JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1990, 53, 189-200 NUMBER 1 (JANUARY)

UNIFICATION OF MODELS FOR CHOICE BETWEEN DELAYED REINFORCERS

PETER R. KILLEEN AND EDMUND FANTINO

ARIZONA STATE UNIVERSITY AND UNIVERSITY OF CALIFORNIA, SAN DIEGO

Two models for choice between delayed reinforcers, Fantino's delay-reduction theory and Killeen's incentive theory, are reviewed. Incentive theory is amended to incorporate the effects of arousal on alternate types of behavior that might block the reinforcement of the target behavior. This amended version is shown to differ from the delay-reduction theory in a term that is an exponential in incentive theory and a difference in delay-reduction theory. A power series approximation to the exponential generates a model that is formally identical with delay-reduction theory. Correlations between delay-reduction theory and the amended incentive theory show excellent congruence over a range of experimental conditions. Although the assumptions that gave rise to delay-reduction theory and incentive theory remain different and testable, the models deriving from the theories are unlikely to be discriminable by parametric experimental tests. This congruence of the models is recognized by naming the common model the delayed reinforcement model, which is then compared with other models of choice such as Killeen and Fetterman's (1988) behavioral theory of timing, Mazur's (1984) equivalence rule, and Vaughan's (1985) melioration theory.

Key words: choice, concurrent chains, models, delay-reduction theory, incentive theory, delayed reinforcement

The introduction of a second operandum into experimental chambers has more than doubled the richness of data that issue from them. Providing animals access to concurrent schedules of reinforcement permits experimenters to study not only the effects of two schedules but also the interaction between them. Just as gravity cannot be studied with only a single body, schedule control cannot be studied with only a single reinforcement schedule. This is because there are always implicit alternatives to the reinforcers we offer animals, and these alternatives will control behavior whether or not they are under our control. The most successful model of single-schedule control is a model of concurrent-schedule control, with the alternative schedule entering as a hypothetical construct (Herrnstein, 1970).

Thus, to simplify and strengthen our understanding of control by reinforcement we must weaken the control exerted by a single schedule by introducing an explicit alternative that is under our control. Although this does not remove the opportunity for still other implicit sources of control, it does provide an introduction to the nature of interactions between schedules; in turn that improves our ability to model the nature of control by the other, implicit reinforcers.

The nature of interactions between concurrent reinforcement schedules depends critically on the nature of the contingencies between the schedules and the behavior they control, as well as the nature of contingencies that govern switching between the schedules. The recent literature analyzing these relations is rich and steeped in theory (see Davison & McCarthy, 1988, for reviews).

It is possible to use concurrent schedules in such a way that they provide valuable information on the nature of the control of behavior by reinforcement without first resolving many of the controversies surrounding choice. We may keep the concurrent schedules identical and study the relation between behavior allocation and the nature of the alternatives chosen. For instance, when the magnitude of a reward is increased, preference for it increases, although not proportionately (Killeen, 1985). What happens when the alternatives chosen are not primary reinforcers, but rather are themselves schedules of reward? There are a number of approaches that have attempted to unify the empirical effects of manipulation of

---

Research and preparation of manuscript supported in part by NIMH Grant MH-20752 and NSF Grant BNS 83-2963 to the University of California at San Diego and NIMH Grant MH39496 to Arizona State University. Reprints may be obtained from Peter R. Killeen, Department of Psychology, Arizona State University, Tempe, Arizona, 85287-1104 or from Edmund Fantino, Department of Psychology, C-009, University of California at San Diego, La Jolla, California 92093-0109.

such terminal-link schedules; this paper attempts to unify two of those approaches.

LAWS, THEORIES, AND MODELS

The attempt to unify various approaches must first come to terms with the nature of the things to be unified. To that end, we indulge a brief digression on the issue of theories, models, and laws. We make categorical distinctions while recognizing that in practice the categories blend. A *theory* is a way of viewing a set of phenomena. It involves commitments to certain assumptions and world views, selection of a data base, agreement as to which questions are relevant, and, as the theory grows, it includes a corpus of data and empirical laws. It also involves *models,* some verbal, some physical, some graphical, and some mathematical. In all cases, the models are ways of conveying the thrust of theoretical assumptions and converting them into specific predictions. *Empirical laws* are statements of regularities in the data, cast in a language that does not rely heavily upon the assumptions of theories that hope to incorporate them. Theories are neither better nor worse than models. Sometimes hypotheses and speculations are incorrectly dignified with the title *theory*; sometimes theories grow out of such speculations (top-down), and sometimes they grow out of empirical laws that are closer to the data (bottom-up). Optimal foraging theory and signal detectability theory are examples of the former, and reinforcement theory is the prime example of the latter.

Theories, models, and laws may be tentative or well confirmed. Ideally, this status should be identified by an independent modifier. Unfortunately, this has not happened, and each of the basic terms has picked up a different connotation concerning degree of confirmation. Laws are presumed intrinsically complete and permanent (Ohm's law, laws of nature); models are things that are always being tinkered with (model cars, model ships); theories are somewhere between the two, equivocally profound and permanent or specious and ephemeral (theory of relativity, "just a" theory).

The approaches to understanding choice between reinforcement schedules include tentative laws (Mazur, 1984), tentative theories (Fantino, 1977), and tentative laws growing into theories (Killeen, 1982a, 1982b). In all these cases, a mathematical treatment of the data is the core of the approach. For convenience the mathematical expressions will be referred to as models, although there are important differences among the models, depending on the context from which they issued. Unification of laws has few implications, other than the increased power or simplicity that results. Unification of theoretical models begs the question of which theoretical assumptions get carried forward with the unified model; it is not possible to unify theories the way models are unified, unless some basic assumptions are changed.

DESCRIPTION OF THE MODELS

Delay-Reduction Theory (DRT)

The delay-reduction theory (DRT) states that effectiveness of a stimulus as a conditioned reinforcer may be predicted most accurately by calculating the reduction in the length of time to primary reinforcement correlated with the onset of the stimulus in question relative to the length of time to primary reinforcement measured from the onset of the preceding stimulus (Fantino, 1969, 1977, 1981; Fantino & Davison, 1983; Squires & Fantino, 1971). The simplest form of the parameter-free DRT may be stated as:

$$\text{Reinforcing strength of stimulus } A = f\left(\frac{T - t_A}{T}\right) \qquad (1)$$

where t_A is the temporal interval between the onset of stimulus A and primary reinforcement, and T is the total time between reinforcer presentations. Expressed differently, the greater the improvement, in terms of temporal proximity or waiting time to reinforcement, that is correlated with the onset of a stimulus, the more effective that stimulus will be as a conditioned reinforcer. Although the hypothesis has been extended to areas such as self-control (Ito & Asaki, 1982; Navarick & Fantino, 1976), elicited responding (Fantino, 1982), three-alternative choice (Fantino & Dunn, 1983), observing (Case & Fantino, 1981; Fantino & Case, 1983), the serial-position effect in short-term memory (Wixted, 1989), and operant analogues to foraging (Abarca & Fantino, 1982; Fantino & Abarca, 1985), the hypothesis was first developed to account for choice for two variable-interval (VI) schedules of reinforcement in the concurrent-chains procedure developed by Autor (1960, 1969) and Herrnstein (1964).

Assume the subject is choosing between two reward outcomes, one involving an average delay of t_1, the other an average of t_2. The subject produces the two outcomes by responding in the choice phase (concurrent white lights) of the procedure (Figure 1). Responses in the presence of the concurrently available white lights lead to the colored lights on the left and right associated with t_1 and t_2, respectively, according to VI schedules of reinforcement. The independent variable is typically some difference between t_1 and t_2 or some difference in the food that generally occurs after t_1 or t_2 (e.g., the amount of food). The dependent variable, the measure of choice, is generally the relative rate of responding in the presence of the concurrently available white lights; for example, the number of responses made to the left white light (B_1) divided by the number of responses made to both white lights ($B_1 + B_2$).

Numerous studies have shown that the strength of preference for one alternative is a function of the length of the choice phase (see Fantino, 1977, for a review), a finding that follows from the DRT. For example, we may calculate how far in time from primary reinforcement the organism is during the choice phase (or initial links of the concurrent-chains procedure), and we already know how far reinforcement is when the left and right terminal links are entered: the values of t_1 and t_2, respectively. Let T represent the average scheduled overall time to primary reinforcement from the outset of the choice phase. T equals the average time in the choice phase plus the average time in the outcome (or terminal-link) phase. For example, when the equal VI schedules in the choice phase are VI 120-s schedules, the mean time in the choice phase is 60 s, because the two VI 120-s schedules operate simultaneously. If the VI schedules associated with the terminal links are VI 30 s and VI 90 s (i.e., if t_1 = 30 s and t_2 = 90 s), then the average time in the outcome phase also equals 60 s (the mean of 30 s and 90 s, the two equiprobable outcome times). Thus, T equals 60 s plus 60 s or 120 s. The delay reduction correlated with the onset of t_1 (entering the terminal link on the left) then equals $(T - t_1)$; similarly, the delay reduction correlated with the onset of t_2 equals $(T - t_2)$. Fantino (1969) predicted that the following equation should describe choice:

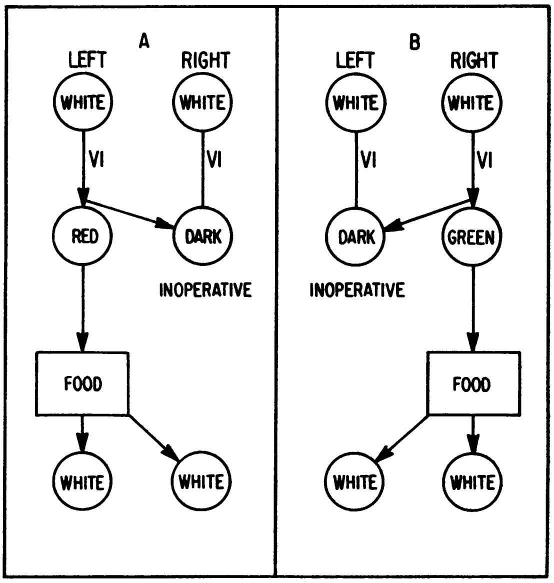

Fig. 1. The concurrent-chains procedure for the study of choice. **A** (left) depicts the sequence of events when responses on the left key are reinforced; **B** (right) depicts the analogous sequence on the right key. Responses in the presence of the colored lights (the stimuli of the outcome phase) are reinforced with food according to some schedule of reinforcement (generally the independent variable; the schedule is often a VI schedule, as shown). The measure of choice is the relative rate of responding in the presence of the concurrently available white lights (number of responses to one white light divided by the sum of responses to both). Typically, equal VI schedules arrange access to the outcome schedules. After Fantino (1969); copyright 1969 by the Society for the Experimental Analysis of Behavior, Inc.

$$\begin{aligned}\frac{B_1}{B_1 + B_2} &= \frac{T - t_1}{(T - t_1) + (T - t_2)} \\ &\quad (\text{when } t_1 < T, t_2 < T) \\ &= 1 \ (\text{when } t_1 < T, t_2 > T) \\ &= 0 \ (\text{when } t_1 > T, t_2 < T). \qquad (2)\end{aligned}$$

Thus, in the example above, because T = 120 s, Equation 2 requires a choice proportion of .75 [because $T - t_1 = 90$, $T - t_2 = 30$ and $90/(90 + 30) = .75$]. With longer or shorter equal VIs in the choice phase, the choice proportions required by Equation 2 vary systematically: The shorter the choice duration the greater the preference. Note that when either outcome represents an increase in average delay to reinforcement (either $t_1 > T$ or $t_2 > T$), Equation 2 requires the subject to respond exclusively for the other outcome. That is, Equation 2 specifies when one outcome should be chosen exclusively.

Note also that t_1 and t_2 cannot *both* be greater than T. Equation 2 has been modified by Squires and Fantino (1971) for the special case of different initial-link durations by weighting the terms in the right-hand side by the overall food rate scheduled on each key so that:

$$\begin{aligned}\frac{B_1}{B_1 + B_2} &= \frac{r_1(T - t_1)}{r_1(T - t_1) + r_2(T - t_2)} \\ &\qquad (\text{when } t_1 < T, t_2 < T) \\ &= 1 \ (\text{when } t_1 < T, t_2 > T) \\ &= 0 \ (\text{when } t_1 > T, t_2 < T), \qquad (3)\end{aligned}$$

where r_1 and r_2 are the overall rates of food on the left and right keys, respectively. This extension reasserts the well-known fact that a reinforcement that is experienced more often than another has more impact in controlling the behavior on which it is contingent. Furthermore, it asserts it in such a way that at the boundary condition where the durations of the terminal links are zero ($t_1 = t_2 = 0$), Equation 3 reduces to the familiar matching law (Herrnstein, 1970).

It follows from Equations 2 and 3 that the longer the choice phase, the greater T is, and as T becomes indefinitely large $B_1/(B_1 + B_2)$ approaches .5. An intuitive justification for this prediction is presented in Fantino and Logan (1979, p. 231). This prediction has been confirmed in many studies (see Fantino, 1977, and Davison & McCarthy, 1988, chap. 10, for reviews). It also follows that for a given choice phase (say one that lasts 30 s, on the average) preference for the favored alternative increases as the outcome durations (t_1 and t_2) are increased while maintaining a constant ratio between them. Thus, increasing the duration of the outcome phases (while maintaining a constant ratio between them) relative to the duration of the choice phase has the same effect as decreasing choice phase duration relative to that of the outcome phase: Preference for the favored alternative should increase. This prediction has also been confirmed (e.g., MacEwen, 1972; Williams & Fantino, 1978).

DRT has been supported in cases in which its predictions have been pitted against predictions from competing viewpoints. For example, Dunn and Fantino (1982) pitted delay reduction against rate of reinforcement as potential controlling variables in a concurrent-chains experiment and found that rate of reinforcement accounted for choice if and only if its predictions were consistent with those of DRT. In their three-alternative version of Equation 3, Fantino and Dunn (1983) showed that Luce's (1977) choice axiom, also known as the "independence from irrelevant alternatives" axiom of formal choice theories in economics and psychology, was consistent with choice if and only if its predictions were consistent with those of DRT.

More recently LaFiette and Fantino (1989) have shown that DRT makes accurate predictions under radically different deprivation conditions (i.e., in both open and closed economies; after Hursh, 1980), and Fantino and Preston (1988) have successfully applied DRT to a foraging analogue in which the following counterintuitive prediction was supported: As the less profitable of two outcomes was encountered more frequently (with the absolute encounter rate for the more profitable kept constant) there came a point at which subjects accepted it less frequently; that is, greater accessibility led to lesser acceptability.

Incentive Theory (IT)

The name of this collection of models derives from the assumption that reinforcement acts not by retroactive strengthening but rather by the combination of two factors, arousal and sign-tracking. Behavior is invigorated by the incentive-motivational arousal associated with a stimulus and is attracted to spatial locations and response topographies that signal reinforcement (Killeen, 1975, in press).

Incitement. The incentive effects are assumed to be proportional to the rate of reinforcement, r_1, for those responses. This assumption was based on the demonstration that in a model for temporal control during periodic reinforcement, one parameter, A (which predicts the rate at which responding would occur if there were no competing responses or temporal inhibition), was shown to be proportional to the rate of reinforcement (Killeen, 1979).

Direction. The directive effects of primary reinforcement depend on a number of factors, such as the salience of the response and the delay of reinforcement. During a delay, alternate responses might occur and block the strengthening effect of the reinforcer on the initiating response (Killeen, 1982a). Assume that the probability of strengthening the response in the case of no delay is p, a parameter whose value is related to the salience of the response. When the reinforcer is delayed the probability of an alternate response occurring and blocking control during any succeeding instant is q. Then the probability of strengthening the original response after a delay of one unit of time is $p(1 - q)$, after two units is $p(1 - q)(1 - q)$, and after t units is $p(1 - q)^t$. If we let the units of time become very small, this geometric progression may be approximated by the exponential:

$$P = pe^{-qt}. \tag{4}$$

Conditioned reinforcement. In most situations the directive effects of a reinforcer are not transferred by the primary reinforcement alone. There are other signs of reinforcement, such as a change in keylight color, to which the animal may be attracted. If these signs are also contingent on the response to be strengthened, their effects will add to those of the delayed primary reinforcer. However, conditioned reinforcers are not as potent as the primary reinforcers that they signal. How does the conditioned reinforcement strength of a stimulus decrease as a function of the delay between its onset and the primary reinforcer? The simplest monotonically decreasing function is:

$$C = 1/t, \quad t > 0 \tag{5}$$

which measures the temporal density of reinforcement and was successful in capturing conditioned reinforcement effects (Killeen, 1982b). The reciprocal of t is a function that, over a substantial range, is not very different from the exponential decay function of t. Its contribution to the direction of behavior will be manifest primarily when experimental manipulations differentially introduce or remove conditioned reinforcers from one of the alternatives.

Putting it all together. Reinforcement incites behavior. The amount of such behavior available is proportional to its rate of reinforcement, r. If responding is not directed to signs of reinforcement, it is manifest as adjunctive behavior. The two directive vectors of reinforcement, P and C, sum to produce the overall directive effects: $S_d = P + C$. Operant behavior will be a product of these two factors: $B = r(P + C)$. The rate of responding on one operandum relative to the rate on another is given by

$$\frac{B_1}{B_1 + B_2} = \frac{br_1(P_1 + C_1)}{r_1(P_1 + C_1) + r_2(P_2 + C_2)}. \tag{6}$$

There are two free parameters, q (which participates in the determination of P, Equation 4) and a response bias parameter, b (which incorporates both relative response biases, p_i, and stimulus biases). A value of $q = 0.12$ provides a good fit to the data from many studies (see Figure 2).

An oversight corrected. There is an inconsistency in the above formulation. The parameter q was stipulated as the probability that an alternative response would intervene between the originating response and reinforcement, and thereby block the conditioning. But that probability cannot be independent of the rate of reinforcement, as regular reinforcement incites a whole panoply of behavior, both directed and undirected. Killeen (1975) and Killeen, Hanson, and Osborne (1978) showed that the theoretical maximum rate of potentially interfering adjunctive responding was proportional to the rate of reinforcement. Therefore the parameter q should also be proportional to the rate of reinforcement. Its reciprocal, which we shall call τ, should be proportional to the average time between reinforcements (T):

$$\tau = 1/q = kT, \tag{7}$$

where k is a free parameter relating rate of reinforcement to rate of responding. Note that the average time between reinforcements (T) is calculated exactly the same here as in DRT. Reanalysis of the data in Figure 2 with

$$P = pe^{-t/\tau} \tag{4'}$$

shows that it is fit just as well (and in some cases better) with this treatment of the primary effects of delayed reinforcement as it was with the original formulation. There is not a big

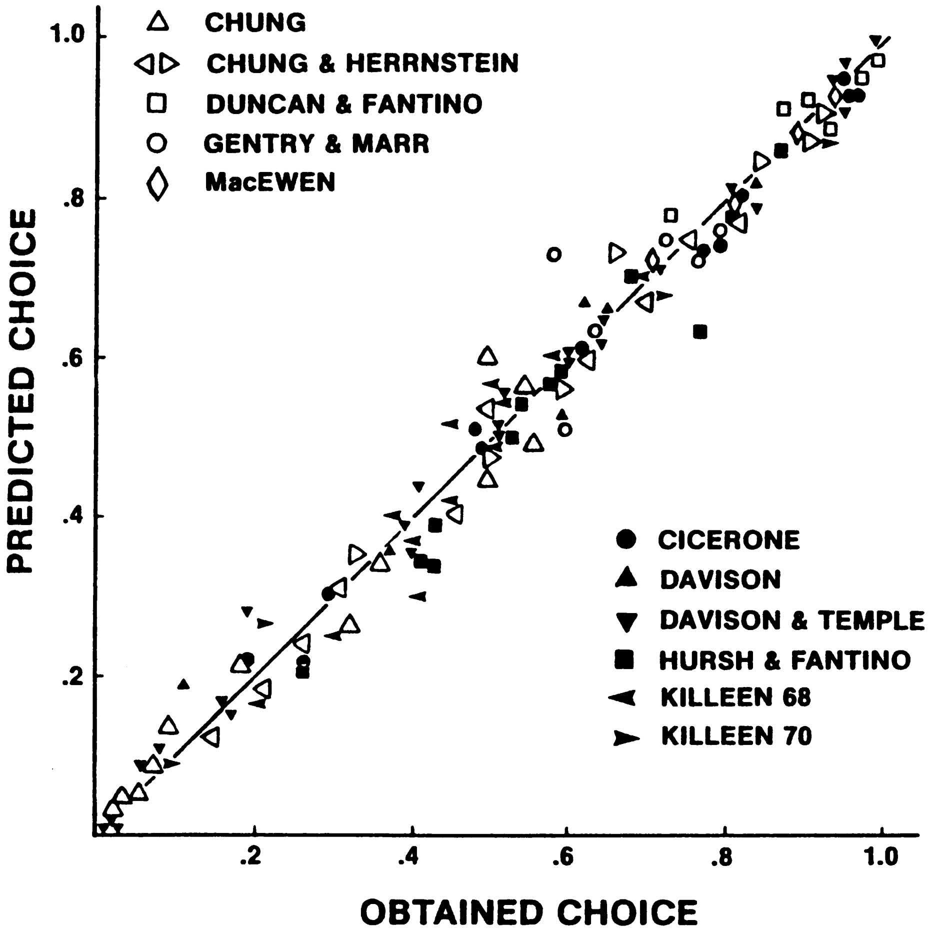

Fig. 2. Relative rates of responding in the initial links of concurrent-chains schedules versus the relative rates predicted by Incentive Theory 1. For filled symbols the parameters *q* and *b* were 0.125/s and 1.0. For the other data the parameters were: Chung (1964)(0.15, 1.3); Chung and Herrnstein (1967)(0.10 and 1.2 [8-s standard] or 1.8 [16-s standard]); Duncan and Fantino (1970)(0.125, 2.4); Gentry and Marr (1980)(0.25, 0.33); MacEwen (1972)(0.08, 1.5). With DRM there is less need for idiosyncratic values of the parameters, as those are automatically adjusted as the overall rate of reinforcement is varied. (Figure from Killeen, 1982c; copyright Elsevier Press. Reprinted with permission.)

difference between the two models because the average rate of reinforcement does not change very much in these studies. We shall call this amended version Incentive Theory II (IT2). There is an important conceptual difference, however, because this formulation has the delay of reinforcement gradient changing as a function of the rate of reinforcement; it is steeper in a rich environment and shallower in an environment where reinforcements are infrequent. This is an important implication and one that generates a bridge to a recent model of time perception.

Relation to a behavioral theory of timing. Killeen and Fetterman (1988) offered a theory of how animals behave when reinforced for discriminating the passage of time. They posit a central pacemaker with a period proportional to the average time between reinforcers in the experimental context. The pacemaker emits pulses that serve as stimuli to move animals from one state to the next in a sequence. These states are correlated with different types of adjunctive behavior. If the animals are called upon to estimate the period of time since the interval began, they do so by making a conditional discrimination based on the type or location of the adjunctive behavior they were emitting at the time of the question. Throughout much of the paper Killeen and Fetterman assume that the pacemaker is a Poisson emitter; that is, that pulses are independent and exponentially distributed. (Some discriminations are too accurate to support such a simple model and require other pacemakers whose period is more regular.) In analyzing the control by delayed reinforcers, they argued that animals stay in the state that corresponds to the initiating response until the next pulse from the pacemaker, whereupon they move to a different state. If the primary reinforcer is delivered before that transition, it strengthens the initiating response, otherwise it does not.

Because the waiting time between pulses is distributed exponentially, these two treatments of control by delayed reinforces—incentive theory (with $1/\tau$ replacing q) and the behavioral theory of timing (BTT)—are formally identical. Instead of intervening responses blocking control with a certain constant probability (IT2), the animal moves out of the state that supported the initiating response with constant probability (BTT). In both cases, the probability of blocking or the probability of moving to a new state is proportional to the overall rate of reinforcement in the context and is constant over short periods of time, so that an exponentially decaying effect of primary reinforcement is predicted.

UNIFICATION OF DRT AND IT

DRT and IT have almost always provided very similar predictions of behavior (Fantino & Abarca, 1985; Killeen, 1985). Why the persistent similarity in predictions, given the very different assumptions underlying the two models? The following paragraphs seek to explain that similarity.

Several basic tactical assumptions—the choice of relative rate of responding as the dependent variable of interest, and the importance of the rate at which each alternative is experienced—are common. Both may be incorporated in the basic equation of prediction:

$$\frac{B_1}{B_1 + B_2} = \frac{r_1 f(x)}{r_1 f(x) + r_2 f(x)}. \tag{8}$$

Table 1

Correlations (r) between the predictions of Equation 8 with the exponential as the kernel versus the predictions of that equation with a series approximation to the exponential as the kernel.

Equation	Variable	Approximation $1 - t/T$	$+ (t/T)^2/2$	$- (t/T)^3/6$
$e^{-t/T}$	r	.848	.993	.999
	k	.2	1.0	1.0
$e^{-t/(kT)}$	r	.950	.994	.999

The models differ in the function $f(x)$ by which the rates of reinforcement are multiplied. For DRT that function is $(T - t_i)$, whereas for IT2 it is $p_i e^{-t_i/\tau}$ (assuming $C_i = 0$, for simplicity). We shall call this parenthetical function the *kernel* of the models. To show how they are related when they have different kernels, we may examine a Taylor Series approximation to the exponential (for the moment we consider the simple case when $p_1 = p_2 = k = 1$ and therefore $\tau = T$):

$$\begin{aligned} e^{-t/T} = 1 &- t/T + (t/T)^2/2 \\ &- (t/T)^3/6 + \ldots \\ &+ (-t/T)^n/n! + \ldots . \end{aligned} \tag{9}$$

The accuracy of the approximation to the exponential increases with the number of terms included in the series, but often just the first few terms provide an adequate approximation. How many terms are necessary for our purposes? To know that, we must know the typical values for t/T, because the smaller that fraction, the more quickly the series converges. To assess the accuracy of approximation, we will correlate the predictions of Equation 8 when the kernel is the exponential function with the predictions of that equation when the kernel is the Taylor series approximation, as the number of terms increases. We shall do that for a simulation of concurrent-chains experiments, using a range of terminal-link values that cover the range typically employed.

The simulated schedules had VI 1-min concurrent schedules in the initial link, and fixed-interval (FI) schedules ranging from 5 s to 200 s in the terminal links. All possible different pairs of values for the terminal-link delays were evaluated ($N = 820$), given a step size of 5 s between values. The results are listed in Table 1.

The approximations in the first column of Table 1 comprise the first two terms of the series $(1 - t/T)$, in the second column those plus a third term, and in the third column those plus a fourth term. We see from the first row that the correlations for three- and four-term approximation are excellent, but the first two terms alone give us only $r = .848$. However, we may improve the performance of the two-term approximation by adding a parameter (k) to the exponential (but not to its approximation). In the comparison mentioned above, $k = 1$, so that it was not operative as a factor. But we may attempt to find some value of k that maximizes the correlation, and that factor is shown in the second row of the table. Clearly the best way to maximize the correlation is to add additional terms to the approximation, not a free parameter to the exponential. And, in fact we see that with three or four terms the correlation is almost perfect, and the free parameter is forced to 1.0. But the point we wish to make here is that the first two terms of the series can provide a very good approximation ($r = .95$), not to the original exponential, but to an exponential where the denominator of the exponent, T, is multiplied by the factor k. The best value for that factor will depend on the range of values of t_1 and t_2 used in the experiment. With a value of the factor anywhere close to optimal, it would be difficult to choose between the model and its approximation, based on experimental data.

Note that the exponential equation in Table 1 is similar to the primary gradient in the kernel for IT2. The only difference is the interpretation of the parameter k: In the above analyses, it was introduced merely to improve the correlation between a function and its approximation. In IT2, it represents the effects of competing types of behavior incited by reinforcement. The larger that k is, the flatter will be the delay gradient for any given rate of contextual reinforcement. If, in IT2, the value of k was large—say significantly larger than one—then the first two terms of the series would be a poor approximation to that function. However, the value reported by Killeen and Fetterman (1988) was around 0.3 for pigeons. For that value of k, the correlation between IT2 and the approximation provided by the first two terms is quite good—.93.

We have demonstrated that the primary reinforcement gradient in the kernel of IT2 may be represented as the first two terms of a Taylor Series approximation, which may be written $P_i = (T-t_i)/T$ (cf. Equation 1). To simplify exposition, let us assume for the moment that there is no differential conditioned reinforcement in the situation ($C = 0$) and that there are no differential biases ($p_1 = p_2$). If we then insert P_i as the kernel in Equation 8, we find that the T's in their denominators cancel, leaving us:

$$\frac{B_1}{B_1 + B_2} = \frac{r_1(T - t_1)}{r_1(T - t_1) + r_2(T - t_2)}, \qquad (10)$$

where r_i and T are the same as in Equation 3. Compare the result with Equation 3. We see that Equation 8 with the DRT kernel is the same as Equation 8 with the Taylor Series approximation of the IT2 kernel.

What of the conditioned reinforcement addend, $C_i = 1/t_i$, that we omitted? If we had added the complete expression for directive strength as the kernel, $S_d = P_i + C_i$, the approximation would become $(T - t_i + T/t_i)/T$, which reduces to:

$$S_d = (T(1 + C_i) - t_i)/T. \qquad (11)$$

Again, the T in the denominator would cancel out when these kernels are inserted into Equation 8. It can be seen that this factor is very similar to the previous kernel, except when C_i is large, that is, when t_i is small (say, less than 5 s). The correlations between Equation 3 and IT2 in the above simulations ($k = 0.3$) were .93 both when C_i was set to zero and when it took its proper value of $1/t_i$. For small values of the terminal-link delays, conditioned reinforcement enhances the preference for the preferred alternative.

DRT and IT2 thus differ only in that one is an approximation of the other, and this analysis shows that the approximation can be a good one. Of course, the premises of the original theories that generated the models remain different and testable. These theoretical distinctions will be discussed in the final paragraphs of the paper.

Correlations with Real Data

How do the models compare when confronted with real data? Fantino and Davison (1983) report pigeons' preferences for a range of terminal-link schedules. The first 11 conditions were quite similar to the conditions studied in the above simulation: 60-s initial links and various FI terminal links. Fantino's DRT accounted for 99% of the variance in those conditions, leaving an average deviation of 2.8 percentage points between the data and the predictions. IT2 fit the data as well, accounting for 99% of the variance and leaving an average deviation of 2.2 points, with $k = 0.39$.

Variations in the initial links. As noted above, the initial links do more than provide an opportunity to observe extended sequences of choice responses; they also control those responses. When the initial links are short, relative response rates are more extreme than when they are moderate, and this trend continues to long values of the initial links. Both DRT and IT2 predict this effect. To test the similarity of predictions, another simulation was conducted in which concurrent schedules were designed with the values of each of the initial links and the values of each of the terminal links varied randomly on each trial between 1 s and 200 s. The parameter k was given a value of 0.3. In a run that evaluated 10,000 conditions the correlation of the models was $r = .98$. Because of this close correlation, the DRT and IT2 will henceforth be treated as a single model, to be termed the *delayed reinforcement model,* or DRM. The presence or absence of conditioned reinforcement can be reflected in the unified model by the presence or absence of the addend $1/t$. Because its effect will not be significant except in those experiments that vary the conditioned reinforcement stimuli (e.g., Dunn, Williams, & Royalty, 1987), it will be omitted from further consideration in this paper. It may be represented either with the difference kernel of DRT or with the exponential function of IT2.

Despite its general success with concurrent-chains experiments in which the terminal links are varied and with concurrent-chains studies in which equal initial links are varied, DRM does less well coping with data from experiments in which the initial links are of unequal duration. When all 56 conditions of the Fantino and Davison (1983) series are analyzed, the percentage of variance that DRM accounted for decreased to 82%, leaving an average deviation of 8 points. Although there are reasonable ways to alter DRM that would improve this fit, they come at the cost of greater complexity and will not be developed in the present paper. Instead we now turn to a brief discussion of how the DRM relates to other models addressing choice in the concurrent-chains procedure.

Mazur

An alternate Taylor Series for the exponential kernel of IT2 is:

$$\begin{aligned} e^{-t/\tau} &= 1/e^{t/\tau} \approx 1/(1 + t/\tau) \\ &= 1/(1 + Kt). \end{aligned} \qquad (12)$$

This converges on the exact value for the exponential from above, whereas the previous approximation converges from below. The final term is Mazur's (1984) empirical model for choice when both initial links are continuous reinforcement (CRF) schedules. The earlier simulation with concurrent schedules with intervals ranging from 5 s to 200 s in 5-s steps was replicated, this time with 1-s initial links. The predictions of Mazur's model are highly correlated with those of DRM—$r = .84$ with $k = 0.3$. But there are systematic differences, and his predictions with CRF initial links fall closer to his data than do those of DRM. Why should this be? The basic difference between the models is the multiplication of the kernels by the rates of reinforcement in Equation 8 for DRM, whereas for Mazur, only the kernel itself (Equation 12) appears. For such short initial links we therefore predict preferences that are more extreme than Mazur's predictions and more extreme than pigeons' performance in that paradigm.

We believe the critical factor is the same one that causes DRM inaccuracy with very short initial links, especially when they are asymmetric (as was the case in the Fantino & Davison, 1983, data mentioned above): For discrete-trials procedures such as Mazur's (1984) CRF initial links, or for any initial-link schedules that last for less than 10 s, transition to the terminal links can reinforce only one or a few of the responses leading up to it, and the size of the response unit that is strengthened by that reinforcer is thereby diminished. Catania, Sagvolden, and Keller (1988) and Killeen and Smith (1984) have argued that the decreased marginal efficacy of high rates of reinforcement is due to the truncation of opportunities for one reinforcer, following close on the heels of a previous reinforcer, to strengthen responses that occurred before the previous reinforcer. Once the time allowed for uninterrupted responding becomes small enough, there may be no further control by reinforcement rate. In discussing their impressive dynamic model of free-operant choice, Myerson and Hale (1988) noted that "different rules may govern transition-state behavior in free-operant and discrete-trials choice situations. . . . It seems likely that, in the absence of switching that is controlled by overall reinforcement rates [in discrete-trial situations], the relevant dependent variable may become probability of responding, which in turn may be controlled by local reinforcement rates" (p. 300). Their basic dynamic model of transitions underlying choice (Equation 4) differs from Staddon's (1988) quasi-dynamic model (Equation 1) of the same phenomena; Staddon's model was developed out of the discrete-trial data of Horner and Staddon (1987). As Williams (1988) notes, "Reinforcement thus appears to change behavior in two separate ways: by selecting the response unit, and by determining the strength of that unit. . . . The correct interpretation of various changes in behavior often depends critically on their dissociation. . . . This issue has not been totally resolved" (p. 173). We believe this issue underlies the distinction between DRM and kinetic theories of choice on the one hand, which are most appropriate for free-operant concurrent and concurrent-chains paradigms, and Mazur's and Staddon's models on the other, which have grown out of discrete-trial paradigms.

If the DRM were to shed the factor relating behavior to the context of reinforcement (r) in the discrete-trial paradigm it would become more similar to Mazur's (1984) model, thus pointing to a possible unification of models for free-operant and discrete-trial choice. However, the contextual factor is critical in accounting for choice in the free-operant case. Indeed, for initial links that are not CRF, Mazur's model would become quite different from DRM because it does not have a factor relating behavior to the context of reinforcement. Thus, an extension of Mazur's model to free-operant paradigms would not be expected to do as well as the DRM, which makes accurate predictions over a wide range of values.

Vaughan

Vaughan (1985) attempts to extend one model of concurrent responding, melioration (Herrnstein & Vaughan, 1980), to concurrent-chains situations. His tactic is to treat all of the strength of the terminal-link reinforcement as being conferred by the terminal-link conditioned reinforcers, whose strength increases as a hyperbolic function of the rate of reinforcement that they signal (compare his Equation 7 with our Equation 1). He is able to provide excellent qualitative accounts of a large range of choice data with no changes in his parameters. This is possible in large part because he makes only one quantitative prediction (and even that only for didactic purposes, as it maps no data) and therefore does not exploit the rich quantitative data base that is available in this area to test and sharpen his model. As he notes candidly in his Appendix 2, some of his basic assumptions falter as soon as they are compared to data; an ad hoc correction "complicates the mathematics without changing the qualitative trends" so it is omitted. He notes the qualitative similarity of melioration's predictions with those of DRT and suggests experiments that might discriminate between them. In addition, we believe that it would be valuable to confront melioration with real data from relevant experiments and see if it can do even as well as the flawed model that generated the account shown in Figure 2.

The Canonical Form of DRM

Which version of the model are we calling DRM, the one with the difference kernel derived from DRT or the one with the exponential kernel derived from IT2? We have seen that over a wide range of data, the predictions of DRT and IT2 are the same, and that there are good formal reasons for that identity. We believe that both forms are useful. The DRT version is simpler and is the preferred form when qualitative prediction or extensions to new paradigms (e.g., Fantino & Abarca, 1985) are desired. The IT2 version has an added parameter and requires nonlinear curve fitting. On the other hand, this provides flexibility in accounting for individual differences. The present paper stresses the usefulness of both forms and does not attempt the very difficult task of choosing between them on empirical grounds. Whatever one's conceptual predilections towards the two forms, however, it is clear that the models based on these forms, DRT and IT2, may be treated in practice as a single model.

Similar Models, Different Theories

As we noted in the introduction, the convergence of theories on a common predictive model does not mean that the theories are also convergent. DRT is a theory of conditioned reinforcement; IT2 is a theory of arousal and primary and conditioned reinforcement. As they stand, a difference in predictions may be found for very short terminal links, where IT2 tends to predict more extreme preferences than DRT. However, the presence of the free parameter *k* in IT2 permits that model to accommodate a range of empirical data, thus blunting empirical distinction. In principle, there remain several avenues of experimental contrast between these theories. Some of these involve the roles of the frequency of conditioned reinforcement, the salience of the response that initiates terminal-link delays (the parameter *p*), and the probability of blocking primary reinforcement of the initiating response by the imposition of alternate classes of reinforceable behavior during the terminal-link delays (the parameter *q*). In practice, however, it may be difficult to conduct definitive experimental tests to distinguish the adequacy of the two theories.

Whereas models may be judged as valid or invalid against relevant data, theories (DRT and IT2 included) are by nature more protean, and because they are evolving continually, come to be viewed as more or less useful, more or less powerful, more or less parsimonious. We believe that our demonstration of the convergence of DRT and IT2 on a common model of prediction, DRM, is important. By showing where the theoretical predictions align, it will permit both simple predictions in standard situations and more efficient tests of those predictions that remain unique to each theory. We conclude by repeating that the models based on these theories may be treated in practice as a single model. That model stresses the importance of the delay within a temporal context of reinforcement as the central determinant of choice.

REFERENCES

Abarca, N., & Fantino, E. (1982). Choice and foraging. *Journal of the Experimental Analysis of Behavior,* **38,** 117–123.

Autor, S. M. (1960). *The strength of conditioned reinforcers as a function of frequency and probability of reinforcement.* Unpublished doctoral dissertation, Harvard University.

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

Case, D. A., & Fantino, E. (1981). The delay-reduction hypothesis of conditioned reinforcement and punishment: Observing behavior. *Journal of the Experimental Analysis of Behavior,* **35,** 93–108.

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.

Chung, S. (1964). Effects of delayed reinforcement in a concurrent situation. *Journal of the Experimental Analysis of Behavior,* **8,** 439–444.

Chung, S., & Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior,* **10,** 67–74.

Cicerone, R. A. (1976). Preference of mixed *versus* constant delay of reinforcement. *Journal of the Experimental Analysis of Behavior,* **25,** 257–261.

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

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

Davison, M. C., & Temple, W. (1973). Preference for fixed-interval schedules: An alternative model. *Journal of the Experimental Analysis of Behavior,* **20,** 393–403.

Duncan, B., & Fantino, E. (1970). Choice for periodic schedules of reinforcement. *Journal of the Experimental Analysis of Behavior,* **14,** 73–86.

Dunn, R., & Fantino, E. (1982). Choice and the relative immediacy of reinforcement. *Journal of the Experimental Analysis of Behavior,* **38,** 321–326.

Dunn, R., Williams, B., & Royalty, P. (1987). Devaluation of stimuli contingent on choice: Evidence for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior,* **48,** 117–131.

Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior,* **12,** 723–730.

Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313–339). Englewood Cliffs, NJ: Prentice-Hall.

Fantino, E. (1981). Contiguity, response strength, and the delay-reduction hypothesis. In P. Harzem & M. D. Zeiler (Eds.), *Advances in analysis of behaviour: Vol. 2. Predictability, correlation, and contiguity* (pp. 168–201). Chichester, England: Wiley.

Fantino, E. (1982). Effect of initial-link length on responding in terminal link. *Behaviour Analysis Letters,* **2,** 65–70.

Fantino, E., & Abarca, N. (1985). Choice, optimal foraging, and the delay-reduction hypothesis. *Behavioral and Brain Sciences,* **8,** 315–362. (Includes commentary)

Fantino, E., & Case, D. A. (1983). Human observing: Maintained by stimuli correlated with reinforcement but not extinction. *Journal of the Experimental Analysis of Behavior,* **40,** 193–210.

Fantino, E., & Davison, M. (1983). Choice: Some quantitative relations. *Journal of the Experimental Analysis of Behavior,* **40,** 1–13.

Fantino, E., & Dunn, R. (1983). The delay-reduction hypothesis: Extension to three-alternative choice. *Jour-*

nal of Experimental Psychology: Animal Behavior Processes, **9,** 132–146.

Fantino, E., & Logan, C. A. (1979). *The experimental analysis of behavior: A biological perspective.* San Francisco: Freeman.

Fantino, E., & Preston, R. A. (1988). Choice and foraging: The effects of accessibility on acceptability. *Journal of the Experimental Analysis of Behavior,* **50,** 395–403.

Gentry, G. D., & Marr, M. J. (1980). Choice and reinforcement delay. *Journal of the Experimental Analysis of Behavior,* **33,** 27–37.

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

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior,* **13,** 243–266.

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

Horner, J. M., & Staddon, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioural Processes,* **15,** 59–92.

Hursh, S. R. (1980). Economic concepts for the analysis of behavior. *Journal of the Experimental Analysis of Behavior,* **34,** 219–238.

Hursh, S. R., & Fantino E. (1973). Relative delay of reinforcement and choice. *Journal of the Experimental Analysis of Behavior,* **19,** 437–450.

Ito, M., & Asaki, K. (1982). Choice behavior of rats in a concurrent-chains schedule: Amount and delay of reinforcement. *Journal of the Experimental Analysis of Behavior,* **37,** 383–392.

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

Killeen, P. R. (1970). Preference for fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior,* **14,** 127–131.

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

Killeen, P. R. (1979). Arousal: Its genesis, modulation, and extinction. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behaviour: Vol. 1. Reinforcement and the organization of behaviour* (pp. 31–78). Chichester, England: Wiley.

Killeen, P. R. (1982a). Incentive theory. In D. J. Bernstein (Ed.), *Nebraska symposium on motivation, 1981: Response structure and organization* (pp. 169–216). Lincoln: University of Nebraska Press.

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

Killeen, P. R. (1982c). A model for concurrent chain performance. *Behaviour Analysis Letters,* **2,** 305–307.

Killeen, P. R. (1985). Delay reduction: A field guide for optimal foragers? *Behavioral and Brain Sciences,* **8,** 341–342.

Killeen, P. R. (in press). Behavioral geodesics. In D. S. Levine & J. S. Levin (Eds.), *Motivation, emotion, and goal direction in neural networks.* Hillsdale, NJ: Erlbaum.

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review,* **95,** 274–295.

Killeen, P. R., Hanson, S. J., & Osborne, S. R. (1978). Arousal: Its genesis and manifestation as response rate. *Psychological Review,* **85,** 571–581.

Killeen, P. R., & Smith, J. P. (1984). Perception of contingency in conditioning: Scalar timing, response bias, and erasure of memory by reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes,* **10,** 333–345.

LaFiette, M. H., & Fantino, E. (1989). Responding on concurrent-chains schedules in open and closed economies. *Journal of the Experimental Analysis of Behavior,* **51,** 329–342.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology,* **15,** 215–233.

MacEwen, D. (1972). The effects of terminal-link fixed-interval and variable-interval schedules on responding under concurrent chained schedules. *Journal of the Experimental Analysis of Behavior,* **18,** 253–261.

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

Myerson, J., & Hale, S. (1988). Choice in transition: A comparison of melioration and the kinetic model. *Journal of the Experimental Analysis of Behavior,* **49,** 291–302.

Navarick, D. J., & Fantino, E. (1976). Self-control and general models of choice. *Journal of Experimental Psychology: Animal Behavior Processes,* **2,** 75–87.

Squires, N., & Fantino, E. (1971). A model for choice in simple concurrent and concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior,* **15,** 27–38.

Staddon, J. E. R. (1988). Quasi-dynamic choice models: Melioration and ratio invariance. *Journal of the Experimental Analysis of Behavior,* **49,** 303–320.

Vaughan, W., Jr. (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior,* **43,** 383–405.

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

Williams, B. A., & Fantino, E. (1978). Effects on choice of reinforcement delay and conditioned reinforcement. *Journal of the Experimental Analysis of Behavior,* **29,** 77–86.

Wixted, J. T. (1989). Nonhuman short-term memory: A quantitative reanalysis of selected findings. *Journal of the Experimental Analysis of Behavior,* **52,** 409–426.

Received December 20, 1988
Final acceptance September 3, 1989