

This commentary will highlight the importance of retrieval processes to reinforcement effects and will suggest an alternative view of response and incentive memory that emphasizes the encoding of particular learning episodes with specific information and the importance of similarity-based retrieval.

The very nature of the learning tasks modeled in the target article begs the question of retrieval. Free-operant schedules do not have explicit cues for reinforcement. Thus, retrieval processes are not emphasized even in a memory-based account. Killeen specifies only a general retrieval parameter rho, which, although pivotal to the model, is not clearly defined. In the free-operant situation, however, various stimuli such as the general context of the operant chamber, temporal cues, reinforcement-generated cues, and so on might serve as retrieval cues.

Direct evidence of retrieval effects comes from studies of operant-Pavlovian interactions (Overmier & Lawry 1979). In a typical study, an operant baseline is established with the reinforcer. Next, a conditioned stimulus (CS) is trained with the reinforcer as the unconditioned stimulus (US). Finally, operant responding is reestablished and the CS tested during ongoing responding. Although the outcome of this procedure can be complicated by CS-elicited behavior that competes with the operant, the CS often increases the rate of operant responding (Karpicke et al. 1977; Krank 1992). This enhancement of response rate is consistent with the CS serving as a retrieval cue for reinforcer memory with incentive effects.

Another procedure providing evidence for retrieval effects uses differential outcome training. Two distinct operant responses are trained in association with two distinct reinforcers. For example, in my own work with alcohol, lever presses in one location are reinforced with an alcohol solution, whereas lever presses in another location are reinforced with liquid sucrose. Once baseline levels are established, half the animals are given response-independent training with a CS paired with the alcohol solution and the other half with a CS paired with liquid sucrose. In the transfer test, the CS is imposed during a choice test with both alternatives. On the choice test, the CS associated with the alcohol solution induces more responding on the operant also associated with the alcohol solution, whereas the CS associated with liquid sucrose induces more responding on the operant also associated with liquid sucrose (Krank 1992; cf. Kruse et al. 1983). Such reinforcement-specific transfer effects demonstrate a qualitative representation of reinforcer attributes (cf. Rescorla 1990a; 1990b).

How does a representation of specific reinforcer attributes translate into selective increases in responding? One approach, different from the target article, is to treat reinforcement as a process of encoding and retrieval of particular learning episodes. The representation of learning as the integrative encoding of specific features of particular learning experiences is common to a number of recent approaches to human performance in various domains (see Brooks 1987; Hintzman 1986; 1988; Jacoby & Brooks 1984; Medin & Schaffer 1978; Whittlesea 1987; Whittlesea & Brooks 1988; Whittlesea & Dorken 1993). Retrieval, in these exemplar-based approaches, consists of the parallel activation of particular memories based on the principle of similarity (e.g., Hintzman 1986; 1988; Whittlesea 1987; cf. Pearce 1987). Memory is a weighted sum of all memory traces similar to the present retrieval conditions. Hintzman's (1986; 1988) quantitative episodic model simulates data from studies in concept learning, cued recall, recognition memory, frequency judgments.

Hintzman's (1986; 1988) model can account for specific transfer effects seen with differential outcome training. On a test trial, CS presentation retrieves memories of similar past experiences, including the particular features of the reinforcer paired with the CS during the training phase. Initially, because this

Pavlovian training involves response-independent reinforcement, the response memory is ambiguous. Hintzman (1986) solves the problem of ambiguous recall by repeating the retrieval process using the memory generated by the original cue. The result is retrieval of the response alternative associated with the same reinforcer as the CS cue.

An episodic memory model can generate free-operant behavior assuming that context cues are stored in memory traces and act as retrieval cues. In the free-operant situation, higher rates of reinforcement increase the number of memory traces that include reinforcement attributes. Thus, retrieval of reinforcement memories and incentive-based responding become more likely. By contrast, responding falls off at high ratio requirements and long interval requirements because of an increase in the number of memory traces encoded without reinforcement.

The episodic model also predicts the retrieval of specific response attributes. Ratio schedules accordingly generate higher rates of behavior because response attributes such as short IRTs (interresponse times) or larger numbers of the target response (cf. Platt & Day 1979) are encoded and retrieved with reinforcement. Interval schedules generate lower rates of behavior because response attributes such as long IRTs (Platt 1979) or other nontarget behaviors (see target article) are encoded and retrieved with reinforcement.

Taken together, these observations indicate that the episodic approach can, in principle, account for free-operant responding. Whether the principles of integrative encoding and parallel, similarity-generated retrieval can generate parsimonious accounts of existing data as impressive as Killeen's model remains to be seen. Nevertheless, grouping retrieval factors into coupling coefficients does not allow an adequate level of discourse for studies involving operant-Pavlovian interactions. Details of encoding and retrieval processes such as those described here are essential for a general memory-based theory of reinforcement.

## Memories and functional response units

Kennon A. Lattal and Josele Abreu-Rodrigues

*Department of Psychology, West Virginia University, Morgantown, WV 25506*

**Electronic mail:** lattal@wvnvm.bitnet

Constructs like memory are used often by psychologists to organize and summarize functional relations between behavior and its controlling variables (e.g., Zuriff 1985). Behaviorists have criticized the memory construct and related constructs that are defined circularly, imply causation, or implicate processes taking place in other "universes of discourse" (Branch 1977; Palmer & Donahoe 1992; Schlinger 1992; Skinner 1974; for a related critique of the memory construct by a cognitive psychologist, see Watkins 1990). Killeen returns to the memory construct as a way of organizing behavioral history effects as they contribute to reinforcement schedule performance. Specifically, memory, and the action of reinforcement on memory, is proffered as a way of taking into account the "animal's point of view" in a reinforcement schedule performance. Here we examine Killeen's use of the memory construct and then consider other ways of taking into account the animal's point of view. [See also Dawkins: "From an Animal's Point of View" *BBS* 13(1) 1990.]

Killeen's memory construct goes beyond a summary of behavior-environment relations as he makes several inferences about its physiological basis (e.g., sects. 2.1 and 9.3.5) and suggests that reinforcement operates directly on memory processes (representations) rather than on behavior. Although the physiological basis of behavior is not in question, describing behavioral mechanisms in physiological terms can undermine the role of environmental events in determining action (Skinner

1938; 1950). Remembering has ultimate, structural limitations; however, the present analysis seems to emphasize such physiological limitations without fully acknowledging the largely unexplored influence of environmental variables on remembering. For example, within the context of a reinforcement schedule performance, the number of responses that can be affected by reinforcement seems to be determined at least as much by environmental factors as by constraints on the organism's physical structure.

When the animal's point of view is represented as a memory process, the action of the reinforcer is not on the response but on its representation in memory (see the concluding section of the target article). Following the logic of a reinforcement analysis, the memory of the response is strengthened by reinforcement. Thus, on future occasions of a discriminative stimulus, the memory of a response is evoked and this in turn must cause the behavior. As Killeen states it, the environment is "filtered" by memory processes. Even if we assume that the reinforcement of a memory representation (as opposed to a response) is a logical extension of the reinforcement concept, the familiar problem of circularity remains. The evidence of memory-strengthening is the changed behavior, which in turn defines the strengthened memory.

Most successful experimentalists in psychology experience early in their careers first-hand evidence of the dictum that "the subject is always right." This aphorism emphasizes the importance of the subject's point of view and has been discussed more generally and formally as the distinction between procedural or operational points of view and functional ones (e.g., Catania 1969). Killeen's resolution of the frequently observed discrepancy between structural operations imposed by the experimenter and the functional effects of those operations on the subject's behavior is to invoke the memory construct: the animal responds to its memory of the event and not simply to the objective event. It does not follow, however, that taking the animal's point of view in the formation of behavioral units requires invoking memory or representational processes.

Progress in the understanding of both human and nonhuman animal behavior seems more likely if one considers "what the animal knows" in functional terms rather than in memory ones. The spirit of Killeen's solution thereby might be restated simply in nonmemory terms as follows: functional effects are more important in understanding behavior than are structural, experimental arrangements. For example, putative reinforcers and punishers sometimes do not respectively increase or decrease target responses, leading the experimenter to reformulate their taxonomic status in functional terms. Moreover, regardless of the care with which contingencies are arranged, their behavioral effect may be different from that expected. There is nothing new in such observations, for they are the foundation of the experimental analysis of behavior – and they have been formulated clearly without invoking events taking place at other levels of discourse.

There are numerous experimental analyses of the animal's point of view that do not involve invoking memory constructs or processes. Indeed, Page and Neuringer (1985, Experiment 4) offer experimental evidence of the *absence* of memorial processes in the learning of complex sequences of behavior. Both Killeen and the first author of this commentary concurrently developed techniques for determining the animal's point of view empirically by "asking" the animal to report different reinforcement contingencies (Killeen 1981b; Lattal 1975; 1979). Arbuckle and Lattal (1988) showed how multiple response sequences are organized into functional behavioral units by the systematic reinforcement of response sequences when brief, unsignaled delays of reinforcement are added to what were previously immediate-reinforcement schedules. Rather than considering the unexpected increases in response rates sometimes obtained with briefly delayed reinforcement a perversion of the reinforcement contingency, a fine-trained analysis of successive

interresponse times in relation to each reinforcer was fruitful in accounting for the increased response rates.

There is significant value in integrative theories of reinforcement that take into account more precisely the role of the organism's history in describing current performance. The experimental literature on behavioral history effects is largely qualitative and analytic (e.g., Barrett 1986; Freeman & Lattal 1992). Killeen's mathematical formulations are synthetic and may add quantitative rigor to the understanding of behavioral history effects. However, substituting a rather vague memory construct analysis for a functional analysis of response units seems unnecessary for describing behavior controlled by reinforcement contingencies.

## Killeen's theory provides an answer – and a question

Mary Ann Metzger<sup>a</sup> and Terje Sagvolden<sup>b</sup>

<sup>a</sup>Department of Psychology, University of Maryland, Baltimore, MD 21228  
and <sup>b</sup>Department of Neurophysiology, University of Oslo, Oslo, Norway  
Electronic mail: <sup>a</sup>metzger@umbc.edu; <sup>b</sup>terjesa@ulrik.uio.no

For some time we have been analyzing the behavior of hyperactive rats (SHR) in comparison to their nonhyperactive progenitor strain (WKY). In part, we have looked for differences in parameters of either response or reinforcement processes. To do this, we have applied the equation appearing as Equation 4 in the target article. As Killeen points out, this equation has a long history in which the constant  $k$  has been shown to measure primarily some aspect of the response requirements and the constant  $c$  to measure some aspect of the effectiveness of reinforcement processes. In a recent study (Sagvolden et al. 1993), we applied Equation 4 to the results of VI (variable interval) experiments with SHR and WKY over a wide range of reinforcement rates to determine the values of  $k$  and  $c$  for each of the two strains. We found SHR had lower  $c$  but not independent difference in  $k$ . We interpreted these results as indicating that SHR, compared to WKY, were more sensitive to reinforcers but had similar asymptotic response rates. However, one result in particular led us to question whether Equation 4 correctly represented the processes for SHR: for these rates, very high rates of reinforcement were associated with gradually decreasing response rates. This was inconsistent with Equation 4 and was reflected in low values of adjusted  $R$ -squares for the fit of Equation 4 to SHR data. Equation 12 in the target article gives a basis for explaining these findings. We accordingly applied Killeen's theory to our original data.

Since Equation 12 is an approximation, we first attempted to apply the exact theoretical equation calculating  $\zeta$  according to Equation C29. We used nonlinear estimation procedures based on least-squares and generally found that the exact equation failed to converge. We do not know whether the problem was the increased number of parameters or the autocatalytic property of the coupling coefficient (sect. 13.3.4). We therefore took Killeen's advice, set  $\beta = 0$  as an approximation, and applied a slightly modified Equation 12 to our data.

The modification we made to Equation 12 was to replace  $a$  with  $p/a$  (sect. 9.2). Although this modification is not pertinent to Killeen's main presentation, it enabled us to calculate a value for  $a$  that should transfer directly to analyses of other types of experiments for SHR and WKY. We refer to the modified equation also as Equation 12. Applying Equation 12 enabled us to calculate values for individual rats for each of the parameters  $p$ ,  $\delta$ ,  $a$ , and  $\lambda'$ .

The first finding concerned the value of  $p$ . For virtually all rats in both groups, we found  $p = 0.7$ . This agrees with the value of 0.71 calculated by Killeen for VI schedules (sect. 14.1.5). We therefore set the value of  $p$  to 0.7 for all rats. In terms of the

concepts of the author's theory, this value of  $p$  indicates that for each rat, 70% of their STM (short-term memory) capacity was occupied by the target response at the time of reinforcement.

The second finding concerned the value of  $\lambda'$ . Unlike the value of  $p$ , the estimates of the value  $\lambda'$  varied among individuals, with 5 cases having very large values. Of those 15 individuals for which nonlinear estimation converged, however, mean  $\lambda'$  for SHR was 18.1 (1/min) and for WKY was 18.2. We therefore set the  $\lambda' = 18$  and reestimated the remaining parameters for the 5 special cases. Fixing the value of  $\lambda' = 18$  had very little apparent effect on the estimated values of the other parameters and statistical analysis showed no evidence that those estimates were different from the otherwise-obtained estimates for either SHR or for WKY. In terms of Killeen's concepts, this value of  $\lambda'$  indicates that there is, on average, no difference between SHR and WKY in the rate of decay of STM.

The third finding concerned the values of the free parameters  $\delta$  and  $a$ . Means for these are given in Table 1. Since there were only two remaining free parameters in the final analysis, the comparison of the fit of Equation 12 vs. Equation 4 can be made directly by comparing the adjusted  $R$ -square for each model, presented in Table 2. These show that Equation 12 fits our data better than Equation 4.

According to Killeen's theory, the group difference in  $\delta$  indicates that SHR have responses of duration (minimum inter-response time, IRT) approximately twice as long as WKY. This is a surprising result because it is apparent that SHR generally produce response rates much higher than WKY and that they also seem to produce many responses with very short IRTs, on the order of one second. The explanation seems to rest in the group difference in  $a$ . A single incentive incites approximately 1 minute of responding in SHR compared to only approximately 15 seconds in WKY.

Overall, we found that Killeen's theory did provide a better fit for the relation between response rate and reinforcement rate. The theoretical account is especially persuasive by virtue of the remarkable consistency of the estimates of  $p$  and  $\lambda'$  for the two strains. The conclusion with respect to effectiveness of the reinforcer was similar to our previous conclusion in which we applied Equation 4 and found the reinforcer to be about four times as effective for SHR as for WKY. It was quite a different matter with respect to the conclusion about the value of  $\delta$ : here the answer provided by the new theory was accompanied by a new question.

The question concerns how SHR can so regularly exhibit greater response rates while having the larger value of the minimum IRT. The question is even more interesting when we add that we often find in SHR bar-press behavior what appears to be a separate group of short-IRT responses, on the order of 1 second or less, that may be affected differently by reinforcement rates than the more normal-appearing longer IRT responses (Sagvolden et al. 1992). We look forward to applying Killeen's theory with explicit acknowledgment of the possibility of a

Table 1 (Metzger & Sagvolden). *Means of estimated values of  $\delta$  and  $a$*

Group	N	$\delta$ (min./resp.)	$a$ (min./reinf.)
SHR	10	.015	.933
WKY	10	.007	.219

*Note:* According to a factorial analysis of variance (group by type of estimation of  $\lambda'$ ) there were no differences or interactions for either  $\delta$  or  $a$  associated with type of estimation of  $\lambda'$ . Differences attributable to group were significant for  $\delta$ ,  $F(1,16) = 15.29$ ;  $p = .0001$ , and for  $a$ ,  $F(1,16) = 9.36$ ;  $p = .0008$ .

Table 2 (Metzger & Sagvolden). Adjusted R-squares for Equations 4 and 12

Group	N	Range		Median	
		Eq. 4	Eq. 12	Eq. 4	Eq. 12
SHR	10	.19-.54	.56-.91	.34	.78
WKY	10	.59-.98	.69-.95	.85	.89

Source: Data from Sagvolden et al. (1993).

special group of short-IRT responses, and to applying the theory to rates of measured behavior as well as to the nominally reinforced response. We have been collecting such additional measures for some time now. Killeen has presented a rich and intuitively appealing theoretical context in which to analyze them. For now, we want to challenge him to explain the intuitively unappealing conclusion that the rapid-responding SHR has a longer minimum IRT.

## Extension to multiple schedules: Some surprising (and accurate) predictions

John A. Nevin

Department of Psychology, University of New Hampshire, Durham, NH 03824

Killeen's model accomplishes something quite remarkable: the comfortable accommodation of results that were not considered in his target article. The results I have in mind are the interactions that arise in multiple schedules of reinforcement and the differences between the component performances in their resistance to change.

A multiple schedule arranges two successive schedules of reinforcement, alternated in time by the experimenter, where each schedule is signaled by a distinctive stimulus. Each stimulus-schedule combination defines a component of the multiple schedule. Many experiments have shown that the rate of responding in a component with a constant variable-interval (VI) schedule depends inversely on the rate of reinforcement in the alternated component. For example, if a pigeon receives food for pecking at a red key according to a VI 60-sec schedule (arranging 60 reinforcers/hr, the constant component), the rate of key-pecking is lower if the alternated green-key component schedule is VI 12-sec, arranging 300 reinforcers/hr, than if it is VI 360-sec, arranging 10 reinforcers/hr (see accompanying Fig. 1; for a general review of behavioral contrast, see Williams 1988).

How can Killeen's model account for the fact that response rate depends on the context of reinforcement as well as the operative schedule? His expression for response rate on a VI schedule in the constant component (Equation 12) is:

$$B = \frac{\rho/\delta R}{R + 1/a} - \frac{R}{\lambda}$$

There is no term for the reinforcer rate in the alternated component, so any change in  $B$  (response rate) must result from a change in one or more of the equation's parameters. Let us proceed by a process of plausible elimination. Rho, the coupling term, cannot change because the schedule contingency in the constant component is unaffected by the reinforcer rate in the alternated component. Delta, the minimum response duration, cannot change because it is determined by the response definition, which is unaltered. Lambda, the "measured rate of decay of short-term memory," also remains constant by definition.

Killeen (sect. 5.6) suggests that delta and lambda may also depend on motivational variables, but these are unaltered in the standard contrast experiment. This leaves  $a$ , seconds of activation per reinforcer, to account for the effects of the alternated schedule.

To see whether changes in  $a$  might explain contrast, I compared the predictions of Killeen's model with the average data from conditions 2 and 3 of Experiment 1 by Nevin (1992a). Those conditions employed a constant VI 60-sec schedule (60 reinforcers/hr) in one component alternating with a VI 12-sec schedule (300 reinforcers/hr) in condition 2, and with a VI 360-sec schedule (10 reinforcers/hr) in condition 3. Response rate in the constant component was inversely related to the reinforcer rate in the alternated component, as shown by the ordinate values of the unfilled data points over BL (baseline) in the upper left panel of Figure 1.

Another reliable multiple-schedule result is that when experimental conditions are altered, response rate in the component with the richer VI schedule is more resistant to change than that in the component with the leaner schedule (for review, see Nevin 1992b). This finding is illustrated by the effects of prefeeding (average of three sessions that were preceded by feeding 60 g of food 30 min before a session) shown in the upper left panel of Figure 1. The logarithmic scale of the y-axis permits direct comparison of the relative effects of prefeeding on response rate. For both conditions, prefeeding produces a smaller relative decrement in the richer component.

Finally, I have reported that resistance to change, like re-

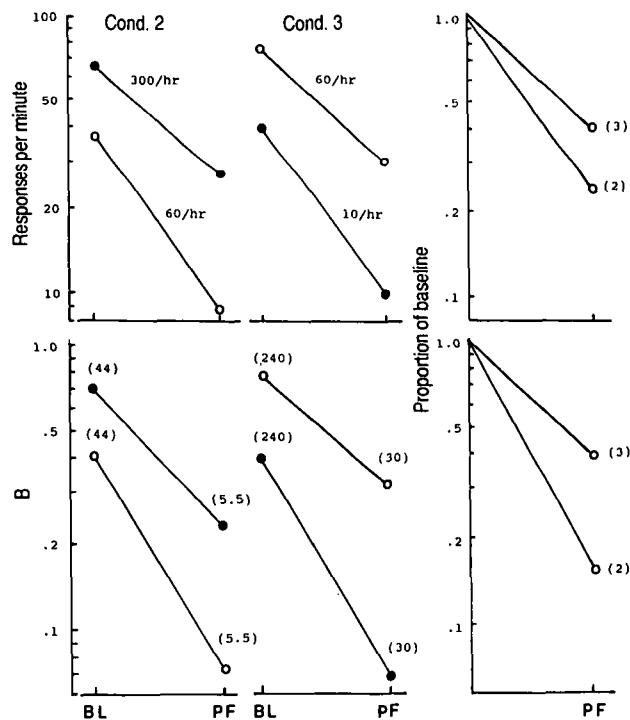


Figure 1 (Nevin). The upper left panel presents average response rates in two multiple-schedule conditions (Nevin 1992a, Experiment 1) for baseline (BL) and a prefeeding (PF) test of resistance to change, on semilogarithmic axes. Unfilled and filled circles represent performance in the constant and alternated components, respectively, with reinforcers/hr indicated by each set of data. The right-hand panel compares the relative decrements in constant-schedule response rate between conditions. The lower panels show the predictions of Killeen's model in Equation 12 for performance maintained by these reinforcer rates. All parameters were set at 1.0 except  $a$ , which is indicated for each data point.

sponse rate, exhibits contrast effects. Specifically, resistance both to prefeeding and to extinction was relatively greater for a constant VI-schedule component when the alternated schedule was leaner (condition 3) than when the alternated schedule was richer (condition 2) than the constant schedule (Nevin 1992a). This comparison with prefeeding is shown in the upper right panel of Figure 1.

The pattern of these results is well modeled by Killeen's Equation 12, as shown in the lower panels of Figure 1. I set  $\rho/\delta$  to 1.0 and  $\lambda$  to 1.0 for convenience and estimated values of  $a$  that would give a pattern of baseline response rates corresponding roughly to my baseline data. The baseline  $a$ -values are the same for the two components within each condition, but they differ substantially between conditions. It is interesting that the baseline values differ by a factor of about 6 between conditions, and the overall reinforcer rates differ by about the same factor in the opposite direction. Since  $a$  is construed as seconds of activation by a reinforcer, these estimates suggest constant total activation per session in the two conditions. This outcome may be coincidental, but it deserves consideration in future developments of this model.

If the effect of prefeeding is construed as reducing the activating power of reinforcement and  $a$  is reduced by the same proportion in both schedule components in both conditions, the resulting decrease in response rate is relatively smaller in the component with the richer schedule, as shown in the lower left panel of Figure 1. Contrast in resistance to change also falls out of these estimates, as illustrated in the lower right panel. This prediction of Killeen's model comes as a surprise, because it is exactly opposite to the plausible predictions of any previous theory of operant performance in multiple schedules (see Nevin 1992a, for full discussion).

I have long believed that a decent account of behavioral transitions would follow from an adequate theory of asymptotic performance. However, my own analyses of resistance to change seemed to defy integration with steady-state theory. Yet now, apparently without trying to do so, Killeen has developed a model that can account for resistance to change as well as steady-state performance. Bravo!

## Problems and pitfalls for Killeen's mathematical principles of reinforcement

Joseph J. Pear

*Department of Psychology, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2  
Electronic mail: pear@ccu.umanitoba.ca*

Killeen's mathematical theory is impressively comprehensive, apparently incorporating all major variables in basic operant-conditioning experiments, including response duration, reinforcement rate, motivation, incentive strength, and degree of conditioning of the response. It is not a simple theory, but it is realistic in its complexity and appears to provide a reasonable fit to the data addressed in the target article; that is, data on the rate of occurrence of a single measured response. Although I am favorably impressed in many ways with the theory, I will concentrate here on its potential problems and pitfalls.

In moving away from the conceptualization of the response as a "punctate" event that so dominates most theorizing about operant behavior, Killeen views the "behavior stream" as a continuous succession of responses with greater-than-zero durations. There is little evidence, however, that behavior is segmented in the manner required by the theory. If an animal is responding at its maximum rate, the segments that make up the individual "responses" in Killeen's sense are presumably the behavior that occurs between the animal's manipulandum operations (i.e., "responses" in the usual manner of speaking and

theorizing about operant behavior). If the animal is not responding at its maximum rate, however, then it must be emitting other responses between operations of the manipulandum, and identifying the boundaries of these responses becomes highly problematic. Suppose, for example, that an animal occasionally makes some sort of twisting loop about the experimental chamber between operations of the manipulandum (cf. Pear 1985). Does this loop constitute one response or is it composed of several responses strung together (i.e., the individual twists, turns, or "jags" in the loop). It is not clear from the target article how the theory deals with the fact that there is often no clear demarcation between responses when behavior is observed continuously. This issue must be dealt with by the theory at some point in its development, if only because it is on continuous data that it bases the rationale for predictions about standard discontinuous operant data.

Another way responses are problematic in Killeen's theory is in the assumption that they occur randomly in time, with reinforcement simply altering the probability of responses preceding it as an inverse function of the number of intervening responses (e.g., sect. 12.2). However, reinforcement can also generate sequential dependencies in responding that do not appear to be taken into account in the theory. For example, when reinforcement is made contingent on the emission of a particular sequence, animals will tend to emit the responses in the sequence specified; that is, the responses form a behavioral chain. Consider, for example, a situation in which reinforcement is contingent on a *B* response followed by an *A* response. A sequential dependency will tend to develop between the two responses such that an *A* response will tend not to occur in the absence of a prior *B* response. If the sequential dependency is perfect, the overall probabilities of *A* and *B* responses will be the same even though reinforcement follows *A* responses more immediately. The failure of the theory to take sequential dependencies into account could be the reason for the consistent underestimation by the theory of the proportion of *B* responses in Catania's (1971) experiment, as seen in Figure 6 of the target article. It should be noted that sequential dependencies may occur even when no explicit contingency exists for emitting responses in a particular sequence (e.g., Catania & Cutts 1963); thus, they are a potential problem for all the schedules to which Killeen wishes to apply the theory. For example, sequential dependencies are clearly evident in the patterns of behavior that animals engage in between responses on interval schedules (e.g., Pear 1985). Since sequential dependencies are clearly what Shimp (e.g., 1976a; 1976b) has in mind when he refers to the "structure" of response units, it is strange that Killeen suggests that his theory, which does not explicitly consider sequential dependencies, is a contribution to formalizing Shimp's concept of the response unit (sect. 9.3.5).

The fate of Killeen's theory depends ultimately on how well it predicts data in comparison with competing theories, such as those stemming from Herrnstein's (1970) theoretical treatment of the matching of relative response rates to relative reinforcement rates for concurrent operant responses. Clearly, Killeen's theory predicts behavior on ratio schedules better than Herrnstein's does; this, as Killeen points out, is because of an apparent suppressive, as well as enhancing, effect of reinforcement on response rate. But before we conclude that this suppression is due to an erasing of memory for recent behavior, as Killeen does, other explanations should be considered. For example, the suppression might be due at least in part to the time taken to consume the reinforcer or to momentary satiation. Because of its comprehensiveness, the theory could no doubt deal with these possibilities, although this very comprehensiveness is problematic with regard to determining which parts of the theory to modify to account for discrepant data.

A pivotal concept in Herrnstein's (1970) theory is that of behavior maintained by "other" reinforcers, that is, reinforcers not under the control of the experimenter. This concept is

present in Killeen's parameter  $a$ , which, as it becomes small, activates less of the behavior maintained by reinforcers under the control of the experimenter so that uncontrolled reinforcers activate relatively more behavior. Herrnstein extends his theory by giving different sources of qualitatively identical reinforcers controlled by the experimenter the same theoretical status as uncontrolled reinforcers; this allows a simple and very straightforward derivation of response-reinforcement matching on concurrent schedules of reinforcement. Killeen does not make this assumption, however, and it is not clear that this matching law can be derived from his theory. Indeed, because all responses that are followed, however remotely, by qualitatively identical reinforcers are strengthened by those reinforcers, Killeen's theory may predict more severe undermatching than the data warrant (Baum 1979). Regardless of whether this is the case, the extension of Killeen's theory to concurrent schedules will probably be far more complex than with Herrnstein's theory.

An interesting facet of Killeen's theory is the mechanism it provides for phylogenetic contributions to behavior, thus permitting some "prepared" responses to be more conditionable than other responses (sect. 9.3.2). It predicts that stereotyped prepared responses will emerge under noncontingent reinforcement but also that reinforcement for an unprepared response will lead to hysteresis when noncontingent reinforcement occurs. This is sometimes observed, but not always (e.g., Eldridge et al. 1988). A similar prediction of stereotyped species-specific behavior under noncontingent reinforcement can be derived from a generalized form of Herrnstein's response-reinforcement matching theory (Baum 1974), but hysteresis is not an inevitable outcome when an animal is shifted from contingent to noncontingent reinforcement (Fear 1988). Research is required to determine which theoretical approach more accurately describes interactions between phylogenetically determined behavior and reinforcement.

The final problem area I will mention is the prominent place of the term *memory* in the description of the theory. That reinforcers have a stronger effect on increasing the probability of responses they follow more rapidly has been known for some time without requiring memory as a theoretical construct. It is not clear why adding to that the notion that reinforcement may also have a decremental effect on responding makes it necessary to invoke the term. As Branch (1977) pointed out, using a term such as *memory* that is defined in many different ways by different individuals could hinder our understanding of the actual processes that determine behavior under various contingencies of reinforcement.

## From overt behavior to hypothetical behavior to memory: Inference in the wrong direction

Howard Rachlin

*Psychology Department, State University of New York, Stony Brook, NY 11794-2500*

**Electronic mail:** hrachlin@psych1.psy.sunysb.edu

Staddon (in press) quotes the physiological psychologist, Jeffrey Gray, as follows:

No one is immune to cognitive correctness. Jeffrey Gray, in his commentary on the Ciba Foundation Symposium [Bock & Marsh 1993], writes "I once asked a radical behaviourist what, in his view, is the difference between two awake individuals, one of them stone deaf, who are both sitting immobile in a room in which a record-player is playing a Mozart string quartet? His answer: their subsequent verbal behavior. Mercifully, there were no radical behaviourists at the symposium."

A more honest behaviorist would have answered: "One person is hearing the Mozart symphony while the other is not." The problem is that the honest behaviorist *seems* to (but in fact does

not) concede that perception and overt behavior are two different things. To see why they are one thing, consider what it means to hear or be deaf. If a person can hear, it will be possible to parse his (verbal and nonverbal) *overt* behavior into two classes, one contingent on the presence and the other on the absence of sounds, as in Figure 1, where each dot stands for an act of a given person; S is presence of a sound; NS is no sound;  $B_1$  and  $B_2$  are any conceivable division of overt behavior into two categories so that there appears a significant preponderance of acts in the diagonal boxes in Figure 1a as opposed to Figure 1b, where it is not possible to match behavior and sound. In words, the hearing person normally discriminates sounds from their absence but the deaf person normally does not. Such discrimination is what we really *mean* in everyday life when we say that a person can hear; the failure to discriminate is what we really *mean* when we say a person is deaf.<sup>1</sup>

The circled points in the corresponding upper-left boxes in Figure 1 represent Gray's example – two people doing the same thing under corresponding circumstances *at the moment* yet doing different things when their behavior is considered in its context. Another example of the same sort might be two orchestras playing the same bar of notes, one in the context of a Mozart symphony and the other in the context of a Haydn symphony (assuming one could find a single bar common to a Mozart and Haydn symphony). True, both are playing the same bar, but the orchestra playing the Mozart symphony is *in no sense* playing the Haydn symphony. Correspondingly, even though the two people in Gray's question are both doing the same thing at the moment, the hearing person is *in no sense* deaf and the deaf person is *in no sense* able to hear.

The behavioristic presumption behind Gray's question is that the two "listeners" both have lived long enough and will live long enough for a discriminative process to occur, for the acts to fill in the contingency table. Otherwise – if both listeners were born the moment before the Mozart piece was played and died the moment after – it would mean nothing to say that one heard it and the other did not.

Gray apparently had something else in mind. As a physiological psychologist, what he probably meant by deaf – versus normal – was that the hearing mechanism in the normal person's ear and brain was a normal one and that of the deaf person was abnormal in some way. Of course some such physiological story must be true; however, that story cannot be what we normally mean when we talk about "hearing" and "deafness," because these words were bandied about long before people knew about the existence of nerves, let alone about their nature. Even now, the precise neural mechanism underlying the behavior in Figure 1a is not completely known. Were Gray and his colleagues to

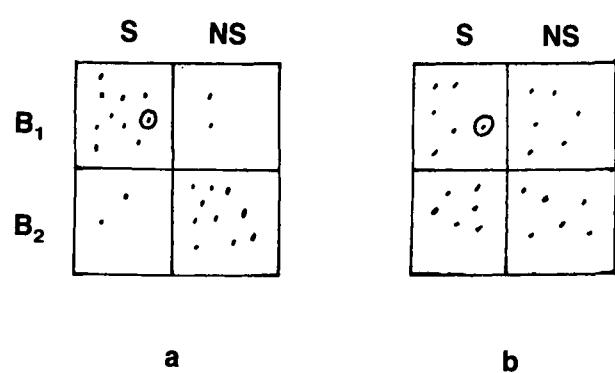


Figure 1 (Rachlin). Correspondences (points) of two classes of actions ( $B_1$  and  $B_2$ ) with presence and absence of sound (S and NS) for two people, one with normal hearing (a) and one deaf (b). The circled points in the two tables represent conceivably identical acts of hearing and deaf people in the presence of a particular auditory stimulus.

propose a hearing-mechanism candidate, and were the candidate (simulated on a computer, say) to produce an output such as depicted in Figure 1b, we (or the journal editors) would send Gray and his collaborators back to the drawing board.

Another conceivable referent for Gray's "hearing" and "deafness" would be a certain conscious experience in the mind of the hearing person, absent in that of the deaf person. There are several problems with this conception, not the least of which is the fact that deaf people may well be conscious of sounds (say, upon being struck in the head) but still be called deaf – popularly and technically – if they produce an output like that in Figure 1b. Beethoven was deaf and yet must have been conscious of sounds. Suppose the deaf person in Gray's example is Beethoven and at the very moment when the hearing person is hearing the Mozart symphony, Beethoven was humming it to himself. He would be nonetheless deaf.

Another problem with the identification of hearing and consciousness is that we sensibly attribute hearing and deafness to species of animals to which we are loath to attribute consciousness. Identification of hearing with consciousness goes against religion and philosophy (ancient if not modern) as well as science, and I do not charge Gray or (when we come to it) Killeen with seriously holding it.

What the above discussion has to do with Killeen's elegant target article is this: Killeen, like Gray, is only peripherally concerned with behavior. His model is fundamentally physiological and cognitive from its beginning, where it is justified as "plausibly attributed to simple neural networks," to its conclusion, where reinforcement is said to apply to "the animal's [internal] representation of its behavior, not the behavior itself." The crucial interactions in Killeen's model take place not between the animal and its environment but within the animal, in its memory.

Killeen, like Gray, is thus interested in the mechanisms underlying behavior. But, unlike Gray, Killeen's evidence for his model is entirely behavioral. Killeen's model is brilliantly conceived and highly refined and accounts for a lot of behavioral data. It is much as if Gray were to develop a theory of hearing that predicted exactly which points would go where in Figures 1a and 1b. The difference would be that Gray's theory would (presumably) be based on physiological data – on taking the organism apart and physically studying particular (spatially definable) mechanisms, whereas Killeen's data are the points themselves. In judging a purely descriptive theory (such as Herrnstein's [1970] matching theory) one would look for simplicity and applicability to ecological or human concerns. But Killeen's model is neither simple nor applicable in any clearly discernable way.

From a strictly behavioral viewpoint it is far from clear how one would apply this model, where reinforcement "isolates" responses from other responses, to even very simple extensions of present procedures. (Consider progressive-ratio schedules where the rate of responding depends, as it frequently does in life, on events *beyond* the next reinforcer; Hackenberg & Axtel 1993). Moreover, if Killeen's goal were just to describe the data with a minimum number of parameters, there are certainly too many here. Other theories describe as much data or more with far fewer parameters. But of course descriptive simplicity and applicability are not Killeen's goals and it would be unfair to judge his model on those grounds. How then?

The model that Killeen has built – elegant, ornate, one might even say baroque – is exposed to winds from two directions. First is physiological plausibility. Given the state of current knowledge of the physiology of memory, the model is safe. Second is the plausibility of the parameters. Are they independently constant? It is on this ground that Killeen's model may be attacked and on this ground that he defends it, in his discussion of Tables 2 and 3.

Without going over every instance, it appears that there are some interesting constancies among the parameters and so the

model has some promise. But there are also some disappointing variabilities. One example is the occasional variability of the parameter  $\lambda'$  (the basic memory parameter) with (presumably motivational) body weight. This is particularly disappointing in the light of the constancy of this cognitive parameter over motivational variables in other experiments. Killeen explains the variability by saying, "a highly motivated animal may live more in the present." But when does this maxim apply and when does it not? Some studies of self-control (Logue et al. 1988) seem to disconfirm the maxim.

Another problem, perhaps more basic, lies in the model's treatment of responses, nonresponses, and response patterns. Interval schedules are said not to reinforce long interresponse times. Fine; that was a bad idea in the first place. But instead, in Killeen's model we have reinforcers reinforcing a series of unobserved hypothetical responses during the putative inter-response interval. Like the target response, these hypothetical responses must take time. Do they then fill up the interval? Apparently not. It is possible in the model for an animal to be not responding at all – even hypothetically. So time is conceived as filled with the duration of the target response, hypothetical other responses (themselves with durations), and intervals between responses, all of which must be inferred from the rate of the target response. Surely this is too much inference, even if it were in the right direction.

#### ACKNOWLEDGMENT

This commentary was prepared with the assistance of a grant from the National Institutes of Health.

#### NOTE

1. To hear a symphony, finer discriminations would be required. S might then stand for "symphony" or "music" and NS, for "nonsymphony" or "nonmusic." There would presumably be no problem, except in cases of borderline deafness, in coming to agreement on a set of reasonable criteria.

## Memory and the integration of response sequences

Phil Reed

*Department of Psychology, University College London, London WC1 6BT, United Kingdom*

**Electronic mail:** ucjtspr@ucl.ac.uk

The proposed marriage between concepts from cognitive and behavioral psychology is to be welcomed. Especially valuable in this enterprise is the dowry from cognitive psychology, which stresses the important role played by memory processes in shaping behavior. That formal mathematical expression has been given to the concept of memory must increase the probability of its acceptance by behavioral psychologists. To this extent, the ideas contained in the target article are important and could serve as the foundation for a lasting relationship between reinforcement and cognitive theories. However, to continue the metaphor, there are two cautionary notes that should be heard along with the wedding banns for the present proposal; one is conceptual and the other empirical.

To consider the most important point, the empirical one, first; the model is silent regarding a range of phenomena concerned with the integration or chunking of responses emitted on schedules of reinforcement. For a model that attempts to utilize cognitive concepts like memory, this may be an important omission. The model appears to account for much schedule-controlled behavior with relatively few assumptions. Critically, the memory window over which reinforcement operates is assumed to vary in length with the reinforcement magnitude and in content with the schedule of reinforcement and the species used. It is these assumptions that do the work in this

model, not the putative reinforcement of particular *patterns* of responding. This appears to lead to the conclusion that the responses that fall into the memory window upon which the reinforcer operates retain their self-integrity and do not become part of a larger response unit such as an interresponse time or stereotyped response sequence (see Schwartz 1984). Postulating functional response units comprising a number of individual responses allows a solution to at least one problem that the version the model expressed in the target article does not. Under some circumstances, particularly on second-order schedules of reinforcement, the same reinforcer (presumably with the same window of effect in each case) appears to control in a similar way the rate of emission of a single response in some cases, and in other cases the rate of emission of an extended response sequence (see Marr 1979). For example, a response sequence can be reinforced, and after such a pattern is well established it can subsequently be reinforced according to a second contingency, say, a fixed-interval or an extinction schedule.

The result is that the second-order contingency often acts to alter the rate of emission of the *response sequence* rather than the individual response (Schwartz 1984). If it is assumed that the response sequence takes longer to emit than the single response, then the sequence must fill a greater portion of the memory store on which reinforcement acts than does the single response. Despite this, the reinforcer still supports differential rates of emission of the entire *sequence*, not of individual responses (Reed et al. 1991). This implies that not only has the entire sequence of responses been captured by the memory window but also that the nontarget behaviors that compete with the target responses to determine response rate have been captured in the memory window along with the sequence. This has occurred in the absence of any alteration to the value of the reinforcer or any reason to suppose that the operative memory window has altered. Thus, the present model does not appear to accommodate such findings, and without the acceptance of some form of mechanism that would allow chunking of responses to occur it may have difficulty in accommodating a relatively large set of findings. It could presumably be argued that the sequence has become chunked, takes less time to process in memory and, consequently, takes up less of the limited memory store. In the present model of reinforcement, however, there appears to be no mechanism that would allow the integration of responses to occur, and in some places it appears that the argument runs counter to such an interpretation – especially in the passages relating the model to Baum's (1973) correlational-based law of effect.

The second issue is conceptual and relates not only to this model in particular but to all mathematical and quasi-mathematical models: it is that the apparent precision in terms of prediction offered by mathematical models may become too seductive. The precise nature of the gradients relating to the gradient of reinforcement effects over a delay and the nature of the function relating reinforcement to behavior strength are two examples from the present model that have produced in the past large amounts of theory-led research with little bearing on contemporary theorizing. It is for this reason that such models should constantly be referred back to the range of empirical evidence that has not been specifically generated to test them to see whether a general theory of behavior is offered (i.e., one that allows description at the appropriate level using as few terms as possible). The above argument concerning response units is one example that may point to some changes or additions for the present model.

Lest the above sound too critical, the notion that stimuli will affect the preceding behavior to differing degrees depending upon the salience of those stimuli does allow an integration of reinforcement theory, not only with existing associative phenomena such as the marking effect but also with cognitive psychology. In studies on the effect of outcomes of differing motivational significance on judgments about the causal effec-

tiveness of the responses that precede them, it has been noted that humans integrate the responses made prior to the outcome over different time spans, depending upon the "value" of the outcome (Reed 1992). Thus, with arbitrary outcomes such as a triangle flash, integration is over a relatively short time span whereas with more valued outcomes (e.g., money) integration is over longer time spans. Hence, the present model is on the right track and should, with an appropriate eye to generality of explanation, generate future developments that may help to integrate many areas of psychology.

## Awareness and reinforcement

Charles P. Shimp

*Psychology Department, University of Utah, Salt Lake City, UT 84112*

Killeen has provided a useful service by clearly articulating some implications of an assumption that is attractive by virtue of its simplicity and its intuitive plausibility. This assumption is that reinforcement operates on what an organism remembers having recently done. It is an assumption to which one sensibly turns after discarding other even simpler but demonstrably inadequate possibilities, such as that reinforcement operates on the behavior contiguous with it (Guthrie 1959) or on recent behavior so that a delay-of-reinforcement gradient strengthens temporally more remote behaviors to lesser degrees (Hull 1943). Killeen's assumption is a perfectly reasonable step on a ladder starting at the bottom with the simplest possible assumption and ascending to more complicated and more powerful ones.

In the minds of some, short-term memory for what one has recently done is a very short step away from one's awareness of what one has recently done. Interpreted this way, the assumption Killeen considers is similar to the assumption that trial outcomes affect the recent behavior one is aware of having engaged in. Or, in short, the idea is that reinforcement affects aware behavior. This is a classic position in the history of psychology, where it is linked with views of William James and Sigmund Freud. The fundamental idea has been that awareness facilitates an organism's adaptation to the environment, or that "you're better off if you know what you're doing." More recent scholars have also promoted versions of this same category of assumption. The "buffer model" of Atkinson and Shiffrin (1968) held that short-term memory was a limited-store buffer that temporarily held representations of recent events and thus gave them a period of time during which they had a chance of being associated with trial outcomes and thereby stored in long-term memory. The famous patient H.M. was originally said to have neuropsychological complaints explainable in these terms. Killeen's elegant version of this assumption shows it still has a lot of life left in it.

However, the elegant simplicity of Killeen's assumption has been replaced in many current accounts of memory by a more complex approach that assumes trial outcomes have multiple effects on various categorically different types of memories, including unaware as well as aware memories (Jacoby et al. 1993; Schacter 1987). In general, the memory system is now often assumed to have modularized component systems. A simple, one-dimensional short-term memory that is a function only of the number of intervening behaviors does not seem adequately powerful.

In my own case, I have made frequent use of virtually the same assumption Killeen makes, and for some of the same reasons, including simplicity and the want of anything that is both better and tractable (Shimp 1978; Shimp et al. 1990). The assumption has appeared to me to be sufficiently central to my own theoretical work that, throwing caution to the winds, I submitted it to actual empirical investigation. The results have

been much more compatible with the more complex accounts of Jacoby et al. (1993) and Schacter (1987) than of the simpler assumption made by Killeen and myself. Thus, the degree to which pigeons (Shimp 1982; 1983) and rats (Shimp 1984) seem to remember their recent behavior, or to know what they have just done, does not seem to account for the effects of reinforcement on their recent behavior. It would seem likely that there is a corresponding lack of correlation between what an animal "intends" to do and what it actually does (Shimp et al. 1989). Quantitative models will apparently have to adjust for this complexity.

**Structure.** Killeen's acknowledgment of the importance of local structure to an understanding of behavior is most welcome and encouraging. He has developed his model to the point where it does indeed seem that, at least in principle, it could handle basic facts about the local organization of behavior as reflected in interresponse times. This feature of his model sets it apart from many contemporary models of behavior that do not have this capability. It will be important for Killeen to push ahead and to show that the model can handle these facts in practice as well as in principle.

**Verisimilitude.** Like most other quantitative models, Killeen's does not attempt to capture the rich detail in behavior streams of individual organisms. It attempts, instead, to capture the average characteristics, averages over both organisms and time. This is an immensely fruitful approach. However, it is disadvantaged in the contemporary context of the social and behavioral sciences where "everyday" behavior, social-constructionist perspectives, as well as research goals and methods of ethnography are increasingly valued. Naturalistic observational data and detailed descriptions of everyday behavior are playing vastly greater roles than heretofore, but Killeen, appealing to a different scientific tradition, explicitly eschews this degree of verisimilitude. It is exciting that Killeen's approach offers the potential of unifying these research traditions that so often are seen as incompatible. His model can, at least in principle, be developed into a "behaving theory" that could generate the richly detailed behavior streams of real organisms (Shimp 1989) and thus unify the ethnographic and mechanistic approaches.

## Practical effects of response specification

Richard L. Shull

Department of Psychology, University of North Carolina at Greensboro,  
Greensboro, NC 27412  
Electronic mail: shullr@uncg

Killeen reminds us of an important distinction: the response that produces a reinforcer is not necessarily identical to the response whose frequency is changed by that reinforcer. This lack of correspondence can cause serious problems for those who must deal with reinforcement at a practical level. Suppose, for example, that a young child offers the parent a gift of freshly picked flowers. The parent is pleased at the generous act but notices that the flowers have been pulled up by their roots and are from a forbidden part of the garden. Should the parent reinforce the act or punish it? If the parent reinforces the behavior, will the child become more likely to act generously or will the child become more likely to pick flowers indiscriminately? If the parent punishes the act, will the child learn to become more careful in the garden or simply to be less thoughtful and generous? A teacher frequently faces a similar dilemma. A normally quiet child speaks out of turn in class and says something only tangentially related to the material. If the teacher reinforces that action, will the child become more likely to engage in appropriate classroom discussions or will the child just become more likely to make inappropriate and disruptive comments?

The practical problem is to predict what actions will actually

change in frequency as a result of reinforcement or punishment. One approach – often favored by cognitively oriented theorists – is to assume that reinforcement strengthens the response as specified "from the organism's perspective." What is critical, then, is the *intended* or the *remembered* response. If, for example, the child with the flowers was intending to offer (or remembered offering) a thoughtful gift, and only incidentally picked in forbidden ways and places, then reinforcement should strengthen acting generously.

The problem with this approach for the practical person is that it is not always clear how one is to identify the organism-referenced response (i.e., the intent or the contents of working memory) in advance of performing the reinforcement operation and observing the effects on behavior. Parents or teachers need to decide beforehand whether or not to reinforce a behavior, and to do that, they must be able to anticipate the effects of a particular episode of reinforcement or punishment. It would not help much to be told that intentions or the contents of working memory are to be inferred after the fact from changes in behavior.

In the target article, Killeen deals with a simpler and more restricted range of phenomena than is suggested by the everyday examples described above. For example, verbal influences are not involved. Yet his basic approach lies partly within the cognitive tradition, because reinforcement operates to strengthen the response as represented in the contents of working memory. Killeen's model breaks from that tradition, however, in ways that should appeal to persons with practical interests. Most of the terms in his model represent events whose values can be specified in advance, thus supporting the development of a priori prediction. The major exceptions are the constants (i.e., the free parameters) that must be extracted *post hoc* from the data. It will be interesting, therefore, to see whether the values of these constants remain stable across situations or at least change in consistent ways as a function of certain measurable variables. If they do not, the genuinely predictive power of the model will be somewhat limited.

Clearly, the model represents an initial, introductory step. Further work may extend the approach to cover a wider range of phenomena. Whether or not such extensions prove successful, however, will probably depend on the extent to which there are, in fact, *fundamental* constraints on the size and content of operant units. The exponential decay of memory elements is treated as one such constraint. Might questions be raised, however, about the generality of the constraints? What about the fact that the size and content of the "memory element" (i.e., the "chunk") is highly variable, depending on experience and circumstances? And how might the idea of fundamental constraints be reconciled with the persistent view that the operant is a very flexible unit whose size and content vary greatly, depending on selection by reinforcement contingencies and other environmental conditions (e.g., Baer 1981; Skinner 1957)?

## The scale of nature: Fitted parameters and dimensional correctness

D. W. Stephens

School of Biological Sciences, University of Nebraska at Lincoln, Lincoln,  
NE 68588  
Electronic mail: dws@forager.unl.edu

Behavioral ecologists look to animal psychologists for insight into the mechanisms that govern the ecologically relevant behavior we study. At least for this behavioral ecologist, Killeen's target article succeeds at this level, because it challenges my ideas about how reinforcement works. It has a key property of the best models in science; after reading it, you wonder how you could have failed to consider it before. As a behavioral ecol-

ogist I wonder whether Killeen's ideas could be applied when animals have meaningful alternatives to key-pecking available. However, my strongest reactions stem from my background as a modeler rather than my background as a behavioral ecologist.

When a model "explains" data in terms of a fit to a predicted curve, one wonders whether the parameters are fitted to the data, or known *a priori*? This is a central question, and not an issue to be relegated to fragmentary notes. I gather, primarily from Killeen's Note 4, that he estimated all three parameters of the FR (fixed-ratio) model (and presumably the other models) from the response-rate data. This diminishes the power of his interpretation. Killeen argues that different coupling functions (the  $\zeta$ 's) explain the differences between different schedules of reinforcement. The curves in Figure 1A are fits of both Killeen's FR and VR (variable-ratio) model to Powell's (1968) FR data. Both models fit the data well, so we do not know whether the schedule differences that Killeen attributes to differences in "coupling" are really coupling differences or parameter differences. Note also that for the same data I find a fitted "specific activation" that is twice as large for the FR model as for the VR model. The fitted parameters have *numerically* different meanings in the different models.

There are many dimensional errors in the target article, that is, Killeen writes equations such as:

$$1 \text{ second} = 1 \text{ hour}$$

It just is not so. Recall the elements of dimensional correctness: (1) the units on both sides of an equality must agree, (2) the units of added (or subtracted) quantities must agree (the apples and oranges principle), and (3) the arguments of transcendental functions (i.e.,  $e^x$ ,  $\sin x$ , etc.) must be dimensionless (so-called pure numbers with no units). Rule (3) gives Killeen the most trouble; he writes expressions (e.g., Killeen's Equation 18) such as:

$$e^{-1 \text{ sec}}$$

To see the nonsense here, recall the series identity for  $e^x$ :

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \dots$$

A dimensionless  $x$  is required, because otherwise each element of the sum would have different units and we would be adding apples and oranges.

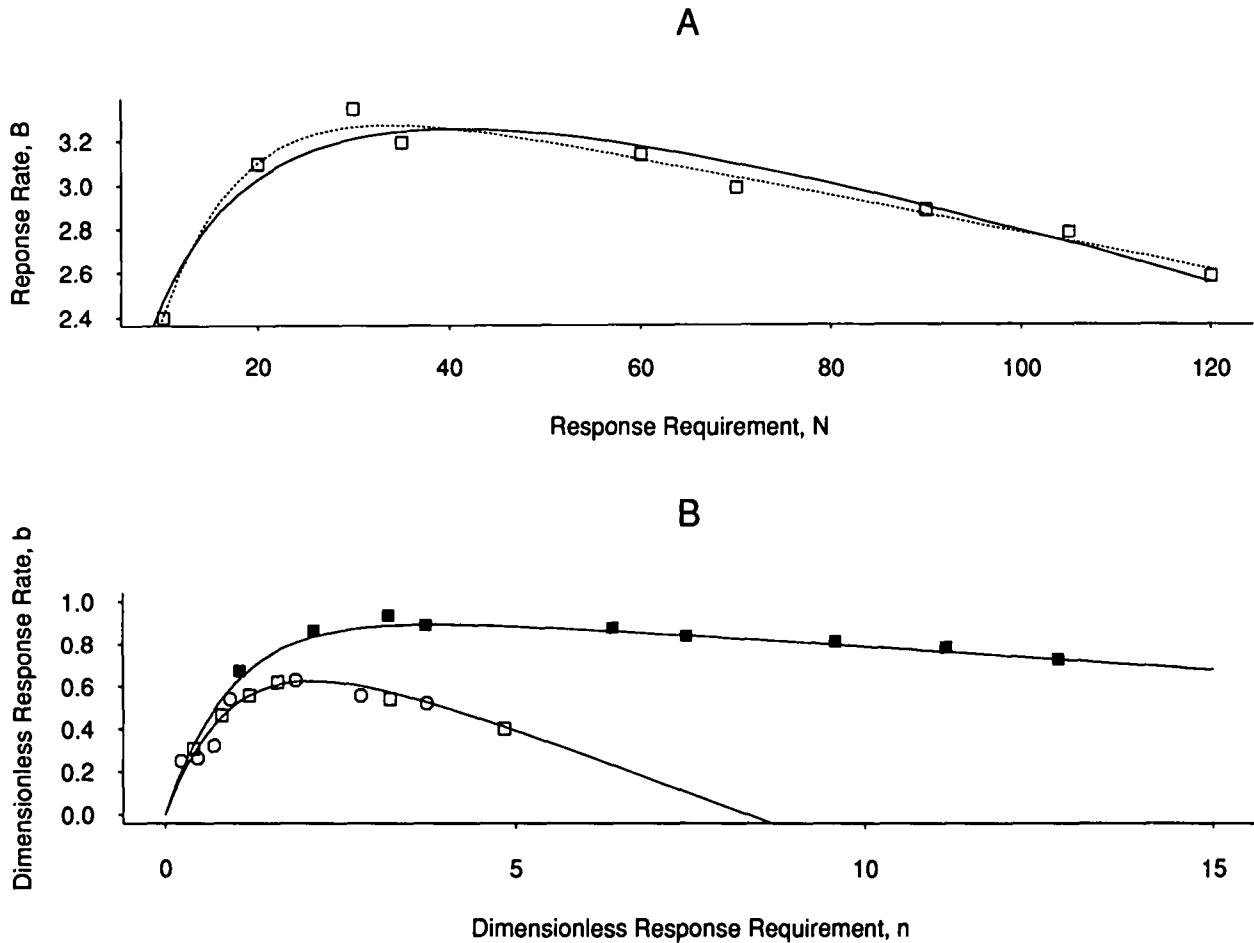


Figure 1 (Stephens). *Panel A* shows Powell's (1968) FR data as open squares (extracted by eye from Killeen's Fig. 7), and two fitted curves. The solid curve is a fit to Killeen's VR model and the dashed curve is a fit to Killeen's FR model. I used the nonlinear regression routines in the S-plus statistical system to fit both models. Estimated parameters were  $a = 120$ ,  $\lambda = 0.116$ , and  $\delta = 0.275$  for the FR model; and  $a = 71.2$ ,  $\lambda = 0.137$ , and  $\delta = 0.221$  for the VR model. The standard error of the residuals was 0.09 for the VR model and 0.06 for the FR model.

*Panel B* shows a dimensionless, one-parameter version of Killeen's FR model (Equation 3). All the data were extracted by eye from Killeen's Figure 7. The solid squares show Powell's data, the open squares show Barofsky and Hurwitz's (1968) rats 1 and 3, and the open circles show Barofsky and Hurwitz's rats 4 and 5. The curves take their parameters from Killeen's Table 2. The upper curve plots Equation 3 for  $\alpha = 0.02$  (from the Powell fit), and the lower curve uses  $\alpha = 0.12$  (from the Barofsky & Hurwitz fits).

Consider Killeen's Equation 1:

$$\beta = \lambda e^{-\lambda d}$$

We are told  $d$  is a duration with units of seconds. Rule (3), therefore, requires that  $\lambda$  must have units of inverse seconds, and this in turn tells us that  $\beta$  must have units of inverse seconds. However, in the next paragraph Killeen treats  $\beta$  as a pure number when he writes:

$$\lambda = -\ln(1 - \beta) \quad (1)$$

(cf. Rules [2] and [3]). At least one of these statements is false.

In equations 7 and 8, Killeen is calculating a coupling coefficient for an FR schedule, when he passes from a discrete to a continuous case. In the discrete case, the weight of remembered responses of an FR with requirement  $N$  is given by

$$1 - (1 - \beta)^{N-1} \quad (2)$$

and Killeen tells us, this is approximated by

$$1 - e^{-\lambda N} \text{ given } \lambda = -\log(1 - \beta)$$

The problem here is that the  $N$ 's in these equations are not the same. When we move from the discrete to the continuous response case, we imagine that there exists some step size, say  $h$ , relating the discrete quantity to the continuous quantity. Let  $N_d$  and  $N_c$  be the discrete and continuous response requirements, respectively; evidently  $N_c = hN_d$ . We see that  $h$  must have the same units as the continuous response requirement because it divides the  $N_c$  axis into chunks of length  $h$ . This implies that  $N_d$  is a pure number; because  $N_d = \frac{N_c}{h}$ .

Next we assume that  $\beta \approx \lambda h$  as  $h$  becomes small. Dimensional consistency now requires that  $\lambda$  have dimensions that are the inverse of  $N_c$ 's. If we substitute this and our definition of  $N_d$  into expression (2) and rearrange it, we have

$$\lim_{h \rightarrow 0} \left( 1 - \frac{(1 - \lambda h)^{\frac{N_c}{h}}}{1 - \lambda h} \right) = 1 - e^{-\lambda N_c}$$

Now, we see that the dimensionally incorrect expression  $\lambda \approx -\log(1 - \beta)$  arises because Killeen has improperly dropped the dimension bearing stepsize  $h$ ; the correct expression is

$$\lambda \approx -\frac{\log(1 - \beta)}{h}$$

Note that Killeen may suppose that  $h$  equals  $1 \frac{\text{response}}{\text{reinforcement}}$ , but it is no more correct to drop the dimensioned term in this case than it is to say that 1 minute equals 1 second.

These dimensional errors reduce Killeen's mathematical credibility, and they further weaken his interpretations of the parameters. However, there is little numerical effect when the parameters are fitted, because dimension-correcting factors can be absorbed into the fitted terms. The real tragedy of this dimensional sloppiness is that Killeen misses an opportunity to understand more deeply the "hidden structure" of his own ideas. Consider his model for FRs:

$$B = \frac{1 - e^{-\lambda N}}{\delta} - \frac{N}{a}$$

where Table 1 gives the terms and their units. Buckingham's  $\pi$  theorem tells us that all mathematically correct relationships can be written in a "maximal set" of dimensionless terms (often called the  $\pi$ 's; see Huntley 1958; Isaacson & Isaacson 1975; Logan 1987; Stephens & Dunbar 1993).

When viewed through the  $\pi$  theorem's lens, the FR model is not a three-parameter curve relating the two variables  $B$  and  $N$

Table 1 (Stephens). Terms and dimensions for Killeen's FR model

Name	Symbol	Units
Response rate	$B$	responses second
Schedule requirement	$N$	responses reinforcer
Memory-decay rate	$\lambda'$	reinforcers response
Minimum interresponse time	$\delta$	seconds response
Specific activation	$a$	seconds reinforcer

<sup>1</sup>Killeen gives the relation  $\lambda = \frac{\lambda' \delta}{\rho}$  and he gives units of 1/seconds for  $\lambda'$ , and seconds/response for  $\delta$ ;  $\rho$  must be a pure number if Equation 9 is to be dimensionally correct, so this implies units of 1/response for  $\lambda$ . This choice of units violates rule 3, so I've changed it as indicated.

but a one-parameter curve relating similar rescaled quantities. Let  $b = \delta B$ , so that  $b$  measures response rate as a proportion of maximal response rate; let  $n = \lambda N$ , so  $n$  is a dimensionless measure of response memory; and let  $\alpha = \frac{\delta}{a \lambda}$  a dimensionless parameter that seems to measure the relative strength of memory and response ceilings. Killeen's FR model becomes the one parameter curve:

$$b = 1 - e^{-n} - \alpha n \quad (3)$$

The magnitudes of  $b$ ,  $n$ , and  $\alpha$  have absolute meanings that are comparable from one study to the next. This allows meaningful comparisons of different data sets (within the context of the model, Fig. 1B). For example, it eliminates the difference between the two groups of rats in Barofsky and Hurwitz's (1968) data, and it lets us say that pigeons respond at a higher relative rate than rats do, apparently because pigeons strike a different balance between memory and response constraint ( $\alpha$  differences). Moreover, dimensionless models are powerful tools for approximation; in this case they show that the linear approximation  $1 - \alpha n$  (cf. Killeen's Equation 6) is accurate to within 0.007 for all  $n$  greater than 5. I am not suggesting that dimensionless models replace their dimensioned counterparts, only that they provide an important tool for revealing the "minimal" mathematical structure of models and for making sensible comparisons and approximations.

## Short-term memory in human operant conditioning

Frode Svartdal

Department of Psychology, University of Tromsø, 9037 Tromsø, Norway  
Electronic mail: frodes@psyk.uit.no

Killeen assumes that reinforcement affects the animal's representation of its behavior, not the behavior itself. This representation is a weighted memory of recent responses. But can the operation of such a memory mechanism be specified independently of functional accounts of response-reinforcer relations? In other words, if only remembered responses are reinforceable, is the criterion for remembered responses then that they

are reinforceable? It may be unnecessary to answer this question on the nonhuman level, but Killeen's model is not restricted to nonhuman data (e.g., Table 2). In human conditioning an added source of information on schedule representation is available, namely, subjects' verbalized representations of their behavior and of contingencies. Since standard models of human cognition would assume representations in short-term memory to be accessible for verbal description, a requirement that only remembered responses be reinforceable may then imply that proper verbalization of contingencies is required for human contingency sensitivity.

Such a requirement is troublesome, however, for at least two reasons. First, verbalized representations of contingencies are strong determinants of human operant performance, whether they represent (e.g., Shanks & Dickinson 1991; Wasserman et al. 1993) or misrepresent (e.g., Catania et al. 1982; Hayes et al. 1986; Lowe 1979) environmental contingencies. The latter possibility presents a complication for the present model in that accounts of environment-behavior relations are then influenced by additional controlling variables that have proved to be very difficult to identify (e.g., Hayes & Hayes 1992; Lowe 1979). Second, findings that environmental contingencies can affect human behavior independently of any reasonable form of conscious representation (e.g., Hayes & Broadbent 1988; Jacoby et al. 1992; Svartdal 1991, 1992) indicate that contingencies need not be represented in memory in the form assumed by "standard" models of human cognition to be effective.

The distinction between contingency-shaped and rule-governed behavior (e.g., Hineline & Wanchisen 1989; Skinner 1969) attempts to encompass some of the problems presented by this diverse situation. Whatever its merits, environmental contingencies are, according to this distinction, assumed to affect human behavior in two ways: through verbalized representations of contingencies and through mechanisms that do not involve such verbalized representations. A "standard" view of cognition and memory, adopted by Killeen, one postulating a limited-capacity short-term memory associated with verbalized representation of contingency information, can account for cases where contingencies are successfully mediated by verbalized representations of response-outcome relations (e.g., Shanks & Dickinson 1991). When environment-behavior relations are overridden by verbal control mechanisms (e.g., Hayes et al. 1986) the model may halt, and when mediational representations of contingencies cannot be identified, postulation of supplementary processing mechanisms may be required (e.g., Eriksen et al. 1990; Hayes & Broadbent 1988; Navon & Gopher 1979; Shanks & Dickinson 1991, p. 359). It remains to be seen whether Killeen's model can deal with these complications as successfully as it deals with operant performance in nonhuman animals. If it can, our understanding of human operant performance will certainly progress.

## Animal-centered models of reinforcement

William Timberlake

*Psychology Department, Indiana University, Bloomington, IN 47405  
Electronic mail: timberla@ucs.indiana.edu*

Killeen's use of a set of equations to describe and predict reinforced behavior follows a long line of efforts to develop general laws of behavior patterned on the equations of Newtonian physics. A difference between Killeen's approach and previous work is that he has not focused as sharply on isolating learning from performance as Hull (1943) or Rescorla and Wagner (1972) and he is more interested in analyzing separable mechanisms involved in learned behavior than are the adherents of the matching law or economic optimization approaches (Herrnstein 1970; Staddon 1980).

Killeen combines basic operant concerns, an incentive arousal theory, tools and concepts from twentieth-century physics and mathematics, and a spot of cognitive psychology to produce a model and a set of equations accounting for reinforcement effects. As with any model, one wonders whether the concepts are well defined and compatible with their previous use. One also wonders whether the model is unique, that is, whether the same data could be predicted from different equations, or more subtly, whether the same equations can be produced from different assumptions.

In developing the complexity required to take us beyond the traditional discriminative stimulus-response-reinforcement approach, Killeen posits a variety of concepts that require intuition, scaling assumptions, and iterative fitting techniques to assemble. An advantage of this approach is that it makes contact with a broader range of disciplines, mechanisms, and data than more traditional reinforcement approaches. A disadvantage is that it can be quite complex in its development and application. It is difficult to predict from inspection how this range of elements should fit together or what various combinations of elements can account for. It will require considerable work to discover the potential of these concepts and the best ways of mapping the fit between the animal and its world.

At a fundamental level, I very much liked the animal-centered aspects of Killeen's approach. I think it is quite important to view the world from the animal's perspective – not in an anthropomorphic (human-centered) way, but in a theromorphic (animal-centered) way (Timberlake 1993a). Killeen asks how the animal defines the response (as opposed to how the experimenter defines it). Using a combination of potential models and data-fitting techniques, Killeen estimates the nature of the short-term response memory that is acted on by reinforcement. This short-term memory adds an important response dimension to the classic temporal delay gradient of reinforcement theory.

A concern about Killeen's memory model has to do with how well the retrospective effects of reinforcement are modeled by a single response-based decay function. For example, Timberlake and Lucas (1989) provided evidence that at least two motivational modes underlie appetitive behavior leading to food, a focal search mode and a more general search mode. It might be that response-based decay functions are independent and different for these two modes, or that the decay function for more general search is time based rather than response based. As evidence for such complexity, consider that a rat does not learn a multiple "T" maze from goal box to start box as might be anticipated from Killeen's approach, but from each end toward the middle.

Though Killeen's combination of a short-term response memory and incentive-based arousal brings the focus of reinforcement closer to the organism, in several respects I do not think it goes far enough. That an approach of such complexity falls short may seem counterintuitive, but I think some limitations stem from its ties to the traditional reinforcement approach. First, the notion of a reinforcer remains unexamined in Killeen's theory; it simply exists and has measurable effects. There are considerable data showing that a reinforcer is not a particular stimulus or response, or even a response with a single characteristic, such as high probability or biological significance (Timberlake 1993b). From the animal's standpoint, reinforcers arise when the local rates or patterns of responding constrained by a schedule differ from the amounts or patterns instigated by the stimulus conditions.

Second, Killeen's approach may not go far enough with respect to how learning fits into the larger framework of an animal's daily life (e.g., Gallistel 1990; Timberlake & Lucas 1989). Traditional learning research has highlighted the acquisition of simple environmentally defined manipulation responses, such as lever-pressing or key-pecking, or learning about stimuli that predict immediate food. This learning is undeniably impor-

tant, but it is only a small portion of the role of learning in survival.

For example, considerable research has shown that the time of day of a meal is a critical determinant of appetitive behavior, as are the spatial location of the meal and the behavior required to obtain it. In addition, these types of learning often appear much more rapid than learning a manipulation response or the color or form of stimuli just preceding a food item. Perhaps the most puzzling point is that these ecologically critical forms of learning are actually manipulated by typical experimental procedures but they are rarely studied apart from their largely undefined contribution to learning the predictiveness of local stimuli and responses.

Killeen's model is complex, but its focus is on manipulation response leading immediately to food items and its complexity is largely abstract and relatively neutral with respect to evolution and ecology. It seems apparent that the study of phenomena such as birdsong learning will require careful consideration to reconcile with his model. Even in dealing exclusively with feeding, it is not clear how much more light the study of manipulation responses can cast on how a cheetah learns to hunt, or how birds learn to anticipate insect eclosions.

Many naturally occurring contingencies are highly interactive and appear to require a more animal-centered approach. For example, predator and prey have different goals, yet the success of the behavior of each can depend most subtly on the other's behavior. It may be possible ultimately to deal with the learning of hunting and capturing techniques as a problem of short-term memory, schedule requirements, and arousal – but the complexity of mapping these concepts onto an animal in its niche is easily underestimated. I suspect that attempts to simulate animals using computing and mechanical techniques (Beer 1990), especially animals in social and interactive circumstances, will provide considerable impetus for a more animal-centered approach.

Finally, in supporting a more animal-centered approach to modeling I am not claiming that real-world learning is too complex to analyze in the laboratory or that basic research on simple mechanisms is a waste of time and cannot account for real behavior. What I am arguing (probably with some agreement from Killeen) is that we have not adequately explored the variety of learning mechanisms and their interactions. Though psychologists have been remarkably analytic about associations between stimuli, responses, and reinforcers, they have not been analytic enough about how special qualities such as memory, location, time of day, and particular stimuli, responses, and motivational states produce learning that fits the ecology and evolution of a species. This is an important direction in modeling behavior.

#### ACKNOWLEDGMENT

I would like to acknowledge the support of grants MH 37892 and IBN-9121647, and the comments of Mark Hoyert and Fran Silva.

## How general is a general theory of reinforcement?

Stephen F. Walker

Department of Psychology, Birkbeck College, London WC1E 7HX, England  
Electronic mail: [ubjta25@cu.bbk.ac.uk](mailto:ubjta25@cu.bbk.ac.uk)

Both the title of the target article and its contents invite comparison with the mathematically formulated principles put forward by Hull (1943; 1952), which were intended to apply to all behaviour but which were stated precisely enough for their lack of generality to be eventually demonstrated. Killeen's first paragraph refers to the "wide-ranging implications" of the principles presented and section 5 is titled "A general theory of reinforcement." There is a proviso in section 9.3.1 that the

target article "addresses only distance along the dimension of homogeneous operant responses," but it is implied that this is an example that will be capable of extension. One of the strengths of the theory presented is that it contains parameters that are good candidates for explanations of differences between response categories and between species. I shall therefore comment first on questions of the relevance of the theory to data outside its base of homogeneous operant responding and second on whether the theory is sufficiently powerful even within this base.

The most immediate difficulty for Hullian response-reinforcement theory was dealing with spatially directed behaviour. A very simple example described by Hull (1934) was that of a rat trained to run a fixed distance past a closed door and then to come back to it, the door then being open to allow the animal to proceed to a food reinforcement. Hull noted that under these conditions highly trained rats would waste time trying to scratch through the closed door as they passed it, and some would go through the door on the first pass if it was left open. This and many other examples provided by Tolman (1932; 1948) could not be directly predicted from Hull's first principles, which were entirely frequency sensitive and response based and he developed elaborate ramifications of his theory ("habit-family hierarchies") to account for them. A much more direct option appears to be open to Killeen, because he suggests that processes of incentive coupling may occur on any dimension of "the organism's psychological space" (sect. 9.3.1). If representations of geographical space are major components of most species' psychological space (Gallistel 1990) then the coupling of incentives to locations rather than only to the responses currently required to reach them would seem to be an important area of a general theory of animal learning. An enormous amount of experimental evidence, for instance from the radial maze (Olton 1979) and memory for hoarded food (Shettleworth & Krebs 1982), is available for the testing of principles in this area and yet Killeen's discussion of foraging (sect. 9.3.4) refers only to memory indexed by responses. Good evidence is presented that response indexing occurs under operant schedules of reinforcement, whose contingencies require it, but other circumstances (including many examples of Pavlovian conditioning and operant discrimination learning) may induce coupling of incentives to "the stimulus as coded" or "the location as coded."

A second area of difficulty for Hullian theory was the rapidity of behaviour change after reinforcement manipulations (latent learning and reward devaluation), and this resulted in the separation of habit formation from incentive learning. Killeen appears to have brought them back together again, which may be constructive, but how in this case can one account for evidence that distinguishes between stimulus-response associations and response-reinforcer associations? Can the treatment of motivation by means of the activation function (which bears an informal resemblance to "incentive motivation": Hull 1952) account for recent evidence on reward-devaluation effects (e.g., Dickinson 1985; Rescorla 1990b)?

The focus of the target article is on schedule effects in conventional operant conditioning, and it particularly provides an alternative to theories that propose direct effects of reinforcement on IRTs. An acknowledged area of vagueness is in the structuring of "response units" (sect. 5.4, last para.) which have to be inferred in some cases but are especially obvious where there is sensitivity to response number, either in FR schedules or in explicit "counting" schedules (Mechner 1958; Davis & Pérusse 1988). In these cases it is arguable that "number" or "run-length" is directly reinforced. Limitations on generality within the domain of repeated operant responses occur insofar as there is no treatment of negative reinforcement, but one would expect that straightforward modifications to the mathematics (or even redefining what constitutes a positive incentive) would allow an interesting extension of the present principles. The issue of IRT reinforcement in free-operant avoidance learn-

ing was raised by Sidman (1954), but later studies (Sidman 1962; Herrnstein 1969) suggest that explanations that do not require reinforcement of specific IRTs should be preferred. There are, however, likely to be difficulties in identifying the terminal response and the point of reinforcement: "contiguity" in free-operant avoidance, especially in probabilistic versions (Herrnstein & Hineline 1966), is by definition less visible than in schedules of positive reinforcement.

An apparent gap in the theory, which applies to punishment of operant responses (Boe & Church 1967) as well as to aspects of extinction, is that there is no separate mechanism of response inhibition. In some cases the weighted moving average might be expected to suffice, but there is compelling evidence from many areas of animal learning, including behavioural contrast effects in multiple schedules, that a theory which includes only excitation and its absence is only half a theory.

Clearly the target article is intended to be narrow but thorough, rather than all-embracing. But part of its appeal is its potential generality, and this can only be realized by testing the theory against a wider range of phenomena than has so far been attempted.

## Fifty years on: The new "principles of behavior"?

J. H. Wearden

*Department of Psychology, Manchester University, Manchester M13 9PL, United Kingdom*

**Electronic mail:** wearden@psy.man.ac.uk

It would be a stern reader indeed who found little of merit in this target article, and most commentators, like Mark Anthony, may come to bury but will surely stay to praise. And there is much to praise: Killeen's theoretical grasp extends to an unprecedented range of schedule effects; he throws out new ideas, like the coupling of memory contents and reinforcement, and gives us new perspectives on many older ones, such as the reinforcement of interresponse times (IRTs) and correlations between behavior and reinforcement. As an intellectual achievement, few single articles on schedule theory are its peers, but is it right?

Goodness of fit between equations and data (as in Figs. 7 to 12) looks impressive, although the complex functions sometimes used would probably account for a respectable proportion of the variance in a relief map of the Himalayas. A more critical test is the behavior of the parameters; do they change in a psychologically reasonable way with experimental manipulations? In general, they do: increasing lever weight lengthens theoretical response duration; each reinforcer stimulates more behavior in naturally obese, but not lesioned, rats; and degree of food deprivation mostly affects the incentive parameter, as it should (see Tables 2 and 3). Just occasionally, there is a hint of derailment: for example, was it really sensible to fit the same equation to the very different behaviors of different human subjects in McDowell and Wixted's (1986) study, then have to puzzle over the wild differences in parameters? Killeen boldly tries to integrate data from humans into his models but might have been more prudent to hold fire until some mysteries about what "reinforcement" does in experiments with humans are explored further (see Svartdal [1993] for just one).

Given that the equations fit and "behave" well, do they account for the important phenomena of schedules? Successes are certainly commoner than failures, although the focus of interest is usually on response rate rather than the patterning of responses over time. This means that the pratio or postreinforcement pause on fixed-ratio schedules, perhaps the single most striking feature of the behavior that this schedule generates, remains mysterious. To be fair, no other theory accounts con-

vincingly (or at all) for this perplexing phenomenon, but can Killeen's treatment deal with such pauses *in principle*, or is some supplementary process needed? Since the pause does not move the animal closer to reinforcement and does not contribute in any way to the fulfillment of the response requirement, the behaviors occurring in it are unlikely to be present in memory at the time of reinforcement, yet pausing persists so strikingly. Likewise, the analyses of possible within-interval response patterns on fixed-interval schedules (sect. 14.1.4) and the effects of reinforcing IRTs (sect. 13.4) are comparatively cursory, leaving the impression that the model might not do as good a job here as when accounting for response rates when animals are responding more or less continuously.

Having expended innumerable megaflops trying to obtain reasonable simulations of behavior using the reinforcement of IRTs (with Wearden & Clark [1988; 1989] just the tip of an unpublished iceberg), I found Killeen's treatment of the difficult problems of IRT reinforcement particularly insightful. In 1988 and 1989 Clark and I wrestled with what we called "pattern memorizers," where a reinforcer caused the storage in memory not just of the single IRT that preceded it but of the preceding "pattern" of IRTs (either limited by number of IRTs or the total time they occupied). These ideas strikingly parallel the way that Killeen's article treats short-term memory, except for two important differences; first, we did not weight the preceding IRTs in the pattern (so the model essentially used the reinforcement process Killeen shows in panel C of his Fig. 1) and second, we could never get the pattern memorizers, despite their computational complexity, to produce convincing simulations of behavior. One moral here might be that attempting to reduce Killeen's local correlations to something more "molecular" may be futile; another is that while having ideas might be difficult, getting them to work properly is even more difficult!

Killeen's treatment of IRTs as a sequence of unmeasured responses terminating in a measured one helps us understand the way in which interval and ratio schedules do or do not reinforce IRTs perhaps more clearly than before, and this basic idea is impressively developed. Data from the 1970s (Annable & Wearden 1979; Shettleworth 1975) suggest, however, that some usually unmeasured behaviors are more difficult to strengthen with contiguous food reinforcement than others. The presence of these behaviors in short-term memory at the time of reinforcement would have different implications for the change of the IRT of which they are part than the presence of others, more readily reinforceable by food, so knowing which unmeasured behaviors make up an IRT may be important. Although this may limit Killeen's mathematical treatment, it may be that behavioral measures are too imprecise for such details ever to make a significant difference to the precision of prediction.

Fifty years ago, Hull (1943) offered us his proposed *Principles of behavior*, and more recently (Wearden 1989) I suggested that contemporary researchers in reinforcement schedule theory are Hull's true intellectual heirs. The current target article strongly reinforces this belief, magnificently illustrating that the search for scientific understanding of behavioral phenomena is a quest still well worth pursuing.

## A mathematical theory of reinforcement: An unexpected place to find support for analogical memory coding

Donald M. Wilkie and Lisa M. Saksida

*Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4*

**Electronic mail:** dwilkie@cortex.psych.ubc.ca; saksida@unixg.ubc.ca

In arguing that "we reinforce the animal's representation of its behavior, not the behavior itself," Killeen has both challenged

the traditional behavioristic view of the reward process and expanded the domain of animal memory research. Several of us who study animal memory have tended to view reinforcement as a very useful tool, but a topic not too relevant to our interests. After Killeen's target article, such a stance will no longer be possible. Already we see interesting relations between these two fields. One, the focus of this commentary, concerns the nature of the memory code for the response.

In the animal memory literature several different types of memory codes have been proposed and argued for. These can be illustrated by reference to the literature on memory for event duration. Animals can retain the duration of an event such as a tone in several possible ways. First, the temporal information may be encoded into a categorical representation on a nontemporal dimension. Kraemer et al. (1985) discussed two ways in which temporal information might be categorically encoded. One involves the formation of an anticipatory code of the correct response or correct stimulus for an event of the specified duration (e.g., after a short event, "choose red" is encoded). This type of code is often called "prospective" (see Honig 1978). A second involves the transformation of the temporal information into a nontemporal dimension (such as "short" or "long"). In both cases, the animal retains the categorical tag but not the specific temporal aspects of the duration itself. Although there is scant empirical evidence for the latter type of coding, there is evidence that animals remember temporal information in a prospective manner under some conditions (see e.g., Grant & Spetch 1991).

Temporal information may also be remembered in a veridical, analogical fashion. One way this could be achieved would be by keeping track of other events or regular physiological processes occurring during the interval being timed. Duration would then be represented as a series of events that are in some way accumulated. Again, under some conditions, there is empirical evidence for this type of coding (for examples of such evidence see Grant & Spetch 1991; Spetch & Sinha 1989; Wilkie & Willson 1990).

Although Killeen does not explicitly discuss memory codes, it seems to us that responses would have to be represented in a veridical/analogical fashion for his model to work. It is difficult to imagine how another type of coding of the response would lead to the increased probability of a stereotyped set of behaviors with specific topographical and temporal dimensions. When animals remember just temporal information there seems to be some flexibility in the nature of the code used. However, when animals remember responses on which reinforcement works, it seems likely that the code will be more specifically analogical.

In sum, Killeen's target article is very useful to animal memory researchers, because it provides us with a concrete model of reinforcement that is not only integrated with, but contributes additional evidence to, at least one area of memory research.

#### ACKNOWLEDGMENT

Preparation of this commentary was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

### The return of the reinforcement theorists

C. D. L. Wynne

*Department of Psychology, University of Western Australia, Nedlands, Perth, Australia WA 6009  
Electronic mail: clive@psy.uwa.edu.au*

To those outside the field it may seem that reinforcement is a fairly simple thing. You give a subject something good for doing something, and it does it more; Give a subject something unpleasant and it will stop what it was doing. Thorndike (1898) already had got that far. The problem is, in nearly one hundred

years, we have not come much further than Thorndike's functional relation. The problem that Killeen sets out to solve in his target article – how does an individual know what it is being rewarded for (the "assignment of credit" problem) – has remained largely untouched until rediscovered in artificial intelligence by Minsky (1961).

The experimental analysis of behavior, as behaviorists like to call their science, has been preoccupied with control at the expense of understanding. Killeen's target article demonstrates that the two are inseparable. He offers an explanatory framework that also gives a prescription for faster conditioning: if our definition of the response matches the organism's definition then our control of the subject's behavior will be much better than otherwise. But Killeen also shows the potential utility of a deeper level of understanding. In section 5, he reanalyzes the effects of lesions of the hypothalamus in rats and is able to show that the lesioned animals' memory remained intact but that they were motor impaired – a result that was not apparent without the kind of analysis made possible by his theory.

What should a general theory of reinforcement look like? What do we want it to account for? In the target article, Killeen deduces an axiom from general considerations (that reinforcement acts on whatever is in short-term memory) which he then tests in an experiment. He goes on to show the applicability of this principle in a variety of familiar experimental paradigms. This, in principle, has to be the way to go, but some of the details leave me a little uneasy.

The idea that reinforcement acts on the contents of short-term memory is an intuitively reasonable one, but Killeen's experimental test of it is rather indirect. If we are to accept the idea of a short-term memory for responses, we need to know more about it. It ought to be possible to manipulate it in some way and observe changes in the conditionability of the subject. Perhaps pharmacological methods might be useful here.

Killeen shows that interresponse times are most rapidly shaped when his percentile reinforcement schedule best couples with the animal's memory for what is being reinforced. This choice of interresponse times is completely arbitrary for his theory. If I understand it correctly, the same process should operate on any dimension of the response. Thus, this account would gain weight if it could be demonstrated that the same process operates in the same way on other dimensions of responding, possibly the location of response, the force of responding, and so on. A general theory should be grounded in observations that are (at least potentially) omnipresent – not only applies fall to the ground.

Having deduced his model, Killeen commits the bulk of the target article to deducing behavior on the standard schedules of reinforcement. Here there is great ingenuity: I especially appreciate the explanation of why response rates are usually higher on ratio than interval schedules (interval schedules reinforce sequences of anything that ends in a target response; ratio schedules preferentially strengthen sequences of target responses that end in a target response). Probably these are just the limitations of space, but I wanted to see how the model would account for the temporal structure of behavior: Why do animals pause after reinforcement on fixed-ratio and fixed-interval schedules, for example, but not on variable-ratio and variable-interval schedules? The theory's explanations for other dynamic effects such as the partial reinforcement extinction effect would interest me too. Can the model be generalized to Pavlovian contingencies also?

Finally, the real proof of the utility of a novel theory is in the novel predictions and tests it generates. Here Killeen is strangely silent: Does this theoretical model not lead to anything new? I hope these queries prompt him to develop his theory further.

# Author's Response

## Rats, responses and reinforcers: Using a little psychology on our subjects

Peter R. Killeen

Department of Psychology, Arizona State University, Tempe, AZ

85287-1104

Electronic mail: killeen@asu.edu

### R1. History

This endeavor is nicely placed into its historical context by Bolles and Hanson. As Bolles notes, in the search for objectivity, behaviorists since Thorndike have attempted to get all the determinants of behavior out into the environment. Whereas this increased our ability to communicate with our colleagues, it decreased our ability to communicate with our subjects, as nonnegotiable operational definitions led to unilateral experimental "discussions." Little wonder that we sometimes misunderstood our subjects' responses!

Like me, Bolles thinks that the subjects should shoulder some of the responsibility for the definition of the response unit, and that stimuli and incentives should also be put "where we can do some psychology with them." Behaviorists are flexible about the effective reinforcer, as epitomized in Lattal & Abreu-Rodrigues's case for "functional analysis," but usually less so for the stimulus, and least of all for the response: "Behavior theory has with few exceptions persisted in treating its R datum as a punctate action of the organism, as an event-thing. The habit comes from our folkways of thinking about behavior, and from our mother tongue's vocabulary and descriptive modes, which are ingrained in us and suffuse our scientific education" (Schoenfeld 1972). Curious, isn't it, that to many for whom an indeterminate "history of reinforcement" is a shibboleth, a weighted history of responding is anathema.

Hanson deftly uses the linear-operator model as a thread to tie these principles into the quilt of modern learning theory. Equation 15 gives rise to an exponential-integral learning curve that is at the core of Hull's (1943) learning theory, Estes's (Neimark & Estes 1967) stimulus sampling theory, and Rescorla and Wagner's (1972) model of discrimination learning – to which Wagner (1981) later added memory. As Hanson notes, Rescorla and Wagner generalized the linear operator model in an important way: they assumed that the total associative strength of an unconditioned stimulus (US) was fixed, and that allocation to one conditioned stimulus (CS) would reduce the amount available for allocation to another. Such competition also occurs in two places in the present theory: (1) the attention given a new item in memory is automatically subtracted from that given to older items by Equations 2 and 14, keeping total "memorial strength" constant, and (2) the arousal available from an incentive can instigate only a fixed number ( $a$ ) of response-seconds (Equation B1; this linear conservation of psychological energy is lost when incipient responses compete and response rates curve under their ceiling). Both constancies – attentional

and motivational – play a role in discrimination learning and in response learning.

Hanson then shows how the next step to more than one memory buffer – the nodes in neural networks – lands us in the middle of the cognitivist revolution (see, especially, Healy et al. 1992). Writing Equation 2 as a vector carries us from representation on one dimension to parallel distributed representations. Given enough nodes, a rule for competition/lateral inhibition, and proper contingencies of reinforcement, we can train such networks to behave in ways an order of magnitude more interesting than is possible with unidimensional representations. Davis and others recognize that we are continually beset by a heterogeneous array of internal and external stimuli. It is only with the more general vector model that we are likely to capture this complexity, as Shimp implies and as Wagner and Brandon (1989) have assayed. I share Hanson's hope that my behavioral argument for memory, in this simplest and most strongly defensible case, will reinforce the "land-bridge" over the sea that divides behaviorists from cognitivists. Other behaviorists (e.g., White et al. 1989) have also stepped, if somewhat delicately, onto the isthmus. Whereas intellectual isolation, like geographical, can foster speciation, intellectual intercourse, like physical, can recombine the best of both worlds.

### R2. Theory

**R2.1. Ideology.** I am correctly identified by Davis as a radical behaviorist, but one may practice the art without subscribing to an *ism*, just as one may be a dentist without subscribing to "dentism" or a flutist without professing "flutism." Behaviorists put behavior first in their priorities, just as this journal does in its title. It is not taken as an indicator of mind, nor an extension of a phenotype, nor a vehicle for mathematics. To understand it completely, we benefit from considering covert processes that underlie it, ecological constraints that shape it, and mathematical formulations that describe it – for behavior is a process, and mathematics is its natural language.

Among Skinner's important roles was Hector against ad hoc explanatory constructs. *Mind* cannot explain behavior until it itself is understood, any more than *God* can explain the problem of good and evil. Indeed, in both cases reliance on salvation by the first can undermine our efforts to understand the second. Cognitivism is no less blinkering a commitment than behaviorism.

Davis worries that because the central construct in my work has been viewed historically as a cognitive process, "philosophical legitimacy is not based on logic but rather on technology." He considers this a "most cynical reading," but I think it most sagacious. Many of the enduring problems of philosophy endure because they are ill-posed; once some technical leverage is gained on them, they are reformulated as a science and moved into another building.

**R2.2. Economy.** I have urged behaviorists to be prudent with their resources, but not parsimonious. The conceptual stinginess of parsimony impoverishes exploration and stifles progress. It puts all the emphasis on cost, rather than cost-effectiveness. Constructs are

like parameters in a regression equation; their cost is the degrees of freedom they absorb from the data. But if the data are sufficiently strong and numerous, such constructs can greatly improve our ability to comprehend them. Gratuitous and *ad hoc* explanations are to be rejected, whether they be mentalistic or behavioral, simply because they don't pay their way. (Killeen 1987, p. 36)

The present theory is my attempt to add a construct to behavior analytics that demonstrably pays its way.

Several behaviorists, however, are not convinced. **Donahoe** notes that just because "longer-term regularities . . . are captured by the notion of 'memory' does not force the postulation of 'memory.'" Correct. Any more than long-term regularities of the tides force the postulation of gravity, but they are not a bad rationale for such constructs; better than that for many others found in the canon. Donahoe believes that my invocation of memory will detract from the search for other causes of behavior, such as "the fact that at small ratios the response preceding the reinforcer . . . may be partially guided by the discriminative effects of the [previous] reinforcing stimulus." It may be, or by many other things; perhaps the tides are partially guided by the winds. **Krank** holds that "the episodic approach can, in principle, account for free-operant responding." In principle, so can my mother. Predictions, please! For starters, what functions do these hypotheses draw through the data in Figures 5–7?

**Pear** suggests that the postreinforcement suppression of rates might be due to momentary satiation. Possibly. If so, we should not see the bitonic function when a time-out follows each reinforcement. Any bets?

Every successful theory changes the way we view the world and therefore makes some relationships more obvious and others, inevitably, less obvious. But securing one good predictor in our regression equation should not discourage the search for others. Nor should it discourage other scientists by reducing the error variance they have left to explain, until that residual becomes sufficiently small that they are in fact better off addressing other questions.

**Lattal & Abreu-Rodrigues** believe my "rather vague memory construct" to be an unnecessary embellishment on the existing "functional analysis" of behavior: the "behavioral effect [of contingencies] may be different from that expected. . . . such observations . . . have been formulated clearly without invoking events taking place at other levels of discourse." Formulated, perhaps, but such handicapping has kept them from being solved, and may even be the cause of the false expectations. Why prefer categories when continua are available? Why prefer "functionally not very reinforcing" over an account that can specify with precision the degree of coupling between an incentive and a response? Qualitative accounts of reinforcement are circular because when data are reduced to "increase or not," that single degree of freedom is exhausted in identifying the stimulus as a reinforcer. The present account is not circular because it says much more than "the rate went up."

"Functional analysis," like "Top Secret," hides deep mysteries and everyday bungles alike. It is the error term in the regression equation; the more variance we can take out of it and represent as structure, the richer our science

becomes. Isn't it time we became more functional ourselves – reinforced by what works – and less tied to Skinner's brilliant but provisional structure?

**Timberlake**, on the other hand, thinks I have not gone far enough; "the notion of a reinforcer remains unexamined in Killeen's theory; it simply exists and has measurable effects." Reinforcement is not just the conjunction of incentives with behavior in memory, he holds, but rather it "arise[s] when the local rates or patterns of responding constrained by a schedule differ from the amounts or patterns instigated by the stimulus conditions." His work with others on regulation, following in the tradition of Premack, has revolutionized the way we think about reinforcement. Things *are* more complex, on a number of dimensions, than stated by the present theory. But that is as it should be. Theories should give comprehensible pictures of nature, and to do that some things must be left out. Newton did not tell all there was to know about apples; not even all there was to know about gravity, concerning whose nature he framed no hypotheses. The notion of gravity remained unexamined in Newton's theory; it simply existed and had measurable effects. I do not think these principles omit anything crucial, however; Timberlake's conditions for reinforcement are nothing other than a practical guide to arrangements that guarantee  $\alpha$  and  $\zeta$  to be nonzero.

**R2.3. Ecology.** An evolutionary metaphor is favored by **Donahoe**. He lingers on the fact that selection acts on genes as they are expressed in their environment; that is, on the phenotype, using this as an argument against response rate, Skinner and others' primary datum (after Skinner abandoned his hypothetical construct "reflex strength"; Killeen 1988). If the environment affected responses as it affects the expression of genes, Donahoe argues, we would not expect reinforcement to have the same effect on responses that are controlled by different stimuli. No argument there. Let me see what I can do with that metaphor: memory is to behavior as the genotype is to the phenotype; selection acts on the second terms, while selections are epitomized in the first. In both cases, the vector of selection is longest when the coupling is tightest.

**Timberlake** notes that the learning of simple manipulation responses reflects "only a small portion of the role of learning in survival . . . We have not adequately explored the variety of learning mechanisms and their interactions . . . how . . . memory, location, time of day, and particular stimuli, responses, and motivational states produce learning that fits the ecology and evolution of a species." Other commentators add to the list, urging me to turn to what animals really do. Just a minute, please! Can't we pause, catch our breath, enjoy the view, see where we have been and how best to move on, before turning away from this new perspective on a half-century of research?

**R2.4. Explanation.** A theory that does not address inhibition is only half a theory, **Walker** believes. But inhibition is an explanatory construct, not a datum. I do not object to invoking it in principle but do not need it for the behavior I analyze, and so eschew it as an uncompensated cost in describing this domain of data.

Wearden notes that this theory goes to press on the semicentennial of Hull's *Principles of Behavior*, and he sees it as part of the same tradition. I am honored by the comparison, as Hull fully understood that models – verbal, pictorial, or mathematical – are the only way we can understand behavior, or anything else. But Hull was ahead of his time: he did not have enough data available to constrain his models nor computers to evaluate competing versions of them. Now we have both; it is time for another try.

### R3. Models

**R3.1. Dynamics.** Commentators such as Galicka and Gregson would rather have the dynamics up front. But a number of dynamic models can lead to the same asymptotic equations, and until I had those tied down, I would not have had a criterion of success for the dynamics. The asymptotic models are central in this theory, the dynamic equations but one route to them, and they are a route that still needs roadwork, as Gregson notes.

What the dynamics provide is one plausible and parsimonious machine that leads to the asymptotic models as attractors of behavior. The validity of the asymptotic models (Tables 4, 5, and 6) does not depend on the validity of the dynamic model outlined here. Nonetheless I have confidence in it and the state space in which I have embedded it. It works, it is part of a larger theoretical canvas (Killeen 1992), and it invites the development of competing dynamical models, and of data adequate to select among them.

**R3.2. Mathematics.** What is seen by Gregson are “three interwoven strands to [the] argument: calculus, time series, and attractor dynamics.” He is focusing on the lens, however, rather than through it to the objective, which is a finer appreciation of the statics and dynamics of behavior. Modern mathematics is not an intrinsically better speculum than classic mathematics; it is subtler and more widely applicable, but it is also more esoteric and often less robust. Ockham’s razor applies to the form of theory as well as to its parameters: no nuance beyond its need.

Gregson worries that I shift back and forth between continuous and discrete processes and suggests I stay with the continuous. This is eminently reasonable from a formal viewpoint, but it is not how memory works. The metered model I present in lines 1 and 5 of Table 4 captures the continuous updating of memory that occurs only during a response. My exposition started with the quantal because that is where I started, and I wished to introduce Equation 18 as it was forced upon me by the internal patterns in the data, not as an a priori assumption. Furthermore, I believed that this tactic of gradual enrichment of the model distributed the conceptual burden more evenly over the manuscript; this attempt at pedagogy may have helped the modal reader, but unfortunately it frustrated this mathematician.

Stephens is right in finding something wrong with Equation 1: it is a density, describing the changes in  $\beta$  as a function of changes in  $d$ , but is missing a dot over the  $\beta$  and a prime after the  $\lambda$ ’s. Please add them. I initially wrote it to exemplify a continuous analog of Equation 2

and failed to revise it after I discovered how memory decay varies with response duration, which led me to introduce the prime notation. No big deal: to use Equation 1, it must in any case be integrated over the duration of a response. Equation 18, a fundamental principle of the theory, is that integral. Its units are correct – the product of  $\lambda'$  ( $\text{sec}^{-1}$ ) times  $\delta$  ( $\text{sec}$ ) over  $\rho$  (1) is  $\lambda$ , which is dimensionless (or, equivalently, carries the dimension of a ratio of absolute numbers), as required.

**R3.3. Parameters.** The theory has too many parameters, says Rachlin. Too many for what? Does he deem different motivational states ( $a$ ), topographical variation in responses ( $\delta$ ), or memory of things before the last response ( $\lambda'$ ) irrelevant to a theory of behavior? Does he understand that the conceptually important  $\rho$ , which distinguishes classes of contingencies, is not generally given liberty in this analysis but is kept fixed at 1, letting variations in its value be absorbed by variations in  $\lambda$ ? Why does Rachlin prefer one- or two-parameter models, insensitive to such psychologically important variables, over these principles, which make their roles explicit?

As Stephens notes, I estimate the parameters from the data – that is the point of modeling, to reduce a data field, none of whose elements are particularly interesting in themselves, to a smaller set of theoretically interpretable parameters. Stephens is right that this reduces the power of the theory somewhat, but it hardly exhausts the degrees of freedom in these data. There are few models in any science that capture data with fewer than three parameters.

Much effort went into making this theory as parsimonious and yet general as possible. I was therefore amazed when Stephens eliminated two of these three parameters in creating his lovely figures. Normalization to eliminate nuisance parameters is routine; but many psychologists might not consider these a nuisance, as their values teach us useful things (as Metzger & Sagvolden and Nevin show). Furthermore, normalization is not without its cost – it decreases the degrees of freedom in the data exactly to the same extent as models with explicit parameters. Stephens is not discarding bathwater, just babies. This leaves him with fewer mouths to feed but also with fewer voices to speak for the data.

The simplicity of pure theory is inevitably complicated by the realities of application. Whereas the inverse square law in its elegant simplicity motivates the planets in their courses, predicting the orbit of satellites around the earth requires us to expand the gravitational field in a harmonic series. Each successive term reflects different gravitational features of the earth; because most of the higher-order terms are not identifiable in any physical way beyond their roles in perturbing the orbits, they are essentially free parameters. NASA has carried the series out to hundreds of terms involving hundreds of parameters. Truth (and applicability) are, as Bohr had it, complementary to simplicity. The three parameters corresponding to the three principles of the present theory strike, I think, a reasonable balance between those virtues. Subsidiary parameters – the coupling constant  $\rho$ , which is affected by the preparedness of the response, the stage of learning, and the efficiency of reinforcement contingencies; and the degree to which an incentive displaces

memory for responses – stand ready for deployment in applications where they prove necessary.

**Stephens** notes that both FR (fixed-ratio) and VR (variable-ratio) models predict similar patterns of data. This is because the hyperbolic coupling coefficient for VR schedules is the first two terms of a power series expansion of the FR coefficient. And, as Stephens shows, we find similar patterns of data. A nice prediction confirmed! But whereas Stephens would take pleasure in discovering two similar foraging strategies that he could explain as examples of convergent evolution and would never think to hold that as an argument against evolutionary theory, he is nonplussed when a psychological theory actually *predicts* convergent behavioral functions that he then finds to hold true.

It bothered **Stephens** that in fitting the wrong shoe to the data, the parameter  $a$  changed its value to take up the stress. Yet this is exactly what *should* happen! He has backed into a technique of validating the models – *ceteris paribus* all three parameters should remain invariant under different reinforcement schedules if the models are correct, as **Shull** clearly recognized. Conversely, if the contingencies are not captured by the correct models, we expect the optimization program to cope by shifting parameter values to take up the slack – as happened in Stephens's reanalyses. This shift does not indicate that the "parameters have . . . different meanings in the different models," but only that accurate assignment of meanings is contingent on having the correct analytic framework. It is in fact reassuring that Stephens found different values for the parameters, for it shows that the models address the data in testably different ways.

**Rachlin** is unhappy with the fact that  $\lambda'$  is not invariant across all motivational changes. I am too. It comes close, however, being constant in two out of three relevant experiments. Only Snyderman's (1983) data show a difference in memory-decay rate with motivational changes. Perhaps memory is affected by motivation; if the theory were to tell us this fact, repeatedly, that would be a good thing to know. Or perhaps Snyderman's data are anomalous. However future research clarifies this point, I am sure that Rachlin is happy with the parameter invariances and interpretability that characterize these tables in general. I am too.

Much of the pain associated with mathematical modeling occurs because models make it so much clearer when the framework is wrong (as **Walker** notes of Hull's models, as **Wearden** found of his own, and as some commentators think the case for mine). But there are compensating pleasures: a definite "no" is liberating, and an unambiguous "yes," as a framework finally clicks into registration with data, is one of the purest pleasures that science affords. Various commentators have already coded these principles into their computers in its pursuit. Their success will redound to all, for science is not a zero-sum game.

**R3.4. Assumptions.** One needs more assumptions to run an autoregression (AR) of indeterminate order as a model of the coupling process, **Gregson** notes. Of course. My integrated moving average (IMA) is mathematically identical to an AR of infinite order, with the weights diminishing geometrically. This geometric decrease is an intrinsic part of an exponentially weighted moving average and

constitutes a strong hypothesis about the nature of memory. In abandoning that assumption, Gregson buys more flexibility than he needs, at a price the data are unable to pay. As Feynman (1965, p. 168) has noted,

Two theories, although they have deeply different ideas behind them, may be mathematically identical, and then there is no scientific way to distinguish them. However, for psychological reasons, in order to guess new theories, these two things may be very far from equivalent, because one gives a man different ideas from the other. By putting the theory in a certain kind of framework you can get an idea of what to change . . . there are certain ways of changing one which looks natural that will not look natural in the other.

I chose the IMA as the framework because the exponential rate of decay of memory is a natural aspect of it, one which has, happily, already given several commentators "ideas of things to change." Its relationship to the AR gives Gregson ideas of how to develop the dynamics. Splendid!

**Rachlin** objects to the assumption that subjects may be making unmeasured responses of brief or extended duration; or may not be responding at all. But would he really prefer a theory that assumed that no unmeasured responses can occur; or that they must all be of the same duration; or that subjects are always responding even though we do not see anything happening? Rachlin believes that recognition of these possibilities is "too much inference," but I see it as contact with reality. The theory's ability to make useful predictions in the face of these realities is one of its strengths.

**Timberlake** wonders whether the same data could be predicted from different equations, and whether the same equations can be produced from different assumptions. Take pencil in hand, Bill, and find out! Competition to capture data by different theories culls the weak and refines the strong.

**R3.5. Heuristics.** The precision offered by mathematical models may become too "seductive" for **Reed**, who argues that the benefit of models is usually found in the heuristic relationships they entail. Seduction often leads to enduring relationships, however, that are not substitutes for sex but ramifications of it. We can enjoy both, but without the hard core, the heuristics are just talk.

## R4. Memory

**R4.1. Minding.** A lovely illustration of the central problem addressed by this theory is provided by **Shull**: it is memorial representations that we reinforce, but how do we know what's on the subject's mind? Do we thank the child for the flowers, or chastise her for uprooting them? Do we praise the dog/spouse for returning home, or upbraid him for leaving? Our recognition of the problem does not create it but may help to solve it. The most recent act will be weighted most heavily in memory, but how much of memory does it occupy? Are the proffered flowers a quick emergency tactic when the child is discovered rooting in the garden? Or was the child thinking all along of the pleasure they would bring her parent? This is why understanding intention is important: it tells us what has been

on the subject's mind and therefore what is available for reinforcement or punishment. We may appraise children of what they are about to be rewarded/punished for, but this fills memory with an authority describing acts, and the authority easily becomes more tightly coupled to the consequences than the acts themselves. We might have a felon reenact his crimes before punishment, but to the extent the nearby jailers are on his mind, it is they who will in the future be avoided, not theft. Skinner incorrectly believed that punishment did not work; one of the reasons for that belief, I think, is that the agents of punishment are usually more salient than the acts they punish.

**R4.2. Chunking.** Numerous commentators (e.g., Bolles, Galicka, Pear, Shull, Walker) mention, and Reed discusses at length, the problem raised by sequential dependencies for my account. I agree that this is one of the most exciting, and urgent, developments of the theory that lies ahead. Such dependencies occur when perception elaborates glimpses into scenes, as chess masters reconstitute the pieces from a glance at the board; they occur when our memory redintegrates tunes upon receipt of a few bars; they occur when our legs dance steps we cannot describe. This is the brilliance of evolution; it has given us a causal lens – short-term memory (STM) – that can focus on the informative and assume the routine.

How might the theory accommodate these facts? When constituent elements of a response become automatized, we lose awareness of them, they will not index memory, and thus  $\lambda$  will be smaller than expected. Such constituents may also be unavailable for shaping – or for recall, although they may be reconstituted from the rules that made them automatic. As Bolles and Reed note, marking a response with a salient stimulus can reinstate it to STM. Some perceptions, such as the frequency of events, may not be processed in the same way by STM. But if, as Shimp and I suspect, reinforcement primarily affects behavior of which we are aware, then these models of learning and performance may provide the most secure route for understanding awareness and for imputing it to nonverbal animals. Scary thought.

Will such compression of the response also extend the ambit of the incentive (*a*)? Galicka suggests one experiment, Neuringer and Chung (1967) performed another. They found that by segmenting a long FR into components they could greatly enhance response rate. The conditions of segmentation necessary for their technique to work strongly suggest that their subjects were forming response units. Platt and Day (1979) found effects in extinction that also suggest conservation of effort for response units. Introspectively, well-learned skills seem effortless compared to newly acquired ones. The present theory does not explicitly provide a mechanism for such response compression; however, it recognizes two distinct ways in which chunking may occur and provides two parameters whose systematic changes might indicate its occurrence.

**R4.3. Coding.** In recent years there has been a flowering of research among the individuals, cited by Wilkie & Sak-sida and others, attempting to clarify how memories are coded in nonverbal animals. There is evidence for both

analogical and categorical coding (e.g., Santi et al. 1993). Both find a natural place in this theory.

What does an animal do when it observes a stimulus? After it orients, it sees or hears or feels or tastes; if these perceptual acts require  $\delta$  seconds, then during a CS of duration  $t_{CS}$  there are  $N = t_{CS}/\delta$  opportunities for such observations. Substituting into Equation 8 we derive the amount that a CS of fixed duration will fill memory:

$$M_{CS} = \rho(1 - e^{-\lambda' t_{CS}/\rho}). \quad (R1)$$

It follows that coupling, and thus control, will increase with the duration of the CS. This is the case (e.g., Roberts & Grant 1974). It also follows that coupling will be greatest when attention is not overshadowed by more salient stimuli or attracted by other stimuli that are more tightly coupled to reinforcement (i.e., when  $\rho \approx 1$ ). This too is the case. Equation R1 constitutes a tentative extension of the theory to account for perceptual acts, that is, for stimulus control.

A simple analogical code for the *duration* of a CS is the extent to which it occupies memory at the time it is queried,  $M_{CS}$ ; if considerable, it must have been the long stimulus; if little or none, it must have been the short stimulus (or no stimulus at all: Kraemer et al. 1985). If the query is delayed, other responses during the delay will index memory and bias the judgment toward the short stimulus ( $M_{CS}$  will decrease exponentially with the delay; Spetch & Wilkie 1983; Wearden & Ferrara 1993). Figure R1 shows this "choose-short" effect: whereas the probability of calling a probe stimulus "long" increases with its duration, the functions are substantially lower with 5- and 10-second delays before the opportunity to respond. The curves are from Equation R1 multiplied by the exponential loss during the delay; the two parameters are  $\lambda'_{CS} = 0.20$ , and  $\lambda'_{Delay} = 0.06$ . The smaller rate of decay during the delay indicates that fewer observational and other response occurred during it. This model also predicts that proactive interference from presamples will bias subjects to choose long, because it falsely loads memory. It does (Spetch & Sinha 1989). It will be interesting to evaluate this model against data from experiments with varied

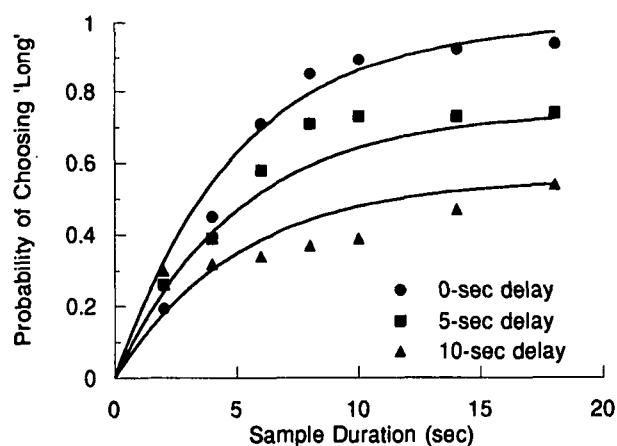


Figure R1. Data from Spetch and Wilkie (1983) demonstrating the "choose-short effect." Note that the probability of characterizing the preceding stimulus "long" increases with its duration and decreases with the delay after the offset of the stimulus. The curves are from a model of analogical memory (Equation R1).

intertrial intervals, varied intensities of sample stimuli, and other manipulations that should affect  $M_{CS}$  in predictable ways.

**R4.4. Memory strength or memory sampling?** On any trial the subject must choose one response or another; we can either build-in a decision mechanism to map  $M_{CS}$  onto action, or more simply treat Equation R1 as the probability that the CS was in memory and make the predictions directly. The latter approach generated the curves in Figure R1, but it changes the interpretation of Equation R1 in a way that makes it much closer in spirit to Estes's stimulus-sampling model (Hanson).

**R4.5. Analogical bisection.** A memory loading halfway between the couplings for long and short stimuli corresponds to a duration of:

$$t_{PSE} = -\frac{\rho}{\lambda'} \ln \left( \frac{e^{-\lambda' t_{CS1}/\rho} + e^{-\lambda' t_{CS2}/\rho}}{2} \right), \quad (R2)$$

which ranges between the arithmetic mean and the value of the shortest stimulus as  $\lambda'$  increases; it takes a value near the geometric mean for representative memory-decay rates. It corresponds to one set of generalized means (Hardy et al. 1959) and constitutes a model of bisection for analog memories.

What of categorical coding? Responses are elicited during intervals between reinforcers, and animals will be emitting different responses during the end of a long CS than during a short one (Killeen & Fetterman 1988). Whether they are orienting to a particular part of the chamber or repeating "long" to themselves, these categorical responses serve to mediate the correct choice response. With categorical encoding we expect less evidence for a choose-short bias, different memory functions (cf. Kraemer & Roper 1992; Wixted 1989), and different points of subjective equality.

**R4.6. Retrieving.** The present theory, Krank notes, does not include an explicit role for retrieval processes. This is true. Can the present theory account for the data Krank relates? I think so. Why should a CS differentially activate responses appropriate to the US it has predicted? As noted toward the end of the target article, instrumental responses often share properties of the consummatory responses toward which they are directed, as do Pavlovian CRs. These may include approach, salivation, search images, and so on. Insofar as these fractional antedating goal reactions are facilitated by a CS, they should enhance the operant responses of which they are a component. But is this really very different from Krank's and from Hintzman's (1986) theories? I think not. There exist substantial literatures on both the behavioral and cognitive sides of this issue. Here is yet another land-bridge, whose crossing requires of the traveller only a little practice with a foreign vocabulary. The grammar, insofar as it articulates the phenomena, should be familiar.

**R4.7. Reporting.** Interested in the implications for memory in verbal animals, Svartdal asks whether, because only remembered responses ( $M$ ) are reinforceable ( $R$ ), we can infer that a response must be reinforceable to be remembered? That is, given  $\tilde{M} \rightarrow \tilde{R}$ , may we infer  $\tilde{R} \rightarrow \tilde{M}$ ? Rewriting the premises as  $R \rightarrow M$  and  $M \rightarrow R$ ,

respectively, we see that any attempt to infer the latter from the former runs against the conversion fallacy. Empirically, both implications may hold, but neither is entailed by the other.  $M \rightarrow R$  is an interesting proposition, however, one which, with proper qualifications concerning perceptual acts, may be true.

The problem with Svartdal's later development of my presumed affirmative reply to his question is that by "remembered" he means "reportable," whereas I mean present in STM. As he has shown (Svartdal 1991; 1992), it is possible for reinforcement to strengthen a response that an organism cannot report a few seconds later – indeed, the very act of reinforcement erases the memory (while at the same time increasing the propensity to emit the response again; Killeen & Smith 1984). Furthermore, we remember movements that we cannot verbalize. For such reasons we may have reinforcement without the ability to talk about the actions that was reinforced, even though  $R \rightarrow M$  is true. Reinforcement, by moving events to long-term memory while erasing them from STM, is one of the agents of implicit memory.

Savvy humans rehearse acts that have led to pleasure, and rehearsal that pleases improves performance; but not all organisms, man or beast, do this. And not all acts are easily re-presented. The easiest acts to rehearse are the ones most easily described, for example, "press the lever fast," and it should not be surprising if instructions like that override the more complex control by contingencies. Simple and easily repeated instructions and aphorisms are themselves likely to be present in memory contemporaneously with actions; they share in their reinforcement and thereby warp the behavioral trajectory into a local minimum. If habit is the great flywheel of society, tutelage aims its vehicle, always imprecisely.

## R5. Behavior

### R5.1. Present analyses.

**R5.1.1. Segments.** There is a slight misconstrual of the theory by Pear: it does not require segmentation so much as respect it; nor does it assume that between segments the animal is engaged in other responses. The animal is *not* always behaving (and *that* is an intriguing commitment of this theory). Responses may be brief or extended (and that is why the expansion of  $\beta$  in Equation 18 is important). As I noted in section 14.1.3. and elsewhere in this response, some responses are mixtures of elements, whereas others are compounds; in the former case, various elements may be differentially shaped as a function of their particular coupling, whereas in the latter the compound is strengthened as a whole (Reed et al. 1991). Action patterns/ethograms, systems of movement notation, Hanson's (1991) theory of differential activation levels, and Pear's own system of continuous recording (e.g., Pear 1985) are all important approaches to this problem.

Pear finds it difficult to parse continuous movements into constituents. As Bolles suggests, it may be worthwhile to shift the burden of definition onto the subject, akin to the way in which backpropagation programs shift the burden of algorithm design onto the computer. If we find differences between  $\lambda$ 's for the same cluster of re-

sponses over the course of training, or between sequences with and without strong sequential dependencies, it will provide evidence for, and an index of, chunking. If we probe a stream of behavior with an extrinsic stimulus, memory for the stimulus should decrease quantally at the boundaries of chunks.

**R5.1.2. Sequences.** Is the failure of the theory to take sequential dependencies into account, **Pear** wonders, the cause of the systematic underprediction of Catania's data in Figure 6? I suppose this is possible, but I do not see how. I do see an alternative hypothesis that accounts for the discrepancies reasonably well: the birds have a key bias. We know that these are quite common. Allowance for a bias toward the A key of one or two percentage points per peck would bring the columns into alignment.

Some commentators wondered how animals "average" a heterogeneous sequence of responses. Good question. As my father told me, you cannot add apples and oranges. Unless you are simply tallying pieces of fruit, I guess. What are the animals tallying? In section 14 I discuss the multidimensional space that frames trajectories of responses. Each dimension – each response class – is strengthened according to the position of its exemplar in memory when an incentive enters; this is represented by the projection of the trajectory onto that dimension. For operant classes whose elements have been compounded by automatization, it is possible that all constituent elements will be strengthened equally.

In the target article I have focused on the simplest cases, where contingencies drive the majority of behavior over one dimension (ratio schedules), or where we may get away by assigning all the other behaviors to an "other" category that absorbs the proportion  $1 - p$  of memory. Our measure of the target response depends on how many of the dimensions we count – whether we tally left-paw presses and right-paw presses separately or add them. Our measure of coupling depends on whether the contingencies constrain only one of the dimensions, permit their union (either paw), or require their intersection (both paws), and so on. I agree with **Timberlake** and other commentators that this extension to the full response space of the organism is an exciting prospect. It will involve some nontrivial issues in the dynamics of behavior and will lead to the discovery of unsuspected response-response constraints. There remains plenty of work yet to be done on the statics of behavior – the asymptotic and equilibrium conditions of traditional experimental designs. Plenty of excitement for all, each to their own tastes and skills.

**R5.1.3. Scatter.** I do not share **Gregson's** concern about the scatter around the predictions in three of the figures. In Figure 5, the curve is predicted from the model associated with Figure A1; it requires only one parameter (memory-decay rate), which for display I have fixed at  $\beta = \frac{1}{4}$ ; I consider its general accord with the data of four different animals a marvelous vindication of the model. The curve in Figure B1 is also a one-parameter model, and I fixed that parameter at the modal response duration for pigeon's key-pecks ( $\delta = 0.3$ ). The data are noisy, but I do not consider that the model's fault. Whereas the deviations in those figures are not systematic, that is not the case in Figures 14 or D2; but here the data are

themselves simulacra, derived as plausibility proofs for the asymptotic models. The fits are more than good enough for that. Note that the greatest deviation is found in applying the model to overall response probability on long FR schedules. The reason is straightforward: overall probability averages two nonhomogeneous parts of FR performance. Such lumping can be dangerous, but I included the curve to show that the results are not fatal; the model provides a reasonable approximation for attainable ratio values. At long ratio values where the postreinforcement pause carries the simulacra below the model, the error in predicting response rates will be absorbed by slightly decreased values for  $a$ . Thus we expect that if the model is applied to overall rates on FR schedules, the value of  $a$  will be smaller than its true value as inferred from the running rates under these schedules.

**R5.1.4. Species.** Is it really sensible, **Wearden** wonders, to fit the same model to the very different behaviors of human subjects in McDowell and Wixted's (1986) study, and then have to puzzle over the wild differences in parameters? I certainly asked myself that question more than once. Such attempts put theory to the test, and permit one either to generalize or to draw boundaries; but they also put one on the spot. I was relieved when I came up with what was to me a plausible explanation of the memorial effects (the nondisruptiveness of symbolic reinforcement). But what of the huge values of  $a$  for two of the subjects? It looks as if they would go on responding until the cows came home. I think this tells us something about the motivation of human subjects in psychology experiments, whose demand characteristics – helping science, passing the course, unavailability of competing behaviors – often profoundly outweigh our paltry reinforcers. These large values for  $a$  may be veridically reflecting the fact that the motivational processes at work are different from what we might like to think.

**R5.1.5. Prediction versus comprehension.** It is suggested by **Bradshaw** that the addition of the memory term in Equation 12 is hardly justified by the deviations that it permits the theory to account for, and those deviations might just as readily be accommodated by other, quite reasonable processes. Absolutely true. I would never go to press with a theory that started with Equation 12 as a fix for Herrnstein's model – as I would not with a general theory that could not encompass his data. I was greatly relieved when the combination of fundamental equations simply gave me a model for VI performance that worked. I then discovered that the mathematical muses were Skinnerians, for after all of the extinction they had put me through, they initiated partial reinforcement by leading me to Equation 12, which separates the motivational/constraint forces (Herrnstein's hyperbola) from the memorial forces in so lovely a manner.

Like **Galicka**, I think that too much undirected schedule research has been done; but, thanks to that excess, there were ample data to guide construction of this theory, as there are still more than enough left to test its implications. I agree with **Galicka, Wynne, and Smith** (1992) that it is time to recognize that Skinner's criteria of control and orderliness are instruments of understanding, not substitutes for it.

## R5.2. The next generation

**R5.2.1. Loose ends.** The majority of commentators noted the potential for generalizing the theory to different phenomena. The asymptotic models do not account for temporal structure, as **Wearden and Wynne** would like. An account is provided by the dynamic model, which has response probabilities changing over time in ways consistent with the data. Pausing persists because early responses are not in memory at the time of reinforcement, and there are other things to do that are more intrinsically reinforcing (e.g., looking for stray seeds, for ways out of the box, and so on). But these dynamic models require further development (e.g., assignment of predispositions –  $\rho_{\text{Other}}$  – for competing behaviors in that context and motivational state). Nor does the theory provide an account of the partial reinforcement effect (PRE) – although such an account is provided by arguments elsewhere (e.g., Killeen 1984; 1991); it will be an interesting piece of work to try to unify these accounts.

Pear notes that the extension to concurrent schedules will be far more complex than Herrnstein's theory. Thank goodness! As **Galicka** observes, such schedules are themselves far more complex than such a simple analysis allows, as are their effects on behavior. I feel toward them as Darwin did toward parental investment: "the whole problem is so intricate that it is safer to leave its solution for the future" (cited in Cronk 1993).

Herrnstein's theory was a landmark, but off the mark in several ways. The interpretation of  $1/a$  as reinforcement for other behavior was not supported by subsequent research (e.g., Dougan & McSweeney 1985). Although "generalized matching" (Baum 1974; Davison & McCarthy 1988) is ubiquitous, it is not derivable from Herrnstein's theory. Herrnstein provided the first simple, robust, and general mathematical model in this area; as the first quantitative Law of Effect, it ranks with the Rescorla-Wagner (1972) model in theoretical importance. By convincing a generation of young behaviorists that, if one was careful, one could go from measured behavior to inferences about unseen processes back to implications for behavior, it was essentially the first draft of the present theory.

**R5.2.2. Applications.** The theory is inapplicable "to even very simple extensions of present procedures," Rachlin thinks, but I think that Rachlin simply has not tried. Take for example, the data of Hackenberg and Axtell (1993), which Rachlin cites as counterexample to the theory. The authors gave humans a choice between: (1) a progression of delays to an incentive, with each successive delay 5 seconds longer than the previous; and (2) a fixed delay of  $T$  seconds, the choice of which would also reset the progression. How might we go about applying the principles to these contingencies? First, we simplify. We do not know the duration of a response. Assume it is the same for each alternative and only make predictions about *relative* magnitudes of coupling. We do not know the rates of responding. Assume they are about the same on each alternative. Since the incentives were a brief tone and point display, we may assume that they had negligible impact in resetting memory, as was the case when McDowell and Wixted (1986) used similar incentives (Fig. 9). If this is the case, the choice response is coupled not only to its nominal

incentive but also to all others following it from either schedule, with coupling decreasing as an exponential function of the delay between response and reward. Again simplify: assume that only the next incentive, plus those obtained from the next sojourn on the other alternative, affects memory.

This model is shown in Figure R2, which has subjects confronted with a fixed interval  $T = 30$  sec on one alternative (plus the eventual incentives on the other side) versus the third increment on the progressive interval (plus the eventual incentive after the fixed interval on the other side). A choice of the FI will be paired in memory with incentives up to four times in this example, each one at an increased remove; the cumulative strengthening of the memory is given by the sum of the traces at the origin. A choice of the PI side is reinforced in memory twice, by the sum of the traces shown. Which receives the greatest total coupling in memory? In the case pictured, the FI side. The analytic solution at equilibrium gives the amount of time spent on the PI side to be:

$$nt = \ln[(1 + e^{-\lambda T})(1 - e^{-\lambda t})e^{+\lambda T} + (1 + e^{+\lambda t})/2]/\lambda \quad (\text{R3})$$

where  $n$  is the number of choices of the PI side,  $t$  is the increment on that side (5 sec), and  $T$  the FI value. For  $T$ 's

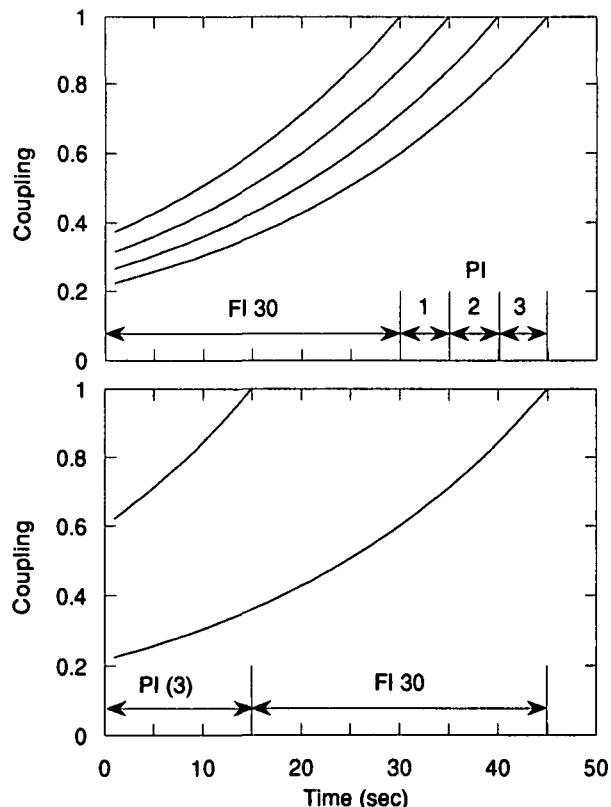


Figure R2. A model for the experimental design of Hackenberg and Axtell (1993). The top panel shows the couplings of a choice response to the incentives that follow it on that (FI) alternative, and during the following sojourn on the other (PI) alternative. The bottom panel shows the couplings of a choice response to the incentives that follow it on that (PI) alternative, and during the following sojourn on the other (FI) alternative. When the sums of these traces at the choice-point (the origin) are equal, the subjects should be indifferent. This equilibrium value for the PI schedule is given by Equation R3.

of 15, 30, and 60, with  $\lambda = 0.03$ , Equation R3 predicts sojourns on the PI key up to 12.1, 14.9, and 24.1 seconds before switching to the FI side. The mean sojourns for Hackenberg and Axtell's (1993) subjects were 10.2, 14.6, and 23.5 seconds. For other conditions in which choices of the FI did not reset the PI, the model predicts the data equally well.

The point is not that these successful predictions vindicate the theory; there were too many simplifications made in arriving at them, too few data points to feel smug about the goodness of it, and other conditions with additional time-outs that will require additional consideration. The point is that a little effort in attempting to apply the principles will often yield dividends. Until one attempts to apply the principles, it is premature to say that they are inapplicable. One need not derive equations such as R3 to apply them: a simple computer evaluation of the tails of the exponential decay functions shown in Figure 3 is adequate to find the critical points.

As I noted in section 14.2, "More accuracy in predicting the effects of frequent brief incentives will be gained by integrating over the incentives and what preceded them (and over subsequent events as well!)." That integral, carried forward to infinity, yields similar predictions but must be solved numerically.

A new parameter has been implicitly introduced here: the extent to which an incentive displaces memory. But even this is susceptible to direct measurement (see sect. 13). As the duration and impact of the incentive increases, these principles predict decreased benefit from the successive incentives, and thus longer sojourns on the PI before switching. That should be an easy prediction to test.

**R5.2.3. Predictions.** It is suggested by Wynne, among others, that "the real proof of the utility of a novel theory is in the novel predictions and tests it generates"; he goes on to ask what these might be. One may argue with Wynne's premise. Brush (1989) claims that scientists historically give novel predictions no more weight in evaluating a theory than the degree to which it successfully accounts for known facts – sometimes even less weight until the nature of the new "facts" is thoroughly understood. Einstein (1930, p. 273) did "not consider the main significance of the general theory of relativity to be the prediction of some tiny observable effects, but rather the simplicity of its foundations and its consistency."

Nonetheless, as Bradshaw notes, the present theory "is not only bristling with testable . . . behavioural experiments; it also points the way towards novel *post hoc* interpretations of neurobehavioural experiments." Some examples: once the relevant parameters are determined for an organism on one type of schedule, its behaviors on all other schedules should be predictable *a priori*. Naive organisms will be more readily shaped on an FR 20 schedule than on an FR 1 with matched rate of reinforcement. Drugs that impair STM will have greatest impact at high rates of reinforcement (notice, for example, that the size of memory is multiplied by the rate of reinforcement in Equation 12). Increasing the durations of a reward will have a diversity of effects in interaction with rate of reinforcement (sect. 5.5.3). The minimum duration of a response interacts multiplicatively with motivational variables, and therefore confounds them – and does so

differentially with the different schedule types (differentiation of the general models in Table 6 shows that response cost is most important under large ratio schedules and small interval schedules). Analogical and categorical coding will have different signatures (Equation R1) and bisection points (Equation R2). Some contributors have demonstrated other predictions that I have not envisioned, and yet others will be educated.

**R5.2.4. Extensions.** These principles provide a useful retrospect as well as a hopeful prospect. Bolles's recording of a left turn for a rat who thought it was going West underscores the utility of the present "theromorphic" (Timberlake) perspective on the classic response-learning versus place-learning controversy.

One can see Walker turning the theory over in his hands like a Rubik Cube, finding more current issues, such as spatial behavior, punishment, and inhibition, that need alignment, and suggesting ways in which this might occur. This is laudable. I also have some notions about these things. For Hull's (1943) rat forced to pass the door before entering it to eat, the door is closely coupled on the return transit, and at some remove in memory on the first pass; but it is the same door in each case, so these couplings summate, generating a strong tendency to sign-track. How might one maximize the coupling to only the second approach? By lengthening the distance by which the rat must pass, so that the vision of the door on first pass gets as little as possible residual strengthening; by changing the look of the door on the two approaches: making the alley narrow so it cannot be observed head-on from both orientations, having the ends of the alley different patterns, the floors on either side different textures, and so on.

Just as organisms attempt to re-create the contents of memory as it was at the time of reinforcement, they avoid anything that reconstitutes a memory similar to that occurring at the time of punishment. Escape, defensive burying, and denial are ways of accomplishing this. In situations such as Sidman avoidance, the coupling coefficient can tell us exactly how tightly coupled the lever-press is to the shock, but we must infer the coupling to the context and to other responses. The theory predicts the elegant results of Hineline (1976), showing that animals will press a lever to forestall a shock, even though that has absolutely no effect on overall shock frequency. Walker is right: extension to the domain of aversive control opens a whole new set of opportunities for theory development as well as problems for it, such as incorporating the intrinsically different couplings for species-specific defense reactions (Bolles 1970), and reconciling this "avoidance as de-coupling" view with Denny's (1991) theory of avoidance as an approach to safety.

Other commentators suggest other extensions. Wynne would like to manipulate STM in some way, perhaps with drugs, and observe changes in the conditionability of the subjects. Bradshaw is also excited about the role that neurobiology can play in helping develop theories of behavior. Wynne and Galicka agree that the theory should be tested against other dimensions of responding. This is all to the good, because no one is interested in studying a cube with only one element. I hope the commentators, having analyzed the openings in the cube, will now realign the parts of it they understand best.

**R5.2.5. Realizations.** Some commentators have, happily, already begun exploring extensions. Galicka notes the consistency of existing research with the premises of the theory concerning the reinforcement of IRTs. Metzger & Sagvolden find Equation 12 useful in making sense of data from hyperactive and normal rats. Hyperactive rats apparently get four times the mileage out of each incentive. But they also have twice the minimum IRT of normal rats, and these commentators wonder why generally faster animals should be slower in this parameter. A similar paradox occurs with me. I am much faster after my third cup of coffee, yet it takes me longer to get my key in the lock. I think the resolution may be similar in both cases: these parameters measure different causal factors in performance, which may act in concert or in opposition. The hyperactive rats may be intrinsically less adept, or the larger value of  $\delta$  may be a byproduct of their hyperactivity, causing more of their responses to be fragmentary or aborted. Metzger & Sagvolden are obviously comfortable with both modeling and behavioral preparations, so I hope they will pursue this question, perhaps by exploring the suitability of Equation B8 as a more appropriate model of ceilings on rates for responses such as lever-presses that may be extended in time.

Nevin has leaped ahead to an implementation of the theory that I thought would require a couple of years to effect. His analysis substantiates my hunch that a large part of behavioral contrast might be due to motivational changes across components, just as McSweeney (1992) has shown large within-session changes in response rates that may be attributable to changes in the incentive value of stimuli (that is, to local changes in  $a$ ). Nevin's demonstration that the theory might predict the difference in resistance to change associated with different reinforcement conditions exceeded my vision. Contrast is a complex phenomenon, and many more analyses will be needed to understand its various aspects. Further theoretical development will also be necessary to permit us to predict the changes in  $a$  that Nevin found, but his observations concerning overall constancies may hold that key. I believe that the effects of reinforcer devaluation that Walker mentions will also be concentrated in  $a$ , although the nuances of the associationist analyses are beyond the theory at this point. If the present theory just makes a contribution to understanding contrast it is clearly beginning to pay the interest on the degrees of freedom it had to borrow to get started.

## R6. Science

**R6.1. Popper.** The falsificationist attitude of several commentators elicits the following observation: if Popper's theory of scientific inference was correct, a single exception should cause us to abandon a theory. But there are exceptions to all theories (Lakatos 1978). Hence, if we believe Popper, we should abandon *his* theory. We should not wholly believe him, however, which lets us appreciate and retain his insights for the domains in which they are appropriate. All theories should entail models that are testable. Whereas models are more or less falsifiable, theories are only more or less powerful. When models fail, it may indicate that some of the theoretical assumptions were wrong, or that some of the data were wrong, or

that the models themselves were wrong in not properly representing the theory or in being misapplied to the data. All of these slippages have occurred – witness the early empirical disproofs of relativity theory, the failure of error bars around estimates of fundamental physical constants to overlap from one generation to the next, the difficulty in finding renormalization models that permit the application of quantum theory to physical systems, and so on.

Like the Red Queen, scientists can devise a half-dozen models before breakfast, and by dinner they can conclude that most of them are impossible. They are slower to accept or reject theories because these have many more instantiations that must be evaluated for utility; eventually, however, not only do we reject theories, we also accept them. We do both tentatively and fallibly, with certitude reserved for philosophers and royalty, not scientists.

There was a dark side to Popper's brilliance; his logic was an apologia for negativism, which, along with the defense of intellectual property, is a ubiquitous proclivity of academics. The question is *not* whether this mechanics of action can be disproved, but whether it is useful or can be made so by our community: What assumptions should be added or removed; what new models may be adduced, what new data become relevant, what reinterpretations of extant data will bring them under the purview of the theory? There are now better versions of the logic of science than Popper's: Laudan (1977) sees science as a problem-solving activity, with theories setting the problems and suggesting the appropriate tools; Hestenes (1992) sees science as the playing of modeling games, with theories providing the rules – selecting the pieces, setting the board, stating the objective and the legal moves – for a family of games. I believe these will prove to be more useful theories of science than Popper's; try them.

**R6.2. Pandora.** It is suggested by Hanson that I have opened Pandora's box. Pandora succumbed to curiosity and unwittingly opened the gold box in which Prometheus had confined all the evils that could trouble mankind. Curiosity likewise moved me to pry the lid of the black box in which Skinner, in the eyes of some behaviorists, had confined all the mentalisms that could trouble mankind. But there is another version of the story, in which it is not evil that is confined in the box but blessings, all of which escaped but hope. I prefer that version, for what is science other than the apparatus for realizing our hope of understanding our world? I prefer it because Pandora's gift of curiosity – of sticking her nose in places it did not belong – motivates much gleefully good research. Finally, I prefer the latter version because choosing it exemplifies a covert activity of all scientists: scientists select among alternative theoretical assumptions, fabricate different instantiating models, and redefine the domain of relevant data so they may discover the parts that fit together best and thereby tell the most harmonious story.

**R6.3. Psychology.** To “use a little psychology on someone” is to be subtle; to intuit the controlling variables. It is to control not by force but by shifting motivation, by redirecting attention, by providing alternatives; it is to understand intention, to “get inside the other person's head.”

As Bolles and others suggest, it is time to use a little psychology on our subject.

## References

Letters *a* and *r* appearing before author's initials refer to target article and response respectively.

- Allen, J. D. & Kenshalo, D. R. J. (1976) Schedule-induced drinking as a function of interreinforcement interval in the rhesus monkey. *Journal of the Experimental Analysis of Behavior* 26:257-67. [aPRK]
- Anger, D. (1956) The dependence of interresponse times upon the relative reinforcement of different interresponse times. *Journal of Experimental Psychology* 52:145-61. [aPRK]
- Annable, A. & Wearden, J. H. (1979) Grooming movements as operants in the rat. *Journal of the Experimental Analysis of Behavior* 32:297-304. [JHW]
- Arbuckle, J. L. & Lattal, K. A. (1988) Changes in functional response units with briefly delayed reinforcement. *Journal of the Experimental Analysis of Behavior* 49:249-63. [aPRK, KAL]
- Atkinson, R. C. & Shiffrin, R. M. (1968) Human memory: A proposed system and its control processes. In: *The psychology of learning and motivation: Advances in research and theory*, ed. K. W. Spence & J. T. Spence. Academic Press. [CPS]
- Atkinson, R. C. & Wickens, T. D. (1971) Human memory and the concept of reinforcement. In: *The nature of reinforcement*, ed. R. Glaser. Academic Press. [aPRK]
- Baer, D. M. (1981) The imposition of structure on behavior and the demolition of behavioral structure. In: *Nebraska symposium on motivation, 1981: Response structure and organization*, ed. D. J. Bernstein. University of Nebraska Press. [RLS]
- Barofsky, I. & Hurwitz, D. (1968) Within ratio responding during fixed ratio performance. *Psychonomic Science* 11:263-64. [aPRK, DWS]
- Barrett, J. E. (1986) Behavioral history: Residual influences on subsequent behavior and drug effects. In: *Developmental behavioral pharmacology*. Vol. V: *Advances in behavioral pharmacology*, ed. N. Krasnegor & D. B. Gray. Erlbaum. [KAL]
- Baum, W. M. (1973) The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior* 20:137-53. [aPRK, PR]
- (1974) On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior* 22:231-42. [rPRK, JJP]
- (1979) Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior* 32:269-81. [JJP]
- (1992) In search of the feedback function for variable-interval schedules. *Journal of the Experimental Analysis of Behavior* 57:365-75. [aPRK]
- Beer, R. D. (1990) *Intelligence as adaptive behavior: An experiment in computational neuroethology*. Academic Press. [WT]
- Bharucha-Reid, A. T. (1960) *Elements of the theory of Markov processes and their applications*. McGraw-Hill. [aPRK]
- Bock, G. R. & Marsh, J. (1993) Experimental and theoretical studies of consciousness. *Ciba Foundation Symposium* 174. Wiley Interscience. [HR]
- Boe, E. E. & Church, R. M. (1967) Permanent effects of punishment during extinction. *Journal of Comparative and Physiological Psychology* 63:486-92. [SFW]
- Bolles, R. C. (1970) Species-specific defense reactions and avoidance learning. *Psychological Review* 71:32-48. [rPRK]
- (1983) The explanation of behavior. *The Psychological Record* 33:31-48. [aPRK]
- (1989) Acquired behavior: Aversive behavior. In: *Ethoexperimental analysis of behavior*, ed. P. F. Brain, R. J. Blanchard & S. Parmigiani. Kluwer. [RCB]
- Bradshaw, C. M. & Szabadi, E. (1988) Quantitative analysis of human operant behaviour. In: *Human operant behavior and behavior modification*, ed. G. C. L. Davey & C. Cullen. Wiley. [CMB]
- (1989) Central neurotransmitter systems and the control of operant behaviour by "natural" positive reinforcers. In: *The neuropharmacological basis of reward*, ed. J. H. Lieberman & S. J. Cooper. Oxford University Press. [CMB]
- Bradshaw, C. M., Szabadi, E. & Bevan, P. (1978) Relationship between response rate and reinforcement frequency in variable-interval schedules: The effect of the concentration of sucrose reinforcement. *Journal of the Experimental Analysis of Behavior* 29:447-52. [CMB]
- Branch, M. N. (1977) On the role of "memory" in the analysis of behavior. *Journal of the Experimental Analysis of Behavior* 28:171-79. [aPRK, KAL, JJP]
- Breland, K. & Breland, M. (1961) The misbehavior of organisms. *American Psychologist* 16:681-84. [aPRK]
- Brooks, L. R. (1987) Decentralized control of categorization: The role of prior processing episodes. In: *Concepts and conceptual development: Ecological and intellectual factors in categorization*, ed. U. Neisser. Erlbaum. [MDK]
- Brown, J. A. (1958) Some tests of the decay theory of immediate memory. *Journal of Verbal Learning and Verbal Behavior* 2:34-39. [aPRK]
- Brush, S. G. (1989) Prediction and theory evaluation: The case of light bending. *Science* 246:1124. [rPRK]
- Capaldi, E. J. (1992) Levels of organized behavior in rats. In: *Cognitive aspects of stimulus control*, ed. W. K. Honig & J. G. Fetterman. Erlbaum. [aPRK]
- Catania, A. C. (1969) On the vocabulary and grammar of behavior. *Journal of the Experimental Analysis of Behavior* 12:845-46. [KAL]
- (1971) Reinforcement schedules: The role of responses preceding the one that produces the reinforcer. *Journal of the Experimental Analysis of Behavior* 15:271-87. [aPRK, JJP]
- Catania, A. C. & Cutts, D. (1963) Experimental control of superstitious responding in humans. *Journal of the Experimental Analysis of Behavior* 6:203-8. [JJP]
- Catania, A. C. & Keller, K. J. (1981) Contingency, contiguity, correlation, and the concept of causation. In: *Advances in analysis of behaviour: Predictability, correlation, and contiguity*, ed. P. Harzem & M. D. Zeiler. Wiley. [aPRK]
- Catania, A. C., Matthews, B. A. & Shimoff, E. (1982) Instructed versus shaped human verbal behavior: Interactions with nonverbal responding. *Journal of the Experimental Analysis of Behavior* 38:233-48. [FS]
- 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* 11:327-83. [aPRK]
- Catania, A. C., Sagvolden, T. & Keller, K. J. (1988) Reinforcement schedules: Retroactive and proactive effects of reinforcers inserted into fixed-interval performances. *Journal of the Experimental Analysis of Behavior* 49:49-73. [aPRK]
- Clark, A. J. (1933) *The model of action of drugs on cells*. Edward Arnold. [CMB]
- Clark, C. W. (1976) *Mathematical bioeconomics: The optimal management of renewable resources*. Wiley. [aPRK]
- Coleman, S. (1984) Background and change in B. F. Skinner's metatheory from 1930 to 1938. *Journal of Mind and Behavior* 5:471-500. [JWD]
- Collier, G., Johnson, D. & Morgan, C. (1992) The magnitude-of-reinforcement function in closed and open economies. *Journal of the Experimental Analysis of Behavior* 57:81-89. [aPRK]
- Cronk, L. (1993) Parental favoritism toward daughters. *American Scientist* 81:272-79. [rPRK]
- Cuthill, I. C., Kacelnik, A., Krebs, J. R., Haccou, P. & Iwasa, Y. (1990) Starlings exploiting patches: The effect of recent experience on foraging decisions. *Animal Behaviour* 40:625-40. [aPRK]
- Davis, D. G. S., Staddon, J. E. R., Machado, A. & Pahner, R. G. (1993) The process of recurrent choice. *Psychological Review* 100:320-41. [aPRK]
- Davis, E. R. & Platt, J. R. (1983) Contiguity and contingency in the acquisition and maintenance of an operant. *Learning & Motivation* 14:487-512. [aPRK]
- Davis, H. & Pérusse, R. (1988) Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences* 11:561-79. [SFW]
- Davison, M. & Jenkins, P. E. (1985) Stimulus discriminability, contingency discriminability, and schedule performance. *Animal Learning & Behavior* 13:77-84. [aPRK]
- Davison, M. & McCarthy, D. (1988) *The matching law: A research review*. Erlbaum. [rPRK]
- Deadwyler, S. A. (1985) A physiological basis for hippocampal involvement in coding temporally discontiguous events. *Behavioral and Brain Sciences* 8:500-501. [aPRK]
- Denny, M. R. (1991) Relaxation/relief: The effects of removing, postponing, or terminating aversive stimuli. In: *Fear, avoidance and phobias*, ed. M. R. Denny. Erlbaum. [rPRK]
- Deutsch, D. & Deutsch, J. A. (1975) *Short-term memory*. Academic Press. [aPRK]
- de Villiers, P. A. (1977) Choice in concurrent schedules and a quantitative formulation of the law of effect. In: *Handbook of operant behavior*, ed. W. K. Honig & J. E. R. Staddon. Prentice-Hall. [CMB]

## References/Killeen: Principles of reinforcement

## References/Killeen: Principles of reinforcement

- Dickinson, A. (1985) Actions and habits: The development of behavioural autonomy. *Proceedings of the Royal Society, B* 308:67–78. [SFW]
- Donahoe, J. W., Crowley, M. A., Millard, W. J. & Stickney, K. A. (1982) A unified principle of reinforcement. In: *Quantitative analysis of behavior II: Matching and maximizing accounts*, ed. M. L. Commons, R. J. Herrnstein & H. Rachlin. Ballinger. [aPRK]
- Donahoe, J. W. & Palmer, D. C. (1993) *Learning and its implications for complex behavior*. Allyn & Bacon. [aPRK]
- Dougan, J. D. & McSweeney, F. K. (1985) Variations in Herrnstein's  $r_o$  as a function of alternative reinforcement rate. *Journal of the Experimental Analysis of Behavior* 43:215–23. [rPRK]
- Einstein, A. (1930) Cited in Pais, A. (1982) "Subtle is the Lord . . .": The science and life of Albert Einstein. Clarendon Press. [rPRK]
- Eldridge, G. D., Pear, J. J., Torgrud, L. J. & Evers, B. H. (1988) Effects of prior response-contingent reinforcement on superstitious behavior. *Animal Learning and Behavior* 16:277–84. [JJP]
- Eriksen, C. W., Webb, J. M. & Fournier, L. R. (1990) How much processing do nonattended stimuli receive? Apparently very little, but . . . *Perception and Psychophysics* 47:477–48. [FS]
- Estes, W. K. (1950) Toward a statistical theory of learning. *Psychological Review* 57:94–107. [SJH]
- Ettinger, R. H., Reid, A. K. & Staddon, J. E. R. (1987) Sensitivity to molar feedback functions: A test of molar optimality theory. *Journal of Experimental Psychology: Animal Behavior Processes* 13:366–75. [aPRK]
- Fantino, E. & Abarca, N. (1985) Choice, optimal foraging, and the delay-reduction hypothesis. *Behavioral and Brain Sciences* 8:315–62. [aPRK]
- Ferster, C. B. & Skinner, B. F. (1957) *Schedules of reinforcement*. Appleton-Century-Crofts. [JWD]
- Feynman, R. (1965) *The character of physical law*. MIT Press. [rPRK]
- Fischer, P. & Smith, W. R. (1985) Chaos, fractals and dynamics. *Lecture Notes in Pure and Applied Mathematics*, No. 98. Marcel Dekker. [RAMG]
- Fodor, J. (1989) Personal communication. Princeton, N. J. [SJH]
- Fountain, S. B. (1990) Rule abstraction, item memory, and chunking in rat serial-pattern tracking. *Journal of Experimental Psychology: Animal Behavior Processes* 16:96–105. [GG]
- Freeman, T. J. & Lattal, K. A. (1992) Stimulus control of behavioral history. *Journal of the Experimental Analysis of Behavior* 57:5–15. [KAL]
- Galbicka, G. (1988) Differentiating the behavior of organisms. *Journal of the Experimental Analysis of Behavior* 50:343–54. [aPRK]
- Galbicka, G., Kautz, M. A. & Jagers, T. (1993) Response acquisition under targeted percentile schedules: A continuing quandary for molar models of operant behavior. *Journal of the Experimental Analysis of Behavior* 60:171–84. [GG]
- Galbicka, G. & Platt, J. R. (1986) Parametric manipulation of interresponse-time contingency independent of reinforcement rate. *Journal of Experimental Psychology: Animal Behavior Processes* 12:371–80. [aPRK, GG]
- Gallistel, C. R. (1990) *The organization of learning*. MIT Press. [WT, SFW]
- Gibbon, J. (1977) Scalar expectancy and Weber's law in animal timing. *Psychological Review* 84:279–325. [aPRK]
- Gibbon, J. & Balsam, P. (1981) Spreading association in time. In: *Autoshaping and conditioning theory*, ed. J. Gibbon, C. M. Locurto & H. S. Terrace. Academic Press. [aPRK]
- Grant, D. S. & Spetch, M. L. (1991) Pigeons' memory for event duration: Differences between choice and successive matching tasks. *Learning and Motivation* 22:180–99. [DMW]
- Greenwood, M. R. C., Quartermain, D., Johnson, P. R., Cruce, J. A. F. & Hirsch, J. (1974) Food-motivated behavior in genetically obese and hypothalamic-hyperphagic rats and mice. *Physiology & Behavior* 13:687–92. [aPRK]
- Gregson, R. A. M. (1993) Learning in the context of nonlinear psychophysics: The gamma zak embedding. *British Journal of Mathematical and Statistical Psychology* 46:31–48. [RAMG]
- Grossberg, S. (1975) A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology* 18:263–327. [aPRK]
- Guthrie, E. R. (1959) Association by contiguity. In: *Psychology: A study of science*, ed. S. Koch. McGraw-Hill. [CPS]
- Hackenberg, T. D. & Axtell, S. A. M. (1993) Humans' choices in situations of time-based diminishing returns. *Journal of the Experimental Analysis of Behavior* 59:445–70. [rPRK, HR]
- Hamilton, A. L., Stellar, J. R. & Hart, E. B. (1985) Reward, performance, and the response strength method in self-stimulating rats: Validation and neuroleptics. *Physiology and Behavior* 35:897–904. [aPRK, CMB]
- Hamilton, W. D. (1964) The evolution of social behavior. *Journal of Theoretical Biology* 31:1–52. [JWD]
- Hanson, S. J. (1991) Behavioral diversity, search and stochastic connectionist systems. In: *Quantitative analysis of behavior: Neural network models of conditioning and action*, ed. M. L. Commons, S. Grossberg & J. E. R. Staddon. Erlbaum. [rPRK]
- Hanson, S. J. & Burr, D. J. (1990) What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences* 13(3):471–518. [SJH]
- Hardy, G. H., Littlewood, J. E. & Polya, G. (1959) *Inequalities*. Cambridge University Press. [rPRK]
- Hayes, N. A. & Broadbent, D. E. (1988) Two modes of learning for interactive tasks. *Cognition* 28:249–76. [FS]
- Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I. & Korn, Z. (1986) Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior* 45:237–56. [FS]
- Hayes, S. C. & Hayes, L. J. (1992) Verbal relations and the evolution of behavior analysis. *American Psychologist* 11:1183–1395. [aPRK, FS]
- Healy, A. F., Kosslyn, S. M. & Shiffrin, R. M. (1992) *From learning theory to connectionist theory: Essays in honor of William K. Estes*. Erlbaum. [rPRK]
- Herrnstein, R. J. (1969) Method and theory in the study of avoidance. *Psychological Review* 76:46–69. [SFW]
- (1970) On the law of effect. *Journal of the Experimental Analysis of Behavior* 13:243–66. [aPRK, CMB, JJP, HR, WT]
- (1974) Formal properties of the matching law. *Journal of the Experimental Analysis of Behavior* 21:159–64. [aPRK, CMB]
- (1979) Derivatives of matching. *Psychological Review* 86:486–95. [aPRK]
- Herrnstein, R. J. & Hineline, P. N. (1966) Negative reinforcement as shock frequency reduction. *Journal of the Experimental Analysis of Behavior* 9:421–30. [SFW]
- Hestenes, D. (1992) Modeling games in the Newtonian world. *American Journal of Physics* 60:732–48. [rPRK]
- Heyman, G. M. (1983) A parametric evaluation of the hedonic and motoric effects of drugs: Pimozide and amphetamine. *Journal of the Experimental Analysis of Behavior* 40:113–22. [CMB]
- Heyman, G. M. & Monaghan, M. M. (1987) Effects of changes in response requirement and deprivation on the parameters of the matching law equation: New data and review. *Journal of Experimental Psychology: Animal Behavior Processes* 13:384–94. [aPRK, CMB]
- Heyman, G. M., Monaghan, M. M. & Clody, D. E. (1987) Low doses of cis-flupenthixol attenuate motor performance. *Psychopharmacology* 93:477–82. [CMB]
- Hineline, P. H. (1976) Negative reinforcement without shock reduction. *Journal of the Experimental Analysis of Behavior* 14:259–68. [rPRK]
- Hineline, P. N. & Wanchisen, B. A. (1989) Correlated hypothesizing and the distinction between contingency-shaped and rule-governed behavior. In: *Rule-governed behavior: Cognition, contingencies, and instructional control*, ed. S. Hayes. Plenum Press. [FS]
- Hintzman, D. L. (1986) "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93:411–28. [rPRK, MDK]
- (1988) Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review* 95:528–51. [MDK]
- Honig, W. K. (1978) Studies of working memory in the pigeon. In: *Cognitive processes in animal behavior*, ed. S. H. Hulse, H. Fowler & W. K. Honig. Erlbaum. [DMW]
- Hull, C. L. (1934) The concept of habit-family hierarchy and maze learning. *Psychological Review* 41:33–54,134–52. [SFW]
- (1943) *Principles of behavior*. Appleton-Century-Crofts. [rPRK, SJH, CPS, WT, SFW, JHW]
- (1952) *A behavior system*. Yale University Press. [SFW]
- Huntley, H. D. (1958) *Dimensional analysis*. Dover Press. [DWS]
- Isaacson, E. D. & Isaacson, M. D. (1975) *Dimensional methods in engineering and physics*. Edward Arnold. [DWS]
- Jacoby, L. L. & Brooks, L. R. (1984) Nonanalytic cognition: Memory, perception, and concept formation. In: *The psychology of learning and motivation*, vol. 18, ed. G. H. Bower. Academic Press. [MDK]
- Jacoby, L. L., Lindsay, D. S. & Toth, J. P. (1992) Unconscious influences revealed: Attention, awareness, and control. *American Psychologist* 47:802–9. [FS]
- Jacoby, L. L., Toth, J. P. & Yonelinas, A. (1993) Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General* 122:139–54. [CPS]
- Jarrad, L. E. (1985) On the hippocampus, time, and interference. *Behavioral and Brain Sciences* 8:503–4. [aPRK]
- Kacelnik, A., Krebs, J. R. & Ens, B. (1987) Foraging in a changing environment: An experiment with starlings (*Sturnus vulgaris*). In:

- Quantitative analyses of behavior VI: Foraging*, ed. M. L. Commons, A. Kacelnik & S. J. Shettleworth. Erlbaum. [aPRK]
- Kapitaniak, T. (1990) *Chaos in systems with noise*, 2nd ed. World Scientific. [RAMG]
- Karpicke, J., Christoph, G., Peterson, G. & Hearst, E. (1977) Signal location and positive versus negative conditioned suppression in the rat. *Journal of Experimental Psychology: Animal Behavior Processes* 3:105–18. [MDK]
- Kelsey, J. E. & Allison, J. (1976) Fixed-ratio lever pressing by VMH rats: Work vs. accessibility of sucrose reward. *Physiology & Behavior* 17:749–54. [aPRK]
- Killeen, P. R. (1969) Reinforcement frequency and contingency as factors in fixed-ratio behavior. *Journal of the Experimental Analysis of Behavior* 12:391–95. [aPRK]
- (1981a) Averaging theory. In: *Recent developments in the quantification of steady-state operant behavior*, ed. C. M. Bradshaw, E. Szabadi & C. F. Lowe. Elsevier. [aPRK]
- (1981b) Learning as causal inference. In: *Quantitative analyses of behavior: Discriminative properties of reinforcement schedules*, ed. M. L. Commons & J. A. Nevin. Pergamon Press. [KAL]
- (1982) Incentive theory. In: *Nebraska symposium on motivation, 1981: Response structure and organization*, ed. D. J. Bernstein. University of Nebraska Press. [aPRK]
- (1984) Incentive theory III: Adaptive clocks. In: *Timing and time perception*, ed. J. Gibbon & L. Allen. New York Academy of Sciences. [arPRK]
- (1985) Incentive theory IV: Magnitude of reward. *Journal of the Experimental Analysis of Behavior* 43:407–17. [aPRK]
- (1987) Emergent behaviorism. In: *B. F. Skinner: Consensus and controversy*, ed. S. Modgil & C. Modgil. Falmer. [rPRK]
- (1988) The reflex reserve. *Journal of the Experimental Analysis of Behavior* 50:319–31. [rPRK]
- (1991) Behavior's time. In: *The psychology of learning and motivation*, ed. G. H. Bower. Academic Press. [arPRK]
- (1992) Mechanics of the animate. *Journal of the Experimental Analysis of Behavior* 57:429–63. [arPRK, RAMG]
- Killeen, P. R. & Fetterman, J. G. (1988) A behavioral theory of timing. *Psychological Review* 95:274–95. [rPRK]
- Killeen, P. R., Hanson, S. J. & Osborne, S. R. (1978) Arousal: Its genesis and manifestation as response rate. *Psychological Review* 85:571–81. [aPRK]
- Killeen, P. R. & Smith, J. P. (1984) Perception of contingency in conditioning: Scalar timing, response bias, and the erasure of memory by reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes* 10:333–45. [arPRK]
- Kintsch, W. (1965) Frequency distribution of interresponse times during VI and VR reinforcement. *Journal of the Experimental Analysis of Behavior* 8:347–52. [aPRK]
- Kraemer, P. J., Mazmanian, D. S. & Roberts, W. A. (1985) The choose-short effect in pigeon memory for stimulus duration: Subjective shortening versus coding models. *Animal Learning & Behavior* 13:349–54. [rPRK, DMW]
- Kraemer, P. J. & Roper, K. L. (1992) Matching-to-sample performance by pigeons trained with visual-duration compound samples. *Animal Learning & Behavior* 20:33–40. [rPRK]
- Frank, M. D. (1992) Alcohol outcome expectancies have selective effects on alcohol self-administration. *Alcoholism: Clinical and Experimental Research* 16(2):396. [MDK]
- Fruise, J. M., Overmier, J. B., Konz, W. A. & Rokke, E. (1983) Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. *Learning and Motivation* 14:165–81. [MDK]
- Lakatos, I. (1978) Falsification and the methodology of scientific research programs. In: *The methodology of scientific research programs: Imre Lakatos philosophical papers*, ed. J. Worrall & G. Currie. Cambridge University Press. [rPRK]
- Lattal, K. A. (1975) Reinforcement contingencies as discriminative stimuli. *Journal of the Experimental Analysis of Behavior* 23:241–46. [KAL]
- (1979) Reinforcement contingencies as discriminative stimuli: II. Effects of changes in stimulus probability. *Journal of the Experimental Analysis of Behavior* 31:15–22. [KAL]
- Lattal, K. A. & Gleeson, S. (1990) Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes* 16:27–39. [aPRK]
- Laudan, L. (1977) *Progress and its problems*. University of California Press. [rPRK]
- Levine, D. (1991) *Introduction to neural and cognitive modeling*. Erlbaum. [aPRK]
- Lieberman, D. A., Davidson, F. H. & Thomas, G. V. (1985) Making in pigeons: The role of memory in delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes* 11:611–24. [aPRK]
- Logan, J. D. (1987) *Applied mathematics: A contemporary approach*. Wiley. [DWS]
- Logue, A. W., Chavarro, A., Rachlin, H. & Reeder, R. W. (1988) Impulsiveness in pigeons living in the experimental chamber. *Animal Learning and Behavior* 16:31–40. [HR]
- Lowe, C. F. (1979) Determinants of human operant behaviour. In: *Advances in analysis of behaviour. Vol. 1: Reinforcement and the organization of behaviour*, ed. M. D. Zeiler & P. Harzem. Wiley. [FS]
- Lubinski, D. & Thompson, T. (1993) Species and individual differences in communication based on private states. *Behavioral and Brain Sciences* 16(4):627–80. [HD]
- Lyttkann, E. (1963) Statistics of regression coefficients in the case of autocorrelated residuals. In: *Time series analysis*, ed. M. Rosenblatt. Wiley. [RAMG]
- Marr, M. J. (1979) Second-order schedules of reinforcement and the generation of unitary response sequences. In: *Advances in analysis of behaviour. Vol. 1: Reinforcement and the organization of behaviour*, ed. M. D. Zeiler & P. Harzem. Wiley. [PR]
- Mazur, J. E. (1982) A molecular approach to ratio schedule performance. In: *Quantitative analyses of behavior. Vol. 2: Matching and maximizing accounts*, ed. M. L. Commons, R. J. Herrnstein & H. Rachlin. Ballinger. [aPRK]
- (1983) Steady-state performance on fixed-, mixed-, and random-ratio schedules. *Journal of the Experimental Analysis of Behavior* 39:293–307. [aPRK]
- (1984) Tests of an equivalence rule for fixed and variable delays. *Journals of Experimental Psychology: Animal Behavior Processes* 10:426–36. [aPRK]
- McDowell, J. J. (1980) An analytical comparison of Herrnstein's equations and a multivariate rate equation. *Journal of the Experimental Analysis of Behavior* 33:397–408. [aPRK]
- McDowell, J. J., Bass, R. & Kessel, R. (1992) Applying linear systems analysis to dynamic behavior. *Journal of the Experimental Analysis of Behavior* 57:377–91. [aPRK]
- McDowell, J. J. & Kessel, R. (1979) A multivariate rate equation for variable-interval performance. *Journal of the Experimental Analysis of Behavior* 31:267–83. [aPRK]
- McDowell, J. J. & Wixted, J. T. (1986) Variable-ratio schedules as variable-interval schedules with linear feedback loops. *Journal of the Experimental Analysis of Behavior* 46:315–29. [arPRK, JHW]
- McDowell, J. J. & Wood, H. (1984) Confirmation of linear system theory prediction: Changes in Herrnstein's  $k$  as a function of changes in reinforcer magnitude. *Journal of the Experimental Analysis of Behavior* 41:183–92. [aPRK]
- McNamara, J. M. & Houston, A. I. (1987) Memory and the efficient use of information. *Journal of Theoretical Biology* 125:385–95. [aPRK]
- McSweeney, F. K. (1974) Variability of responding on a concurrent schedule as a function of body weight. *Journal of the Experimental Analysis of Behavior* 21:357–59. [aPRK]
- (1978) Prediction of concurrent keypeck treadle-press responding from simple schedule performance. *Animal Learning & Behavior* 6:444–50. [aPRK]
- (1992) Rate of reinforcement and session duration as determinants of within-session patterns of responding. *Animal Learning & Behavior* 20:160–69. [arPRK]
- Mechner, F. (1958) Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior* 1:109–21. [SFW]
- Medin, D. L. & Schaffer, M. M. (1978) Context theory of classification learning. *Psychological Review* 85:207–38. [MDK]
- Mesterton-Gibbons, M. (1989) *A concrete approach to mathematical modelling*. Addison-Wesley. [aPRK]
- Minsky, M. (1961) Steps toward artificial intelligence. *Proceedings of the Institute of Radio Engineers* 49:8–30. [CDLW]
- Moore, J. W. (1991) Implementing connectionist algorithms for classical conditioning in the brain. In: *Neural network models of conditioning and action: A volume in the quantitative analysis of behavior series*, ed. M. L. Commons, S. Grossberg & J. E. R. Staddon. Erlbaum. [aPRK]
- Morley, M. J., Shah, K., Bradshaw, C. M. & Szabadi, E. (1988) DSP4 and Herrnstein's equation: Further evidence for a role of noradrenaline in the maintenance of operant behaviour by positive reinforcement. *Psychopharmacology* 96:551–56. [CMB]
- Navon, D. & Gopher, D. (1979) On the economy of the human-processing system. *Psychological Review* 86:214–55. [FS]
- Neimark, E. D. & Estes, W. K. (1967) *Stimulus sampling theory*. Holden-Day. [rPRK]

## References/Killeen: Principles of reinforcement

- Neuringer, A. (1992) Choosing to vary and repeat. *Psychological Science* 3:246–50. [aPRK]
- Neuringer, A. J. & Chung, S.-H. (1967) Quasi-reinforcement: Control of responding by a percentage-reinforcement schedule. *Journal of the Experimental Analysis of Behavior* 10:45–54. [rPRK]
- Nevin, J. A. (1992a) Behavioral contrast and behavioral momentum. *Journal of Experimental Psychology: Animal Behavior Processes* 18:126–33. [JAN]
- (1992b) An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior* 57:301–16. [JAN]
- Norman, D. A. (1970) *Models of human memory*. Academic Press. [aPRK]
- Olton, D. S. (1979) Mazes, maps, and memory. *American Psychologist* 34:583–96. [SFW]
- Olton, D. S. & Samuelson, R. J. (1976) Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes* 2:97–116. [SFW]
- Overmier, J. B. & Lawry, J. A. (1979) Pavlovian conditioning and the mediation of behavior. In: *The psychology of learning and motivation*, vol. 13, ed. G. H. Bower. Academic Press. [MDK]
- Page, S. & Neuringer, A. (1985) Variability as an operant. *Journal of Experimental Psychology: Animal Behavior Processes* 11:429–52. [KAL]
- Palmer, D. C. & Donahoe, J. W. (1992) Essentialism and selectionism in cognitive science and behavior analysis. *American Psychologist* 47:1344–58. [aPRK, KAL]
- Palya, W. L. (1992) Dynamics in the fine structure of schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior* 57:267–87. [aPRK]
- Papini, M. R. & Bitterman, M. E. (1990) The role of contingency in classical conditioning. *Psychological Review* 97:396–403. [aPRK]
- Pear, J. J. (1975) Implications of the matching law for ratio responding. *Journal of the Experimental Analysis of Behavior* 23:139–40. [aPRK]
- (1985) Spatiotemporal patterns of behavior produced by variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior* 44:217–31. [rPRK, JJP]
- (1988) Behavioral stereotypy and the generalized matching equation. *Journal of the Experimental Analysis of Behavior* 50:87–95. [JJP]
- Pearce, J. M. (1987) A model for stimulus generalization in Pavlovian conditioning. *Psychological Review* 94:61–73. [MDK]
- Peterson, L. R. & Peterson, M. J. (1959) Short-term retention of individual items. *Journal of Experimental Psychology* 58:193–98. [aPRK]
- Platt, J. R. (1971) Discrete trials and their relation to free-operant behavior. In: *Essays in neobehaviorism: A memorial volume to Kenneth W. Spence*, ed. H. H. Kendler & J. T. Spence. Appleton-Century-Crofts. [aPRK]
- (1973) Percentile reinforcement: Paradigms for experimental analysis of response shaping. In: *The psychology of learning and motivation: Advances in research and theory*, ed. G. H. Bower. Academic Press. [aPRK]
- (1979) Interresponse-time shaping by variable-interval-like interresponse-time reinforcement. *Journal of the Experimental Analysis of Behavior* 31:3–14. [aPRK, MDK]
- Platt, J. R. & Day, R. B. (1979) A hierarchical response-unit analysis of resistance to extinction following fixed-number and fixed-consecutive-number reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes* 5:307–20. [aPRK, MDK]
- Powell, R. W. (1968) The effect of small sequential changes in fixed-ratio size upon the post-reinforcement pause. *Journal of the Experimental Analysis of Behavior* 11:589–93. [aPRK, DWS]
- Preston, R. A. & Fantino, E. (1991) Conditioned reinforcement value and choice. *Journal of the Experimental Analysis of Behavior* 55:155–75. [aPRK]
- Rachlin, H., Raineri, A. & Cross, D. (1991) Subjective probability and delay. *Journal of the Experimental Analysis of Behavior* 55:233–44. [aPRK]
- Rawlins, J. N. P. (1985) Associations across time: The hippocampus as a temporary memory store. *Behavioral and Brain Sciences* 8:479–96. [aPRK]
- Reed, P. (1992) Effect of local context of responding on human judgment of causality. *Memory & Cognition* 20:573–79. [PR]
- Reed, P., Schachtman, T. R. & Hall, G. (1991) Effect of signaled reinforcement on the formation of behavioral units. *Journal of Experimental Psychology: Animal Behavior Processes* 17:475–85. [arPRK, PR]
- Rescorla, R. A. (1990a) Evidence for an association between the discriminative stimulus and the response-outcome association in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes* 16:326–34. [MDK]
- (1990b) The role of information about the response-outcome relationship in instrumental discrimination learning. *Journal of Experimental Psychology: Animal Behavior Processes* 16:262–70. [MDK, SFW]
- Rescorla, R. A. & Solomon, R. L. (1967) Two-process learning theory: Relationship between Pavlovian and instrumental learning. *Psychological Review* 74:151–82. [MDK]
- Rescorla, R. A. & Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: Current research and theory*, ed. A. H. Black & W. F. Prokasy. Appleton-Century-Crofts. [rPRK, SJH, WT]
- Revusky, S. & Garcia, J. (1970) Learned associations over long delays. In: *The psychology of learning and motivation: Advances in research and theory*, ed. G. H. Bower. Academic Press. [aPRK]
- Reynolds, G. S. (1968) *A primer of operant conditioning*. Scott, Foresman. [JWD]
- Reynolds, G. S. & McLeod, A. (1970) On the theory of interresponse-time reinforcement. In: *The psychology of learning and motivation: Advances in research and theory*, ed. G. H. Bower. Academic Press. [aPRK]
- Richards, R. W. (1981) A comparison of signaled and unsignaled delay of reinforcement. *Journal of the Experimental Analysis of Behavior* 35:145–52. [aPRK]
- Richardson, W. K. & Loughead, T. E. (1974) The effect of physical restraint on behaviour under the differential-reinforcement-of-low-rate schedule. *Journal of the Experimental Analysis of Behavior* 21:455–61. [GG]
- Robbins, T. W. & Everitt, B. J. (1992) Functions of dopamine in the dorsal and ventral striatum. *Seminars in the Neurosciences* 4:119–27. [aPRK]
- Robbins, T. W. & Sahakian, B. J. (1983) Behavioral effects of psychomotor stimulant drugs: Clinical and neuropsychological implications. In: *Stimulants: Neurochemical, behavioral, and clinical perspectives*, ed. I. Creese. Raven Press. [aPRK]
- Roberts, W. A. & Grant, D. S. (1974) Short-term memory in the pigeon with presentation time precisely controlled. *Learning & Motivation* 5:393–408. [rPRK]
- Ross, S. M. (1970) *Applied probability models with optimization applications*. Holden-Day. [aPRK]
- Rumelhart, D. E., Hinton, G. E. & Williams, R. (1986) Learning internal representations by error propagation. *Nature* 323:533–36. [SJH]
- Sagvolden, T., Hendley, E. D. & Knardahl, S. (1992) Behavior of hypertensive and hyperactive rat strains: Hyperactivity is not unitarily determined. *Physiology and Behavior* 52:49–57. [MAM]
- Sagvolden, T., Metzger, M. A. & Sagvolden, G. (1993) Frequent reward eliminates differences in activity between hyperkinetic rats and controls. *Behavioral and Neural Biology* 59:225–29. [MAM]
- Sakamoto, Y., Ishiguro, M. & Kitagawa, G. (1986) *Akaike information criterion statistics*. Reidel. [aPRK]
- Santi, A., Bridson, S. & Ducharme, M. J. (1993) Memory codes for temporal and nontemporal samples in many-to-one matching by pigeons. *Animal Learning & Behavior* 21:120–30. [rPRK]
- Schacter, D. L. (1987) Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:501–18. [CPS]
- Schlanger, H. D., Jr. (1992) Theory in behavior analysis: An application to child development. *American Psychologist* 47:1396–1410. [KAL]
- Schoenfeld, W. N. (1972) Problems of modern behavior theory. *Conditional Reflex* 7:33–65. [rPRK]
- Schwartz, B. (1984) Creation of stereotyped, functional behavioral units. In: *Quantitative analyses of behavior*, Vol. IV: *Discrimination processes*, ed. M. L. Commons, R. J. Herrnstein & A. R. Wagner. Ballinger. [PR]
- Seligman, M. E. P. (1970) On the generality of the laws of learning. *Psychological Review* 77:406–18. [HD]
- Shah, K., Bradshaw, C. M. & Szabadi, E. (1991) Relative and absolute reinforcement frequency as determinants of choice in concurrent variable interval schedules. *Quarterly Journal of Experimental Psychology* 43B:25–38. [CMB]
- Shanks, D. R. & Dickinson, A. (1991) Instrumental judgement and performance under variations in action-outcome contingency and contiguity. *Memory and Cognition* 19:353–60. [FS]
- Shettleworth, D. J. (1975) Reinforcement and the organization of behavior in golden hamsters: Hunger, environment, and food reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes* 1:56–87. [JHW]
- Shuttleworth, S. J. & Krebs, J. R. (1982) How marsh tits find their hoards: The roles of site preference and spatial memory. *Journal of Experimental Psychology: Animal Behavior Processes* 8:354–75. [SFW]
- Shimp, C. P. (1973) Synthetic variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior* 19:311–30. [aPRK]
- (1976a) Organization in memory and behavior. *Journal of the Experimental Analysis of Behavior* 26:113–30. [aPRK, JJP]
- (1976b) Short-term memory in the pigeon: Relative recency. *Journal of the Experimental Analysis of Behavior* 25:55–56. [aPRK, JJP]
- (1976c) Short-term memory in the pigeon: The previously reinforced response. *Journal of the Experimental Analysis of Behavior* 26:487–93. [aPRK]

## References/Killeen: Principles of reinforcement

- (1978) Memory, temporal discrimination, and learned structure in behavior. In: *The psychology of learning and motivation*, vol. 12, ed. G. H. Bower. Academic Press. [CPS]
- (1982) Metaknowledge in the pigeon: An organism's knowledge about its own behavior. *Animal Learning and Behavior* 10(3):358–64. [CPS]
- (1983) The local organization of behavior: Dissociations between a pigeon's behavior and self-reports of that behavior. *Journal of the Experimental Analysis of Behavior* 39:61–68. [CPS]
- (1984) Self-reports by rats of the temporal patterning of their behavior: A dissociation between tacit knowledge and knowledge. In: *Animal cognition*, ed. H. L. Roitblat, T. G. Bever & H. S. Terrace. Erlbaum. [CPS]
- (1989) Contemporary behaviorism versus the old behavioral strawman in Gardner's *The mind's new science: A history of the cognitive revolution*. *Journal of the Experimental Analysis of Behavior* 51:163–71. [CPS]
- Shimp, C. P., Childers, L. J. & Hightower, F. A. (1990) Local patterns in human operant behavior and a behaving model to interrelate animal and human performances. *Journal of Experimental Psychology: Animal Behavior Processes* 16:200–212. [CPS]
- Shimp, C. P., Sabulsky, S. L. & Childers, L. J. (1989) Preference for starting and finishing behavior patterns. *Journal of the Experimental Analysis of Behavior* 53:341–52. [CPS]
- Sidman, M. (1954) The temporal distribution of avoidance responses. *Journal of Comparative and Physiological Psychology* 47:399–402. [SFW]
- (1962) An adjusting avoidance schedule. *Journal of the Experimental Analysis of Behavior* 5:271–77. [SFW]
- Sizemore, O. J. & Lattal, K. A. (1978) Unsigned delay of reinforcement in variable-interval schedules. *Journal of the Experimental Analysis of Behavior* 30:169–75. [aPRK]
- Skinner, B. F. (1935) The generic nature of the concepts of stimulus and response. *Journal of General Psychology* 12:40–65. [aPRK]
- (1938) *The behavior of organisms*. Appleton-Century-Crofts. [aPRK, JWD, KAL]
- (1950) Are theories of learning necessary? *Psychological Review* 52:270–77. [KAL]
- (1957) *Verbal behavior*. Appleton-Century-Crofts. [RLS]
- (1969) *Contingencies of reinforcement. A theoretical analysis*. Prentice-Hall. [FS]
- (1974) *About behaviorism*. Knopf. [KAL]
- Smith, E. (1991) Illusory correlation in a simulated exemplar-based memory. *Journal of Experimental Social Psychology* 27:107–23. [MDK]
- Smith, L. D. (1992) On prediction and control: B. F. Skinner and the technological ideal of science. *American Psychologist* 47:216–23. [rPRK]
- Snyderman, M. (1983) Body weight and response strength. *Behaviour Analysis Letters* 3:255–65. [aPRK, CMB]
- Sober, E. (1984) *The nature of selection: Evolutionary theory in philosophical focus*. MIT Press. [JWD]
- Spetch, M. L. & Sinha, S. S. (1989) Proactive effects in pigeons' memory for event duration: Evidence for analogical retention. *Journal of Experimental Psychology: Animal Behavior Processes* 15:347–57. [rPRK, DMW]
- Spetch, M. L. & Wolke, D. M. (1983) Subjective shortening: A model of pigeons' memory for event duration. *Journal of Experimental Psychology: Animal Behavior Processes* 9:14–30. [rPRK]
- Staddon, J. E. R. (1977) On Herrnstein's equation and related forms. *Journal of the Experimental Analysis of Behavior* 28:163–70. [aPRK]
- (1980) Optimality analyses of operant behavior and their relation to optimal foraging. In: *Limits to action: The allocation of individual behavior*, ed. J. E. R. Staddon. Academic Press. [WT]
- (in press) *Behaviorism: Mind, mechanism and society*. Duckworth. [HR]
- Staddon, J. E. R. & Zhang, Y. (1989) Response selection in operant learning. *Behavioural Processes* 20:189–97. [aPRK]
- Stein, L. & Belluzzi, J. D. (1989) Cellular investigations on behavioral reinforcement. *Neuroscience and Biobehavioral Reviews* 13:69–80. [JWD]
- Stephens, D. W. & Dunbar, S. R. (1993) Dimensional analysis in behavioral ecology. *Behavioral Ecology* 4:172–83. [DWS]
- Stokes, P. D. & Balsam, P. D. (1991) Effects of reinforcing preselected approximations on the topography of the rats' bar press. *Journal of the Experimental Analysis of Behavior* 55:213–31. [aPRK]
- Svardal, F. (1991) Operant modulation of low-level attributes of rule-governed behavior by nonverbal contingencies. *Learning and Motivation* 22:406–20. [rPRK, FS]
- (1992) Sensitivity to nonverbal operant contingencies: Do limited processes resources affect operant conditioning in humans? *Learning and Motivation* 23:383–405. [rPRK, FS]
- (1993) Working harder for less: Effect of incentive value on force of instrumental response in humans. *Quarterly Journal of Experimental Psychology* 46A:11–34. [JHW]
- Terrace, H. S. (1991) Chunking during serial learning by the pigeon: I. Basic evidence. *Journal of Experimental Psychology: Animal Behavior Processes* 17:79–91. [GG]
- Thomas, G. V. (1983) Contiguity and contingency in instrumental conditioning. *Learning and Motivation* 14:513–26. [aPRK]
- Thorndike, E. L. (1898) Animal intelligence: An experimental study of the association processes in animals. *Psychological Review Monograph 2* (whole number 8). [CDLW]
- Timberlake, W. (1983) Rats' responses to a moving object related to food or water: A behavior-systems analysis. *Animal Learning & Behavior* 11:309–20. [aPRK]
- (1993a) Animal behavior: A continuing synthesis. *Annual Review of Psychology* 44:675–708. [WT]
- (1993b) Behavior systems and reinforcement: An integrative approach. *Journal of the Experimental Analysis of Behavior* 60:105–28. [WT]
- Timberlake, W. & Lucas, G. A. (1989) Behavior systems and learning: From misbehavior to general principles. In: *Contemporary learning theories: Instrumental conditioning theory and the impact of biological constraints on learning*, ed. S. B. Klein & R. R. Mowrer. Erlbaum. [aPRK, WT]
- Timberlake, W. & Peden, B. F. (1987) On the distinction between open and closed economies. *Journal of the Experimental Analysis of Behavior* 48:35–60. [aPRK]
- Tolman, E. C. (1932) *Purposive behavior in animals and man*. Century. [SFW]
- (1948) Cognitive maps in rats and men. *Psychological Review* 55:189–208. [SFW]
- van Hest, A., van Haaren, F. & van de Poll, N. E. (1987) Behavioral differences between male and female Wistar rats on DRL schedules: Effect of stimuli promoting collateral activities. *Physiology & Behavior* 39:255–61. [GC]
- Wagner, A. R. (1981) SOP: A model of automatic memory processing in animal behavior. In: *Information processing in animals: Memory mechanisms*, ed. N. E. Spear & R. R. Miller. Erlbaum. [rPRK]
- Wagner, A. R. & Brandon, S. E. (1989) Evolution of a structured connectionist model of Pavlovian conditioning. In: *Contemporary learning theories: Pavlovian conditioning and the status of learning theory*, ed. B. Klein & R. R. Mowrer. Erlbaum. [rPRK]
- Wasserman, E. A., Elek, S. M., Chatlosh, D. L. & Baker, A. G. (1993) Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:174–88. [FS]
- Watkins, M. J. (1990) Mediationism and the obfuscation of memory. *American Psychologist* 45:328–35. [KAL]
- Wearden, J. H. (1989) Mysteries of the organism: Clark L. Hull's *Principles of behavior* and some problems in contemporary schedule theory. *Journal of the Experimental Analysis of Behavior* 51:277–86. [JHW]
- Wearden, J. H. & Clark, R. B. (1988) Interresponse-time reinforcement and behavior under aperiodic reinforcement schedules: A case study using computer modeling. *Journal of Experimental Psychology: Animal Behavior Processes* 14:200–211. [JHW]
- (1989) Constraints on the process of interresponse-time reinforcement as the explanation of variable-interval performance. *Behavioural Processes* 20:151–75. [aPRK, JHW]
- Wearden, J. H. & Ferrara, A. (1993) Subjective shortening in humans' memory for stimulus duration. *Quarterly Journal of Experimental Psychology* 46B:163–86. [rPRK]
- White, K. G., McCarthy, D. & Fantino, E. (1989) Cognition and behavior analysis. *Journal of the Experimental Analysis of Behavior* 52:197–98. [rPRK]
- Whittlesea, B. W. A. (1987) Preservation of specific experiences in general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:3–17. [MDK]
- Whittlesea, B. W. A. & Brooks, L. R. (1988) Critical influence of particular experiences in the perception of letters, words, and phrases. *Memory and Cognition* 16:387–99. [MDK]
- Whittlesea, B. W. A. & Dorkin, M. D. (1993) Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General* 122(2):227–48. [MDK]
- Wilkenfield, J., Nickel, M., Blakely, E. & Ploing, A. (1992) Acquisition of lever-press responding in rats with delayed reinforcement: A comparison of three procedures. *Journal of the Experimental Analysis of Behavior* 58:431–43. [aPRK]
- Wilkie, D. M. & Willson, R. J. (1990) Discriminal distance analysis supports the hypothesis that pigeons retrospectively encode event duration. *Animal Learning & Behavior* 18:124–32. [DMW]
- Williams, B. A. (1972) Probability learning as a function of momentary

## References/Killeen: Principles of reinforcement

- reinforcement probability. *Journal of the Experimental Analysis of Behavior* 17:363-68. [aPRK]
- (1975) The blocking of reinforcement control. *Journal of the Experimental Analysis of Behavior* 24:215-26. [aPRK]
- (1976) The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior* 26:441-49. [aPRK]
- (1978) Information effects on the response-reinforcer association. *Animal Learning & Behavior* 6:371-79. [aPRK]
- (1988) Reinforcement, choice, and response strength. In: *Stevens' handbook of experimental psychology*. Vol. 2: *Learning and cognition*, 2nd ed., ed. R. C. Atkinson, R. J. Herrnstein, G. Lindzey & R. D. Luce. Wiley. [aPRK, JAN]
- Williams, D. C. & Johnston, J. M. (1992) Continuous versus discrete dimensions of reinforcement schedules: An integrative analysis. *Journal of the Experimental Analysis of Behavior* 58:205-28. [aPRK]
- Wixted, J. T. (1989) Nonhuman short-term memory: A quantitative analysis of selected findings. *Journal of the Experimental Analysis of Behavior* 52:409-26. [rPRK]
- Wogar, M. A., Bradshaw, C. M. & Szabadi, E. (1991) Evidence for an involvement of 5-hydroxytryptaminergic neurones in the maintenance of operant behaviour by positive reinforcement. *Psychopharmacology* 105:119-24. [CMB]
- Zeiler, M. D. (1979) Output dynamics. In: *Advances in analysis of behaviour*. Vol. 1: *Reinforcement and the organization of behaviour*, ed. M. D. Zeiler & P. Harzem. Wiley. [aPRK]
- (1991) Ecological influences on timing. *Journal of Experimental Psychology: Animal Behavior Processes* 17:13-25. [aPRK]
- Zeiler, M. D. & Buchman, I. B. (1979) Response requirements as constraints on output. *Journal of the Experimental Analysis of Behavior* 32:29-50. [aPRK]
- Zeiler, M. D. & Thompson, T. (1986) *Analysis and integration of behavioral units*. Erlbaum. [aPRK]
- Zuriff, G. E. (1970) A comparison of variable-ratio and variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior* 12:369-74. [aPRK]
- (1985) Behaviorism: A conceptual reconstruction. Columbia University Press. [KAL]

## THE CHILD'S DISCOVERY OF THE MIND

Janet Wilde Astington

Jean Piaget investigated children's discovery of the mind in the 1920s and concluded that they had little understanding before the age of six. But over the last twenty years, researchers have begun to challenge his methods and revise his conclusions. In *The Child's Discovery of the Mind*, Astington provides a valuable overview of current research in this lively area and of the consequences of this discovery for intellectual and social development.

Developing Child Series \$24.95 cloth \$9.95 paper



## VARIETIES OF MORAL PERSONALITY

Ethics and Psychological Realism

Owen Flanagan

"Flanagan's book prepares a framework for the balanced interchange between empirical studies and their varied interpretations. Rigorous in its search, the value of this book lies in its capacity to ferret out the conditional bonds between human psychology and ethical expectations."

—E. Mark Stern, *Contemporary Psychology*

"This is a rich and elegantly written book...In Flanagan's writing, no view is dismissed out of hand, none is taken over uncritically—thereby exemplifying the same intellectual responsibility he advocates for moral psychology in general."

—Thomas E. Wren, *Journal of Moral Education*

\$16.95 paper

Available at bookstores or from  
**Harvard University Press**

Cambridge, MA 02138 Order toll-free: 800-448-2242