JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1985, 43, 383-405 NUMBER 3 (MAY)

CHOICE: A LOCAL ANALYSIS

WILLIAM VAUGHAN, JR.

HARVARD UNIVERSITY

Analyses of free-operant choice usually employ one of two general procedures: the simple concurrent procedure (i.e., a concurrent variable-interval variable-interval schedule) or the concurrent-chains procedure (i.e., concurrently available initial links, each leading to an exclusively available terminal link). Theories about choice usually focus on only one of the two procedures. For example, maximization theories, which assert that behavior is distributed between two alternatives in such a way that overall rate of reinforcement is maximized, have been applied only to the simple concurrent procedure. In the present paper, a form of the pairing hypothesis (according to which pairings between one stimulus and another affect the value of the first, and pairings between responses and reinforcers affect the value of the former) is developed in a way that allows it to make qualitative predictions with regard to choice in a variety of simple concurrent and concurrent-chains procedures. The predictions include matching on concurrent variable-interval variable-interval schedules, preference reversal in the self-control paradigm, and preference for tandem over chained terminal links.

Key words: choice, reinforcement, conditioned reinforcement, matching, maximizing, melioration, concurrent schedules, concurrent-chains schedules, feedback functions

INTRODUCTION

Choice procedures involve studying how an organism distributes its behavior across two or more simultaneously available alternatives. A result that has proved central was reported by Herrnstein (1961): When exposed to two concurrently available variable-interval (VI) schedules, pigeons distributed their responses between the two keys in approximately the same proportion as reinforcers were received. That is, if twice as many reinforcements were received on the left as on the right, pigeons tended to respond about twice as often on the left as on the right. This may be expressed as:

$$\frac{P_L}{P_R} = \frac{R_L}{R_R} . \qquad (1)$$

Here, P_L and P_R represent responses on the left and right, and R_L and R_R represent reinforcers received on the left and right. Subsequently it was found that the distribution of time also tends to equal, or match, the distribution of obtained reinforcers (Brownstein & Pliskoff, 1968; Catania, 1966):

$$\frac{T_L}{T_R} = \frac{R_L}{R_R} . \qquad (2)$$

T_L and T_R represent times spent on the left and right sides.

Because matching is such an orderly phenomenon, a good deal of effort has gone into both generalizing it to new situations (such as single-key schedules, multiple schedules, and concurrent-chains schedules) and to explaining why it occurs in the first place. For example, it has been said that matching occurs because that is the distribution of behavior which maximizes overall rate of reinforcement (Rachlin, Green, Kagel, & Battalio, 1976; Staddon & Motheral, 1978), a process that may be called global maximization. Others (e.g., Hinson & Staddon, 1983; Shimp, 1969) have suggested that matching results from a

---

Preparation of this manuscript was supported by NSF Grants IST-81-00404 and IST-8409824 to Harvard University. I thank the following for comments and encouragement: W. M. Baum, M. L. Commons, M. C. Davison, E. Fantino, R. J. Herrnstein, S. E. G. Lea, and J. A. Nevin. Reprints may be obtained from William Vaughan, Jr., Department of Psychology and Social Relations, Harvard University, Cambridge, Massachusetts 02138.

process, termed momentary maximizing, of always emitting that choice response with the highest momentary probability of reinforcement.

There are two procedures commonly used in the study of choice, the simple concurrent procedure and the concurrent-chains procedure. In the case of the simple concurrent procedure, two or more alternatives are simultaneously available. Responses on each give rise, according to some schedule, to primary reinforcement. In the case of the concurrent-chains procedure, responses give rise, according to some schedule, to some further stimulus in whose presence primary reinforcement is delivered, according to some different schedule (the "terminal link"). During that terminal link the other alternatives are unavailable (i.e., the alternative keys are darkened). The primary difference between the two procedures, then, is that whereas choice responses are directly reinforced in the case of the simple concurrent procedure, in the case of the concurrent-chains procedure choice responses produce other stimuli, in whose presence primary reinforcement is delivered. The matching relation discussed above was found using the simple concurrent procedure, with VI schedules on each of the alternatives. On the concurrent-chains procedure, it becomes necessary to incorporate the effect of delayed reinforcement into one's analysis of choice. The usual tactic has been to assume that matching of some form is occurring on such schedules, and to hypothesize just what the variables are to which the organism is matching.

However, there has occurred a certain division in the study of choice: In general, those who study choice in the concurrent-chains situation tend to assume some form of matching as an axiom, while that very assumption is being challenged by some others (e.g., advocates of global maximization or momentary maximization) studying choice using the simple concurrent procedure. The tactic pursued in the present paper involves taking a particular form of what is sometimes called the pairing hypothesis, and developing it in such a way that both the process of melioration (Herrnstein & Vaughan, 1980) and the conditioned values of stimuli present during the terminal links of concurrent chains may be deduced. Because, on simple concurrent schedules, melioration leads to matching at equilibrium, the present approach simultaneously addresses two domains of data, and so provides one possible link between them.

MATCHING ON CONCURRENT CHAINS, MAXIMIZATION, AND MELIORATION

Four models of choice on concurrent-chains schedules will be reviewed, primarily for purposes of orientation. In addition, predictions of the model to be developed here will be compared with predictions of these other models in terms of certain known data.

Baum and Rachlin (1969) suggested a particularly straightforward model: The ratio of times in two situations equals the ratio of the products of reinforcer rate, amount (A_i), and immediacy (I_i):

$$\frac{T_L}{T_R} = \frac{R_L A_L I_L}{R_R A_R I_R} . \tag{3}$$

Although this was suggested primarily with respect to simple concurrent schedules, the inclusion of immediacy (the reciprocal of delay) logically implies the concurrent-chains procedure. At the time it was proposed, this formulation provided an integration of data from a number of experiments (e.g., Catania, 1963; Chung & Herrnstein, 1967). For Baum and Rachlin, Equation 3 constitutes an axiom, in the sense that no attempt is made to deduce it from other, more primitive assumptions.

A second model, the delay-reduction hypothesis, was introduced by Fantino (1969a), who pointed out that certain earlier formulations had relied largely on data from concurrent-chains experiments in which the initial links were concurrent VI 60-s schedules. Fantino reasoned that with unequal terminal links (so that one was better than the other), as equal initial links became richer, behavior should shift toward the better side, an outcome inconsistent with these earlier models. He reported data in accord with this intuition, and suggested an equation that would handle these data. As formulated by Squires and Fantino (1971), this reads:

$$\frac{T_L}{T_R} = \frac{R_L(T - t_{2L})}{R_R(T - t_{2R})} . \tag{4}$$

Here, T is the average time to primary reinforcement from the onset of the initial links, t_{2L} and t_{2R} are the left and right terminal-link times, and R_L and R_R are the overall rates of reinforcement on the left and right keys. That is, R_L equals the number of reinforcers during a left terminal link, divided by the sum of the left initial-link schedule and the left terminal link. This formulation has been applied to a wide range of experiments over more than 10 years.

A third approach was taken by Davison and Temple (1973), who reported the results of a large number of concurrent-chains experiments with fixed-interval (FI) terminal links. They proposed the following formulation as an alternative to that of Squires and Fantino:

$$\frac{T_L}{T_R} = E \frac{R_L t_{2R}}{R_R t_{2L}} . \tag{5}$$

Here, R_L and R_R have the same meanings as for Squires and Fantino. Times in the terminal links, inclusive of reinforcer times, are represented by t_{2L} and t_{2R}. E is a factor reflecting the ratio between obtained and programmed terminal-link entries. For example, if the initial links were VI 30 s on the left and VI 60 s on the right (so that the programmed terminal-link entries were in a ratio of 2 to 1), but the left terminal link was entered three times as often as the right, E would equal 1.5.

The fourth approach to be discussed is that of Killeen (1982), whose model of behavior on concurrent chains is based primarily on theoretical considerations (incentive theory). The equation that Killeen suggested is:

$$\frac{T_L}{T_R} = \frac{R_L[\exp(-qt_{2L}) + 1/t_{2L}]}{R_R[\exp(-qt_{2R}) + 1/t_{2R}]} . \tag{6}$$

In this equation R_i and t_{2i} have the same meanings as for Squires and Fantino. Although Killeen usually left the parameter q at 0.125, if changing q and/or introducing a bias term increased the amount of variance accounted for by more than 2%, he made such changes when addressing existing data.

Notice that the four models discussed above all assume some form of matching: The ratio of behavior on two schedules equals the ratio of some function of the schedule parameters. This is the sense in which these approaches take matching as an axiom. Maximization theory, whether global or momentary, assumes that no form of matching is basic; rather, it is assumed that animals behave in ways that maximize overall value or the probability of reinforcement at each response. Matching, by this account, just happens to coincide with maximization in a number of situations. On simple concurrent variable-ratio (VR) schedules, nearly exclusive preference is exhibited for the better side (Herrnstein & Loveland, 1975), a strategy consistent with both forms of maximizing (as well as with matching). Using a computer simulation of concurrent VI schedules, Rachlin et al. (1976) reported that, at matching, global rate of reinforcement is maximized on such schedules. The same result was shown analytically (making certain assumptions regarding the form of the VI feedback function) by Rachlin (1978) and by Staddon and Motheral (1978).

Herrnstein and Vaughan (1980) suggested an alternative interpretation of these data, which they termed melioration. They proposed that rather than maximizing overall rate of reinforcement (or the probability of reinforcement for each response), organisms are sensitive to the local rates of reinforcement on each of two alternatives (i.e., number of reinforcers produced by responding on a side divided by time on that side). If the local rate on one side is higher than on the other, melioration specifies that more time will be spent on the better side. This process can account for maximization on concurrent VR schedules and for matching on concurrent VI schedules. In addition, it is consistent with a number of procedures explicitly designed to discriminate between melioration and global maximization (Boelens, 1984; Herrnstein & Vaughan, 1980;

Vaughan, 1981; Vaughan, Kardish, & Wilson, 1982; also see Herrnstein & Heyman, 1979; Lea, 1976; Mazur, 1981).

A GENERALIZATION OF MELIORATION

The particular generalization of melioration pursued here is based in part on the attempt (Vaughan, 1982) to derive melioration from a strengthening process similar to that of Rescorla and Wagner (1972). Vaughan suggested that the value of a key (i.e., its strength as a conditioned reinforcer) was a function of the rate at which transitions occurred from that key to food. Let V_{1L} represent the value of the left key in a concurrent key condition, and V_{2L} represent the value of food produced by responding on that key. The first assumption to be made here is that V_{1L} is a negatively accelerated function of r_{1L} (short for R_L/T_L—reinforcers on the left divided by time on the left), approaching asymptote at V_{2L} for high rates of food delivery. If reinforcement rate were to fall to zero, it is further assumed that the value of the key would approach its unconditioned value, represented by V'_{1L}. For a key this unconditioned value is assumed to be zero. Figure 1 shows the case in which (local) rate of reinforcement on the left (r_{1L}) is greater than on the right (r_{1R}): The value of the left key (V_{1L}) will be greater than that of the right (V_{1R}). A bent arrow (at V'_1) indicates the unconditioned value of the stimulus under consideration. Arrows leading away (V_{1L} and V_{1R}) signify dependent variables (in terms of the specific function under consideration), and arrows leading to the function signify independent variables. As will be seen below, in some cases independent variables for some functions are dependent on other functions.

The function shown in Figure 1 is a simple hyperbola:

$$V_1 = \frac{rV_2}{r + a} \cdot \qquad (7)$$

Here, V_1 is the value of the stimulus under consideration, V_2 is the value of the stimulus (at the moment it is presented) to which transitions are made, r is the rate of transition (in transitions per second), and a is a parameter (in this case 0.1) that controls the rate at which the asymptote is approached (see Appendix 1 also).

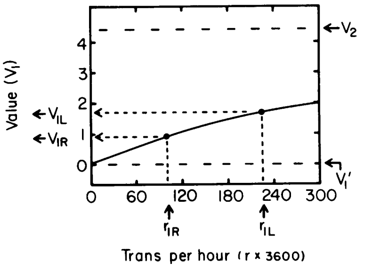

Fig. 1. Value of the left key (V_{1L}) and right key (V_{1R}) as a function of rate of transition on the left (r_{1L}) and right (r_{1R}) to a stimulus whose onset has value V_2. V'_1, assumed to be zero for a keylight, is the unconditioned value of Stimulus 1. Dependent variables for this function are shown as arrows leading away, and independent variables are shown leading to the function. A bent arrow signifies the unconditioned value of a stimulus. In the specific case shown, the rate of reinforcement on the left is 225 per hour and on the right is 100 per hour, leading to a higher value on the left than on the right.

On a simple concurrent VI VI schedule, the number of reinforcers delivered on one side is relatively independent of the amount of time spent there, given that changeovers occur moderately often. Consider, for example, a concurrent VI 40 s on the left and VI 60 s on the right. On the latter, a reinforcer will be received approximately once every minute, or 60 times per hour. If, in an hour's session, only one third of that hour is spent on the right, the local rate on the right will be 60 reinforcers divided by one third, or 180 reinforcers per hour. As this suggests, local rates of reinforcement are a function both of the VI schedules in effect and of the distribution of behavior. As usually programmed, the overall rate of reinforcement obtained on a VI schedule also depends somewhat on the distribution of behavior. However, in the case of what have been called linear VI schedules (Vaughan, 1976, 1982), local rate of reinforcement on a side does in fact equal the overall programmed rate divided by relative time on that side. Figure 2 shows local rates of reinforcement on the left and right, given a linear VI 40 s on the left and a linear VI 60 s on the right, as a func-

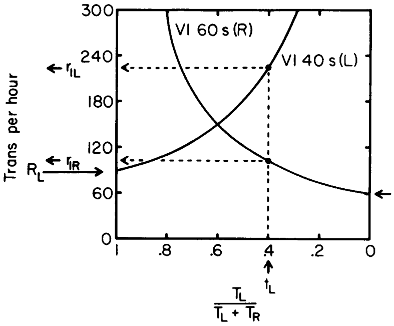

Fig. 2. Rate of transition (from keylight to food) on the left key (r_{1L}) and right key (r_{1R}) as a function of the overall programmed rates of reinforcement on the left (R_L: 90 per hour, or a VI 40 s) and right (R_R: 60 per hour, or a VI 60 s) as well as relative time on the left (t_L, 0.4 in this case). With linear VI schedules, $r_{1L} = R_L/t_L$ and $r_{1R} = R_R/(1 - t_L)$, or R_R/t_R.

tion of relative time on the left. In this Figure, r_{1L} and r_{1R} represent what the local rates of reinforcement would be if relative time on the left (t_L) were to equal 0.4, and R_L and R_R represent overall programmed rates of reinforcement on the left and right of 90 and 60 per hour, respectively.

The dependent variable of Figure 2 is local rate of reinforcement (transitions per hour), the independent variable of Figure 1. This allows the two figures to be conjoined, as in Figure 3, which shows the value of the left and right keys as a function of both the VI schedules in effect and the distribution of time on them. Here it can be seen that at 0.4 relative time on the left, the value of the left key will be greater than that of the right key.

In Vaughan (1982), changeovers from a key lower in value to one higher in value were assumed to strengthen the tendency to make that changeover response, whereas changeovers in the opposite direction were assumed to weaken the tendency to make that latter response. It was further assumed that if the strengths of changeover responses were modified as just mentioned, more time would come to be distributed to the key with greater value, because of those changes in strengths. One formalization of these assumptions may be expressed as:

$$d\frac{T_L}{T_L + T_R}/dt = f_x(V_{1L} - V_{1R}). \qquad (8)$$

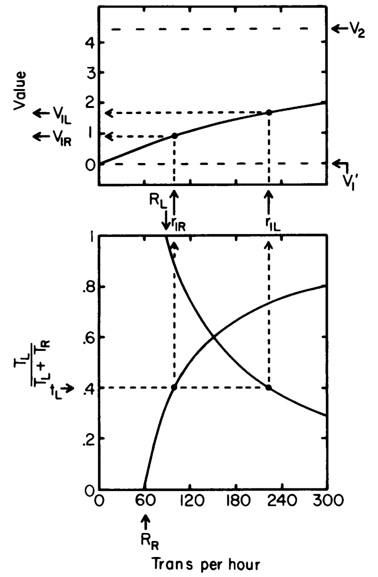

Fig. 3. Conjunction of Figures 1 and 2, showing values on the left (V_{1L}) and right (V_{1R}) as a function of R_L, R_R, and t_L. Note that r_{1R} and r_{1L} are dependent variables with respect to the lower part of the figure and are independent variables with respect to the upper part.

The left side of this equation represents the rate at which relative time on the left changes. The function f_x is assumed to be differentiable, monotonically increasing, and to have the property that $f_x(0) = 0$. According to this equation, if the value on the left exceeds that on the right, the rate of change of relative time on the left will be positive (or under certain conditions, zero), and vice versa. If the value on the left equals that on the right, the rate of change of relative time on the left will be zero. The subscript x represents a vector parameter, consisting of a number of components, possibly including relative times on the alternatives and values of those alternatives. Its only role is to imply that the function f may not be the same under all conditions, though it must always satisfy the above constraints. Myerson and Hale (1984), for example, have shown for some conditions that the rate of change on a side approaches zero as exclusive preference is approached. Equation 8 is represented in two ways in Figure 4, for the case of two keys, of values V_{1L} and V_{1R}. Panel A shows the derivative of relative time on the left, with respect to time, as a function of $V_{1L} - V_{1R}$.

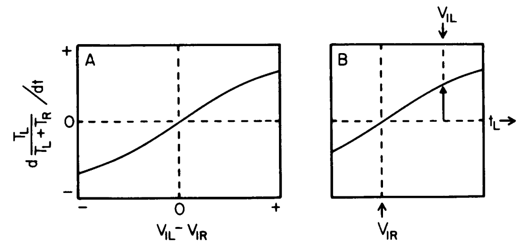

Fig. 4. A: Derivative of relative time on the left, with respect to time, as a function of difference between value on the left and right. B: An alternative representation, with derivative of relative time on the left shown as an upward arrow, indicating rate of change of t_L. Although it is assumed the function must satisfy certain constraints, the function itself may change as a function of certain quantities (e.g., V_{1L} or V_{1R}).

A slightly different representation is shown in Panel B. Here the x axis again represents value, but the origin is shifted to the value on the right, V_{1R}. Rate of change in relative time on the left can then be read off as the value of the function corresponding to V_{1L}. The upward arrow below V_{1L} corresponds to the rate (and direction) at which relative time on the left changes. Scale values are omitted from this figure because few relevant data are available.

Figures 3 and 4 may now be conjoined, as shown in Figure 5. The abscissa of the melioration function has been lined up with relative time on the left (t_L). Note that there are now two functions for value (on the left and right keys), shown separately for simplicity. These two values constitute the input for the process of melioration, whose output constitutes a rate of change of relative time on the left. In the example illustrated, value on the left exceeds that on the right, so relative time on the left will increase (the arrow for change of t_L is nonzero and points up). Equilibrium is reached when the two local rates of reinforcement are equal, because at that point the values of the two keys are equal. The equilibrium is stable because deviations from that distribution of behavior change the local rates in such a way that the above distribution is again approached. Equality of the two local rates of reinforcement is equivalent to Equation 2, matching of time ratios to reinforcer ratios. The model shown in Figure 5, then, constitutes a possible explanation of why such matching occurs.

Before attempting to extend this model to concurrent-chains experiments, it is worth pointing out a number of its properties. The use of linear VI schedules is not integral to the model, but simply serves to make the mathematics easier. In principle, any feedback function could be employed. Second, according to this model the direction in which the distribution of behavior shifts does not depend on overall rate of reinforcement, but only on local rates, as specified in Equation 8. The model is thus to some extent confirmed by experiments that favor melioration over global maximization (e.g., Boelens, 1984; Herrnstein & Heyman, 1979; Lea, 1976; Mazur, 1981; Vaughan et al., 1982). Third, the transition to reinforcement or to a conditioned stimulus is assumed to strengthen immediately preceding behavior (e.g., Ferster & Skinner, 1957; Morse, 1966; Skinner, 1938) or to change the value of contiguous stimuli. The model is thus an example of two-factor theory. Finally, rather than

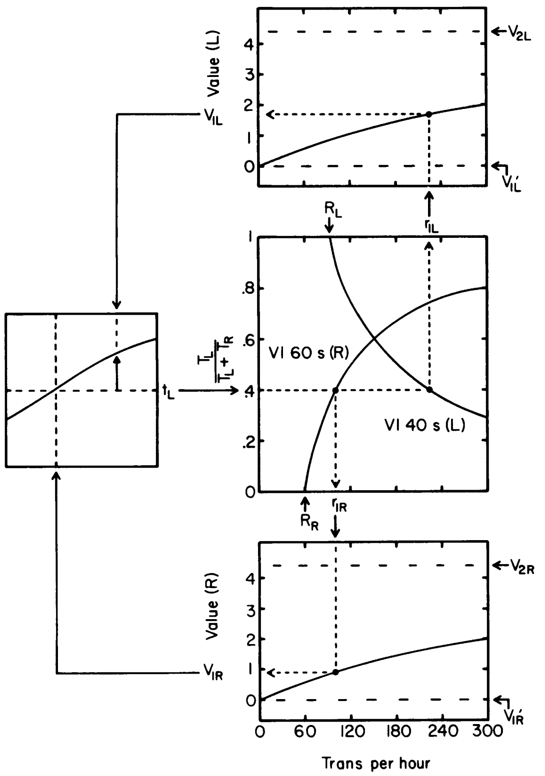

Fig. 5. Conjunction of Figures 3 and 4(B). With VI 40 s on the left key and VI 60 s on the right key, if relative time on the left equals 0.4, the value on the left will be greater than on the right, so relative time on the left should increase. See text for details.

specifying that the distribution of behavior will, say, match the distribution of obtained reinforcers, the model specifies whether or not behavior will tend to shift (and if so, in what direction). This allows the model to make predictions regarding situations that involve unstable, neutral, and stable equilibrium (e.g., Vaughan, 1981; Vaughan & Herrnstein, in press), a property that is usually only implicit in most other models.

As mentioned above, the only difference between simple concurrent VI schedules and concurrent-chain schedules is that in simple concurrents, reinforcers are produced immediately by certain responses during the choice phase, but in concurrent chains a signaled delay of some sort is interposed between those choice responses and the availability or delivery of reinforcement. Figure 6 shows a schematic representation of the concurrent-chains procedure. The two initial links are concurrently available, and a certain proportion of time will be spent on each, as determined by the choice responses. Time spent on each initial link gives rise, with some local rate, to the terminal link for that side (usually signaled by a change in key color), while the other side is darkened. On some schedule (e.g., VI), reinforcement is then made available. Following this, the initial links are usually reinstated. In the model discussed above, during the choice phase each side has a value that is a function of the rate of transititon to reinforcement, as well as of the value of reinforcement itself. A straightforward generalization of this model would involve the interposition of another such function between the choice phase and reinforcement, to parallel the terminal link schedule shown in Figure 6. In this way, the value of a terminal link would be a function of both the rate at which food was presented while it is in force and the value of food. The value of an intital link, in turn, would be a function of both the rate at which the terminal link was presented and the value of that terminal link.

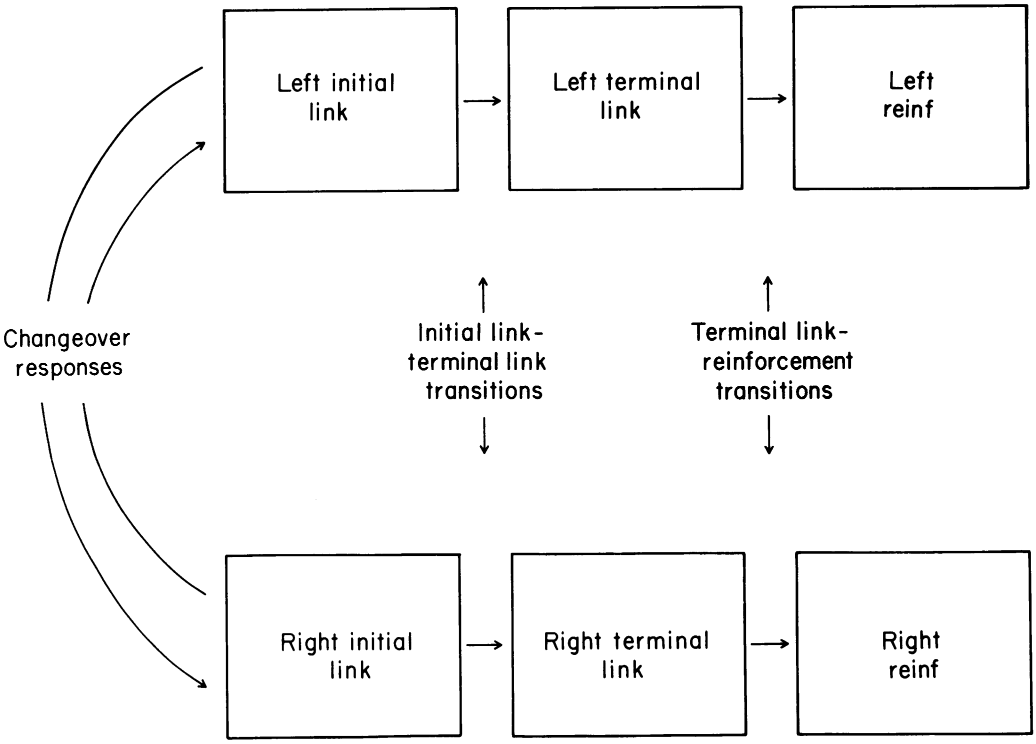

Fig. 6. Schematic representation of the concurrent-chains procedure. The left and right initial links are concurrently available; terminal links are mutually exclusive. Responses on each initial link produce transitions to the respective terminal link with some frequency; each terminal link will then give rise to reinforcement with some frequency. See Appendix 4 for a discussion of the effect of events following reinforcement.

Recall that a hyperbola (Equation 7) was used to model the value of a stimulus as a function of both the rate of transition to a second stimulus and the value of that second stimulus. If it is assumed that the unconditioned value of stimulus 1 (V'_1) need not be zero, Equation 9 is used:

$$V_1 = \frac{r(V_2 - V'_1)}{r + a} + V'_1. \tag{9}$$

(See also Appendix 2.) According to this equation, V_1 (the conditioned value of Stimulus 1) is bounded by V'_1 (the unconditioned value of Stimulus 1) as r approaches zero, and by V_2 (the conditioned value of Stimulus 2) as r grows large. (V_2, in turn, is calculated in an analogous manner, being bounded by V'_2 and V_3.)

Figure 7 shows, for two values of a, a number of graphs of Equation 9. On the left the equation is plotted in terms of r, rate of transition to the next stimulus; on the right it is plotted in terms of $1/r$, or duration of the stimulus (in seconds). For the upper two functions, V_2 is greater than V'_1; for the lower two functions, V_2 is less than V'_1. In the case of the upper two functions, a value of a equal to 0.2 generates a curve below that for a equal to 0.1. In what follows, the former will be used for FI terminal links, and the latter for VI terminal links. This is one of a number of ways to generate a preference for VI over FI terminal links, given equal rates of reinforcement (Herrnstein, 1964).

Figure 8 makes explicit the way in which the value of an initial-link stimulus is assumed to be functionally related to the value of a terminal-link stimulus, which in turn is assumed to depend on the value of reinforcement (V_3) as well as on the rate at which reinforcement occurs (top: r_2), or, equivalently, on the duration of the terminal-link stimulus (bottom: $1/r_2$). V_3, the value of reinforcement at the moment it is presented, constitutes the asymptote (i.e., the maximum possible value) for the value of the terminal link. With r_2 equal to 1800 transitions per hour (or 0.5 transitions per second), V_2 is the value of entering the terminal link. This in turn constitutes the asymptote for the value of the initial link. It is immaterial whether rate (top) or duration (bottom) is employed.

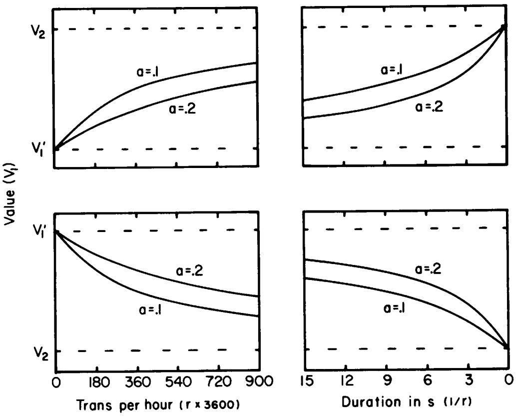

Fig. 7. Plots of Equation 9. Top: $V_2 > V'_1$. Bottom: $V_2 < V'_1$. Left: as a function of r (rate of transition). Right: as a function of $1/r$ (duration). In each case two values of a are shown, 0.1 for the case of VI schedules and 0.2 for the case of FI schedules.

The value of reinforcement as a function of its duration will also fit within this framework. Let the unconditioned value of food (the value of food at the moment it is presented, given that it will be on arbitrarily long) be positive; for convenience, a value of 10 will be employed here. Further, let the value of the keylight following food be zero. Then the

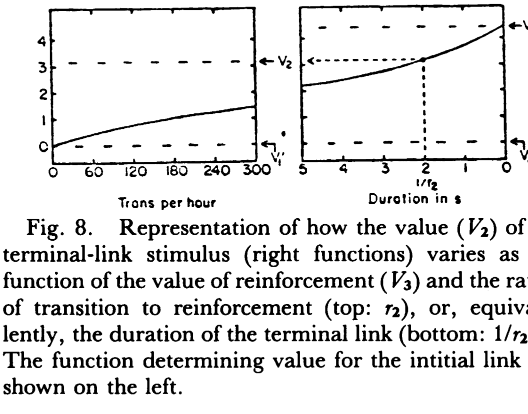

Fig. 8. Representation of how the value (V_2) of a terminal-link stimulus (right functions) varies as a function of the value of reinforcement (V_3) and the rate of transition to reinforcement (top: r_2), or, equivalently, the duration of the terminal link (bottom: $1/r_2$). The function determining value for the intitial link is shown on the left.

lower right portion of Figure 7, with V'_1 equal to 10, V_2 equal to zero, and a equal to 0.2 represents the value of food as a function of its duration (a will equal 0.2 if food is to be kept on for a fixed duration). For example, if food is on for a duration of 4 s ($1/r_3$), it will have a value of 4.44 (V_3), as shown in Figure 9. Note that nothing is being said about the value of terminal-link stimuli, or reinforcers, after their onset. Halfway through 4 s of food, for example, it is not being said that the momentary value of food is equivalent to the onset of 2 s of food. Only the values of these stimuli at their onset is being addressed. With regard to initial-link stimuli, however, the assumption is made that value is approximately constant over time, because reinforcement (or the transition to a terminal link) is approximately equiprobable at each moment of time. These considerations imply that the model applies only to concurrent VI VI initial links, and not to VI FI or FI FI initial links.

Figure 10 shows the complete model, for the case of concurrent VI 60-s initial-link schedules, leading to an FI 2 s (left) and an FI 4 s (right). Function A shows the value of food on the left (V_{3L}) as a function of its duration ($1/r_{3L}$), in this case 4 s. V_{3L} is then the asymptote of the left terminal link (Function B) plotted in terms of duration. With a left terminal link consisting of an FI 2 s ($1/r_{2L}$), V_{2L} represents the asymptote of the left initial link. Function C represents the value of the left initial link as a function of its rate of transition (r_{1L}) to the left terminal link. That value (V_{1L}) is then one of two inputs to the process of melioration. The rate of transition to the left terminal link, r_{1L}, depends on the left initial-link schedule (Function D) in conjuction with the distribution of time on that schedule (0.4 in this case).

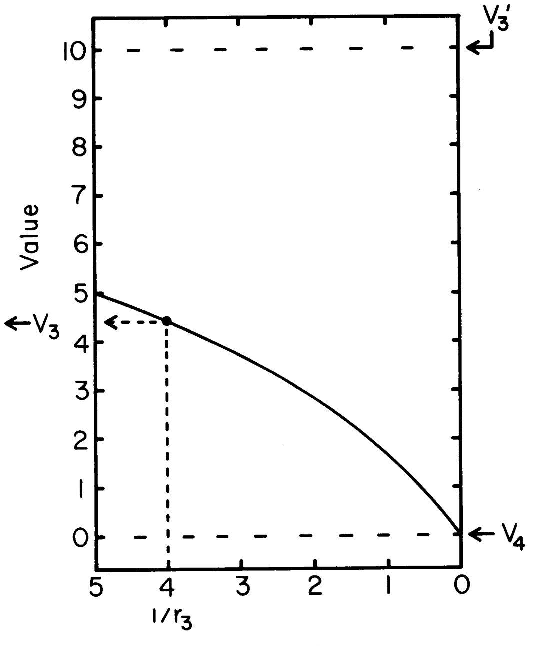

Fig. 9. The value of reinforcement (V_3) as a function of its unconditioned value (V'_3), its duration ($1/r_3$), and the value of the following stimulus (V_4), here assumed zero.

On the right, reinforcement is 4 s in duration ($1/r_{3R}$). G shows the value of the right initial link, which depends on the rate of transition (r_{1R}) from that stimulus to the right terminal link. The right initial-link schedule (H) and relative time on the right (0.6) in turn determine r_{1R}.

The values of the left and right initial links feed into the process of melioration (I), whose output is an arrow showing the rate and direction of change in relative time on the left (t_L). Under the conditions shown, relative time on the left will increase. Equilibrium will be reached when relative time on the left equals 0.58. At this point the local rate of reinforcement on the left (rate of transition from the initial link to the terminal link) will be less than that on the right ($60/0.58 = 103$ versus $60/0.42 = 143$ per hour). These unequal rates of reinforcement are such that they just balance the unequal values of the terminal links, so that the two initial links are equal in value ($V_{1L} = V_{1R}$). At this point the two criteria that are necessary and sufficient for stable equilibrium are met: The initial links are equal in value, and deviations from that point set up conditions tending to drive behavior back to that point. Note that this abstract form of matching is a theorem, rather than an axiom, within the present model.

Given concurrent linear VI schedules, it is possible to represent analytically relative time on the left at the point where behavior is stable, as a function of overall transitions per

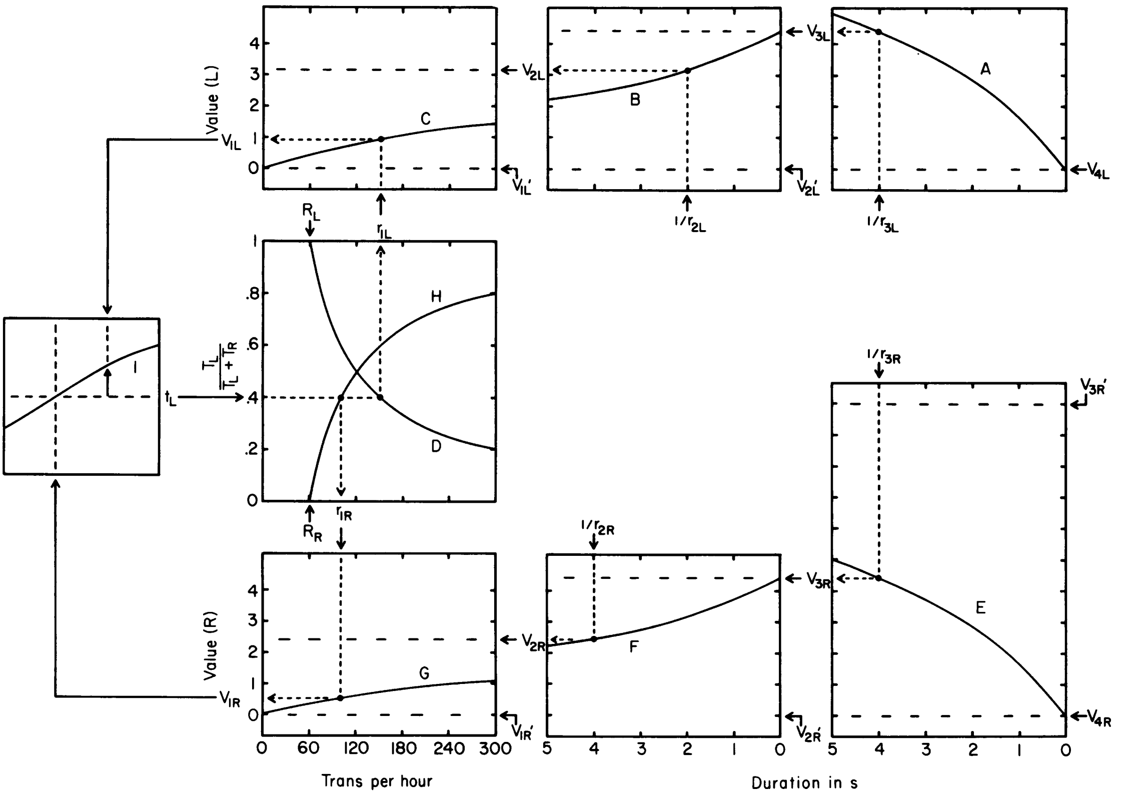

Fig. 10. Analysis of a concurrent-chains procedure, assuming on the left key (top functions) 4 s of reinforcement (A), an FI 2-s terminal link (B), and a VI 60-s initial link (D), giving rise to a value for the left initial link (C). On the right key (bottom functions) is shown 4 s of reinforcement (E), an FI 4-s terminal link (F), and a VI 60-s initial link (H), giving rise to a value on the right initial link (G). Given 0.4 relative time on the left, the present model specifies that relative time on the left will increase (I). See text for details.

second programmed on the left (R_L) and right (R_R) concurrent schedules, and the values on the left (V_{2L}) and right (V_{2R}) to which transitions are made:

$$\frac{T_L}{T_L + T_R} = \frac{R_L R_R(V_{2L} - V_{2R}) + aR_L V_{2L}}{a(R_L V_{2L} + R_R V_{2R})}. \quad (10)$$

(This equation is derived in Appendix 3.) In the present case, with concurrent VI 60-s initial links, $R_L = R_R = 0.0167$ (60/3600) transitions per second. $V_{2L} = 3.17$ and $V_{2R} = 2.47$. Substituting these values into Equation 10 gives $T_L/(T_L + T_R) = 0.58$.

It can be seen from Figure 10 that the present model is both very general (in the sense of being unambiguously applicable to a wide range of experiments) and conceptually very simple. Suppose, for example, that one terminal link consisted of two signaled FI schedules leading to food (chained FI FI). It would be necessary only to insert another function into Figure 10. Similarly, a simple concurrent experiment would involve removing Functions B and F. The same result is approached continuously if $1/r_{2L}$ and $1/r_{2R}$ are made to approach zero continuously: There is no discontinuity upon going from a nonzero to a zero terminal link. Functions A, B, C, E, F, and G are all plots of Equation 9.

Notice that the values of the stimuli following food (V_{4L} and V_{4R}) have been left at zero. Technically, these should be set to the values of the keylights during the initial link. At the level of resolution attempted here, however, doing so would serve only to complicate analyses without significantly affecting the qualitative predictions (see Appendix 4).

To summarize, the model is composed of only two equations (Equations 8 and 9), along with a set of subsidiary assumptions (e.g., that the unconditioned value of a keylight is zero). The use of linear VI schedules is not necessary, but serves to simplify the mathematics. Various experimental situations may be modeled by means of the appropriate concatenation of Equation 9, followed by the application of Equation 8 (or, equivalently, Equation 10). The simplicity of the model derives directly from the fact that only local events affect behavior, or the values of stimuli; the effect of delayed reinforcement is assumed to be mediated by conditioned reinforcers. The primary question remaining, then, is how the model compares with others in addressing actual data.

IMPLICATIONS

For several areas of research, experimental findings will be discussed, the workings of the present model will be explicated graphically, and predictions derived from the present model along with those derived from the previous models mentioned in the introduction will be examined. Finally, three as yet untested implications of this model will be mentioned.

In making predictions from Davison and Temple's (1973) model, their parameter E was left at one. For the Squires and Fantino (1971) model, the suggestion of Navarick and Fantino (1976) was incorporated: If reinforcer durations in the two terminal links differed, and those terminal links were nonzero in duration, the terminal link leading to the longer reinforcer was multiplied by the ratio of the smaller to the larger reinforcer. For Killeen's (1982) model, his Equation 13 was incorporated to account for duration of reinforcement. In additon, his parameter q was left constant at 0.125, and no bias term was employed. No attempt has been made to apply any model to situations for which it was obviously not designed.

Although leaving all parameters fixed may be judged inappropriate in some cases, doing so is equivalent to asking whether a given model could make accurate qualitative predictions for an as yet untested situation. This is consistent with the traditional focus on prediction and control that one finds within the experimental analysis of behavior. An alternative exercise, not undertaken here, would involve incorporating analogous free parameters into each of the models and fitting them to some set of data.

FI Terminal Links Kept at a Constant Ratio

The first area to be considered is that of concurrent-chains experiments with FI terminal links kept at a constant ratio. Given an FI terminal link, the reciprocal of the FI value is equivalent to immediacy of reinforcement, which is one of the variables to which Chung and Herrnstein (1967) suggested that birds match. As mentioned earlier, Baum and Rachlin (1969) built this matching assumption into their model. One implication of this assumption is that, given a concurrent-chains experiment with FI terminal links kept at a constant ratio (e.g., 1 to 2), preference should remain constant as the absolute size of the terminal links is varied.

This approach faces two problems, one logical and one empirical. Consider a concurrent-chains experiment with VI 60-s initial links, an FI 5-s left terminal link and an FI 10-s right terminal link, each leading to 4 s of reinforcement. Matching to immediacy of reinforcement implies a 2 to 1 preference for the shorter terminal link, even if each terminal link is cut in half (inasmuch as their ratio remains constant). But the limit of cutting the terminal links in half is a simple concurrent VI 60-s experiment, which should give rise to indifference. Logically, then, matching to immediacy must address the transition from nonzero to zero terminal links kept at a constant ratio.

On the empirical side, two experiments (MacEwen, 1972; Williams & Fantino, 1978) report an increase in preference for the shorter of two FI terminal links as both increase, given that they are kept at a ratio of 1 to 2. In both cases, the increase was negatively accelerated. Thus, as applied to a concurrent-chains experiment with equal initial links and FI terminal links kept at a constant ratio but varied in absolute durations, a model should predict indifference at zero terminal-link durations and some increase in preference for the shorter FI terminal link as both increase.

In the discussion of Figure 10, it was shown that, with concurrent VI 60-s initial links, an FI 2-s terminal link leading to 4 s of reinforcement on the left, and an FI 4-s terminal link leading to 4 s of reinforcement on the right, the present model predicts that relative time on the left should be 0.58. Figure 11 shows the results obtained if the FI values are systematically varied with their ratio at 1 to 2—for the present model and for the four others under consideration. Three models (Davison & Temple, Squires & Fantino, and the current one) all begin at indifference given zero terminal links and show a negatively accelerated increase in preference for the shorter terminal link as both terminal links increase. The Baum and Rachlin model, as mentioned above, shows a constant preference for the shorter terminal link (the discontinuity at zero is not shown). Killeen's model does not converge on indifference as the terminal links approach zero; furthermore, over part of its range it predicts a decrease in preference for the shorter terminal link as the terminal links increase in duration. On the face of it, then, both the Baum and Rachlin and the Killeen models are lacking with regard to both logic and data.

Killeen (1982) was able to account for MacEwen's data by assuming, in part, that his birds had exhibited bias. There is, however, no independent evidence for such bias (it would be straightforward to rerun the experiment incorporating such a test). Rather, as Figure 11 suggests, Killeen's model requires the use of free parameters under the present circumstance to avoid making predictions that are both counterintuitive and do not accord with known data.

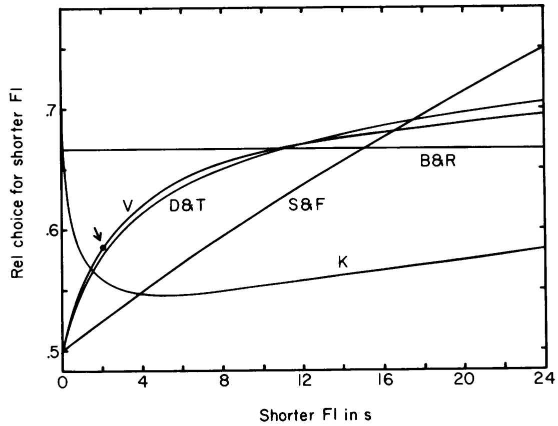

Fig. 11. Predictions of five models for a concurrent-chains procedure in which concurrent VI 60-s initial links lead to FI terminal links that are varied but kept at a ratio of 1 to 2. Relative choice for the shorter FI terminal link is shown as a function of the duration of that terminal link. The arrow indicates the predicted equilibrium distribution for the special case shown in Figure 10. B&R = Baum & Rachlin; D&T = Davison & Temple; K = Killeen; S&F = Squires & Fantino; V = Vaughan.

Self-Control

The second area of research to be discussed is that related to self-control, a topic addressed by a number of models of choice. If an organism is presented with a choice between a small reinforcer delivered immediately and a larger reinforcer delayed for some period, preference for the smaller reinforcer ("impulsiveness") may be exhibited. As equal delays are added to both reinforcers, preference may shift from the smaller to the larger reinforcer ("self-control"), the shift being termed preference reversal. In animals, preference reversal has been exhibited using both direct choice procedures (Ainslie & Herrnstein, 1981; Green, Fisher, Perlow, & Sherman, 1981; Navarick & Fantino, 1976) and "commitment" procedures, which allow the subject to commit itself to one option before the choice is due to be presented (Ainslie, 1974; Rachlin & Green, 1972).

It is sometimes said that preference reversal requires gradients of reinforcement "more concave" than an exponential function (Ainslie, 1975, for example, uses this term). This description, however, is neither mathematically rigorous nor focused on what is essential about gradients that will produce preference reversal. At the point at which preference is reversed, the reinforcement gradients must have the same height but unequal derivatives, with the function closer to reinforcement having the greater derivative. Thus the derivative must not be a function solely of the height of the function. It is sufficient, for example, for the derivative to be an increasing function of the height of the function and a decreasing function of the intercept (when plotted in terms of $1/r$). Expo-

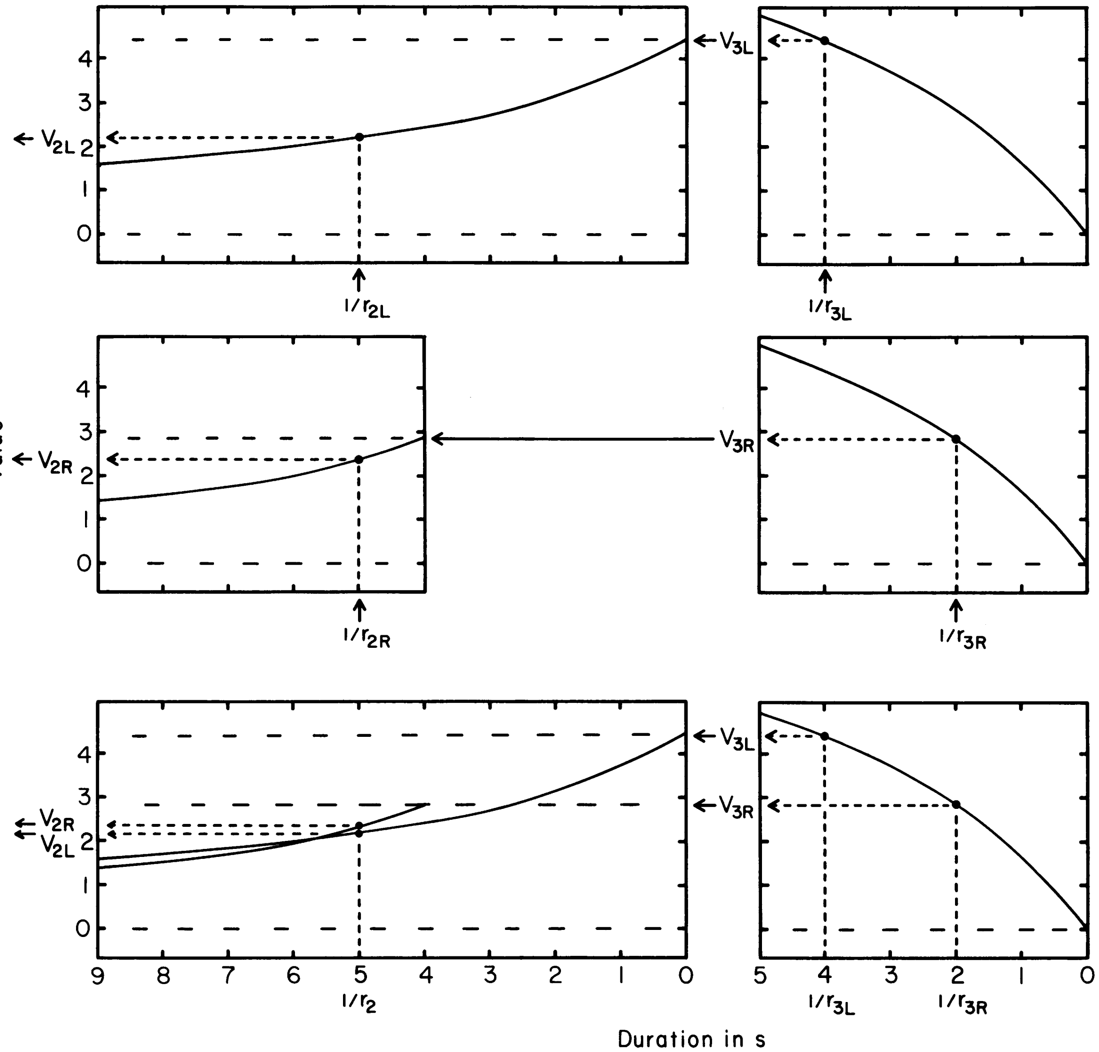

Fig. 12. Top pair: Representation of an FI 5-s terminal link on the left key leading to 4 s of reinforcement. V_{2L} is the value of entering that terminal link, and V_{3L} is the value of reinforcemnent. Middle pair: An FI 1 s on the right key leading to 2 s of reinforcement. V_{2R} is the value of entering that terminal link, and V_{3R} is the value of reinforcement. Bottom pair: Superimposition of the upper two pairs of functions, showing that the value of the terminal links will reverse as a constant amount is added to each FI terminal link.

nential functions are such that the derivative is a function solely of the height of the function, but hyperbolic functions have derivatives that are an increasing function of the height and a decreasing function of the intercept.

The analysis of preference reversal in terms of the present model is shown in Figure 12. The top pair of panels represents the left terminal link, including reinforcement. Reinforcement duration of 4 s ($1/r_{3L}$) produces value V_{3L}, the asymptote (or intercept, as plotted here) for the left terminal link. If an FI 5-s terminal link is in effect ($1/r_{2L}$), the value upon entering that terminal link will be V_{2L}, which in turn is the asymptote for the left initial link.

The middle pair of panels represents the right terminal link. Reinforcement duration is 2 s ($1/r_{3R}$), producing value V_{3R}. The right terminal link shown is an FI 1 s ($1/r_{2R}$), so that the value upon entering the right terminal link is V_{2R}. By shifting the right terminal-link function 4 s to the left in the figure, it is possible to show what happens in general with an FI x s on the right and an FI $x + 4$ s on the left. This is clearer if the two functions are superimposed, as shown in the bottom lefthand panel. With an FI 5 s leading to 4 s of reinforcement versus an FI 1 s leading to 2 s of reinforcement, the FI 1 s will be preferred (i.e., V_{2R} is greater than V_{2L}). As $1/r_2$ increases, however, the two functions cross (at about FI 2 versus FI 6), so that further increases produce a preference for the 4-s reinforcer over the 2-s reinforcer. In other words, the model predicts a shift from impulsiveness to self-control under these conditions. Where the functions cross, their heights are of course the same. The higher derivative for the right side then depends solely on the lower intercept (V_{3R}) for that side as compared with the other side (V_{3L}).

The predictions of the present and four other models under discussion are shown in Figure 13. Here, the initial links are concurrent VI 60-s schedules, and the terminal links are signaled FI schedules. The smaller reinforcer in 2 s and the larger is 4 s. Preference for the shorter reinforcer is shown as a function of the delay to that reinforcer, given that the delay to the longer reinforcer is 4 s greater. The only model not giving rise to preference reversal is that of Davison and Temple. Of the others, both Baum and Rachlin, and Killeen, predict exclusive preference for the 2-s reinforcer at a zero delay; the current model and that of Squires and Fantino predict nonexclusive preference. Experimental data do not yet exist to decide between these two possibilities. Although the preferences predicted by the current model are not great, in theory they could be increased by making the initial links richer (see below).

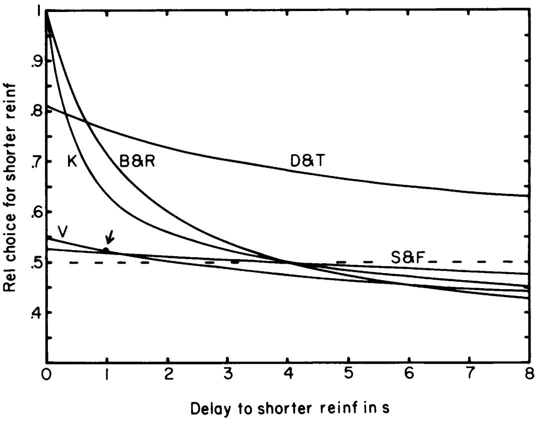

Fig. 13. Predictions of five models for a concurrent-chains procedure in which concurrent VI 60-s initial links lead to an FI x (with 2 s of reinforcement) and an FI $x+4$ (with 4 s of reinforcement), where x is the delay to the shorter reinforcer. Relative choice for the shorter reinforcer is shown as a function of the value of x. The arrow shows the predicted distribution for the special case shown in Figure 12, FI 1 s versus FI 5 s.

Tandem Versus Chained Terminal Links

The third area to be addressed is that of choice between tandem and chained FI FI terminal links. Partly on the basis of work of Gollub (1958), Fantino (1969b) hypothesized that, in a concurrent-chains experiment with terminal links of equal duration, preference should be shown for a terminal link consisting of a single color over one that changed color midway through (e.g., tandem versus chained terminal links). Duncan and Fantino (1972) reported a negatively accelerated increase in preference for FI x over chained FI $x/2$ FI $x/2$ as x increased in duration. Fantino (1977) noted that Wallace (1973) found a preference for fixed-time (FT) x over chained FT $x/2$ FT $x/2$, although the effect was smaller than that reported by Duncan and Fantino. Fantino (1983) determined that the effect depended upon stimulus characteristics, rather than response requirements, of tandem versus chained schedules.

Figure 14 (top two pairs of functions) shows the analysis of tandem versus chained terminal links. The top pair represents a tandem FI 4 FI 4-s terminal link (which is taken to be equivalent to an FI 8-s terminal link, inasmuch as there is no stimulus change; see Fantino, 1983). V_{4L} represents the value of reinforcement on the left, and constitutes the asymptote (or intercept) for the second FI 4-s schedule (which begins as $1/r_{3L}$). The height of the function at this point (V_{3L}) is the final height for the first FI 4, but the mathematical asymptote for the first FI is the same as for the second, V_{4L}. It is this latter property that leads to the equivalence between tandem FI $x/2$ FI $x/2$ and FI x.

The middle pair of functions shows a chained FI 4 FI 4. Again, V_{4R} represents the value of food, this time on the right. Here, however,

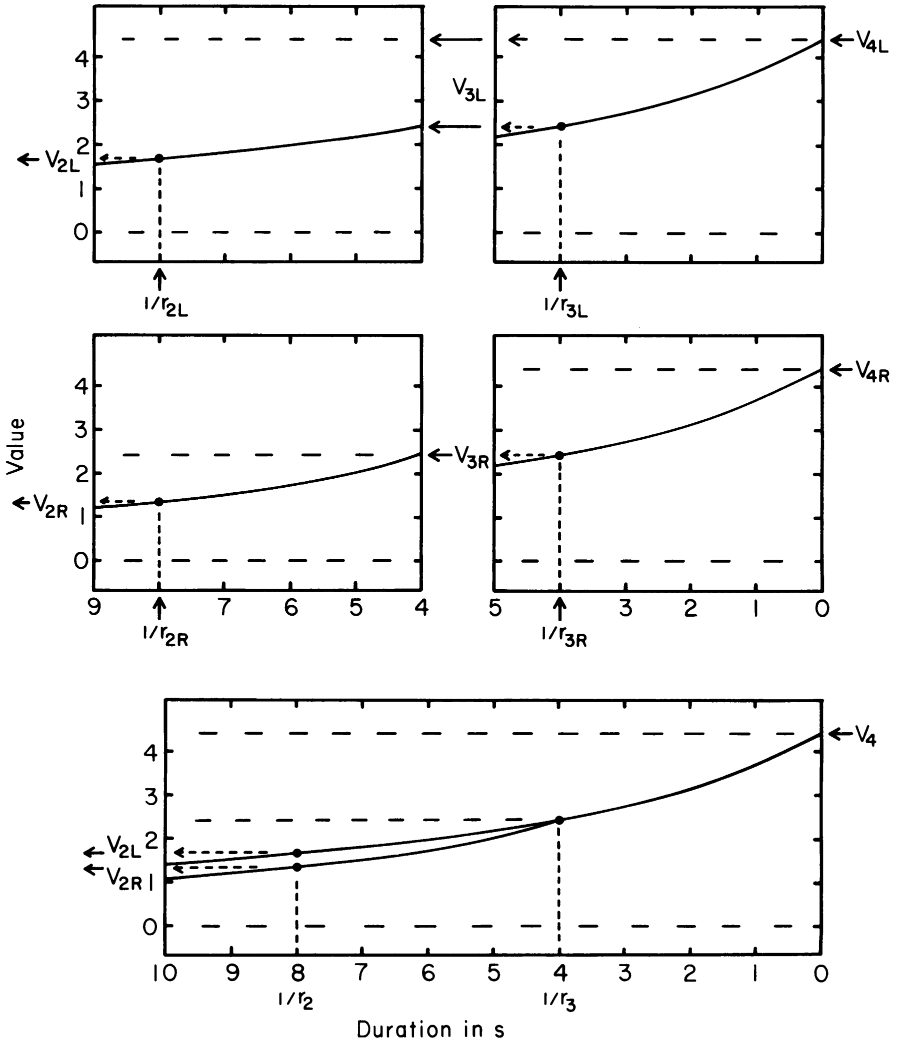

Fig. 14. Top pair: Representation of an FI 8-s terminal link (or a tandem FI 4 s FI 4 s) on the left key. V_{4L} is the value of presenting 4 s of food. Middle pair: A chained FI 4 s FI 4 s on the right key, likewise leading to 4 s of food. Bottom pair: Superimposition of the upper two pairs of functions, showing that the value of the tandem terminal link will be greater than that of the chained terminal link (i.e., V_{2L} lies above V_{2R}).

V_{3R} constitutes the final height of the first link, as well as the asymptote of that function. The way in which this generates a preference for tandem over chained can be seen more clearly in the bottom pair, where the functions have been moved together and superimposed. V_{2L} is the value entering the tandem FI FI, and V_{2R} the value of entering the chained FI FI. Note that at $1/r_3$ the heights of the two functions are the same, but the function for the chained has a greater derivative to the left than the function for the tandem schedule. This is the same property, discussed earlier, that allows for preference reversal—the intercept for the initial link of the chained schedule (V_{3R}) is lower than for the tandem schedule (V_{3L} or V_{4L}). In terms of the present model, preference reversal and preference for tandem over chained schedules each derive from the same properties of this function.

The predictions of the two models under consideration designed to deal with tandem versus chained terminal links are shown in Figure 15, assuming concurrent VI 60-s initial links and 4 s of reinforcement. (No provision for tandem versus chained schedules is made

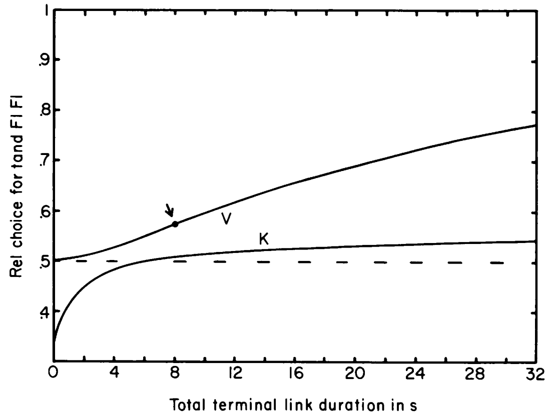

Fig. 15. Predictions of two models for a concurrent-chains procedure in which concurrent VI 60-s initial links lead to tandem FI FI versus chained FI FI. Relative choice for the tandem side is shown as a function of the total terminal-link duration. The arrow shows the predicted distribution for the special case shown in Figure 14, tandem FI 4 s FI 4 s versus chained FI 4 s FI 4 s.

by the models of Baum and Rachlin, Davison and Temple, or Squires and Fantino.) Killeen's model predicts an increase in preference for the tandem side as both terminal links increase in duration, but over part of the range his model predicts a preference for chained over tandem; further, the preference for tandem over chain, when present, is very slight. The present model predicts an at first positively and later negatively accelerated increase in preference for tandem over chained, with indifference given zero terminal links (simple concurrent VI 60-s schedule). If we assume that Duncan and Fantino's birds were not exhibiting bias (so that at zero terminal links their choice responses would be at 0.5), they conform to this pattern of positive and negative acceleration of the change in preference.

With regard to Killeen's analysis, Figure 15 here differs from his Figure 9 in which he attempted to account for the data of Duncan and Fantino. In that figure his variable *q* was changed from the usual value of 0.125 to 0.24; in Figure 15 here, *q* is left at 0.125. This is another case, then, in which Killeen's model appears to require the use of free parameters in order to conform to data.

Fantino and Duncan (1972) reported results from an experiment related to the above analysis. Birds were exposed to schedules in which entering one terminal link did not turn off the other, concurrent, schedule. It was thus possible to encounter concurrent initial links, concurrent terminal links, or an initial link on one side and a terminal link on the other. One of the most striking results was that birds did not appear to average reinforcement rate on a key across components (which would produce matching of responses to reinforcers at all times). Rather, when facing an initial link on one side and a terminal link on the other, almost all responses were directed at the terminal link.

An analysis of this result from the present point of view is straightforward. The value of an initial-link stimulus is maintained by transitions to its terminal link, but the value of a terminal link is maintained by transitions to food. Given equal initial links and equal terminal links (as in Fantino and Duncan's experiment), the value of a terminal link will be substantially greater than that of an initial link, which should produce the trend exhibited by the data.

The analysis of chained terminal links can be related to several experiments in which more than one reinforcer was delivered per terminal-link entry. Fantino and Herrnstein (1968), for example, ran a concurrent-chains experiment with VI 15-s terminal links. On one side a single reinforcer was delivered before returning to the initial link; on the other side the number of reinforcers was varied from 1 to 10. An increase in the number of reinforcers produced a negatively accelerated increase in preference for that side.

Moore (1979) studied a condition in which an FI 30-s terminal link was run in conjunction with an FI terminal link that, following food, led to four more reinforcers on a second FI schedule. If the added FI schedules were short (e.g., 3 to 9 s each), an increase in preference was shown for the side with those reinforcers. If the FI schedules were longer, there was little or no effect.

Shull, Spear, and Bryson (1981) studied a situation in which birds could switch from a VI 120-s schedule to a 240-s component during which four reinforcers were delivered. The first occurred 30 s into the interval, and the last 210 s into the interval. The location of the other two was varied. In general, as the other two occurred closer to the beginning of the interval, rate of entering the 240-s component increased.

Moore suggested that reinforcers following the first enhance its value; Shull et al. suggested that each reinforcer contributed to response strength, although to a decreasing extent as it was further delayed. In terms of the present model, if one terminal-link component is followed by another, differently signaled terminal-link component, the first will have greater value than if it led back to the initial link (see Appendix 4). This analysis is similar to that of Moore, although some additional assumption would be required before the effect of adding components signaled by the same stimulus as the first could be accounted for.

UNTESTED PREDICTIONS

Three as yet untested predictions of the current model might also be examined. Consider the case of equal simple concurrent VI schedules, one leading to 2 s of reinforcement and one to 4 s of reinforcement. As the two VI schedules are varied (keeping them equal to each other), Baum and Rachlin predict a constant preference, matching to relative durations of reinforcement. The present model predicts such matching for a particular pair of VI schedules, and a shift toward exclusive preference for the longer reinforcer as both schedules grow rich (the other models under consideration do not address this procedure). These predictions are shown in Figure 16, which shows preference for the longer reinforcer as a function of the (equal) concurrent VI schedules.

This prediction is similar to one made in a slightly different context by Fantino (1969a). As mentioned earlier, he argued that given a concurrent-chains schedule with unequal terminal links (e.g., VI 30 and VI 90-s schedules), as equal initial links became richer, preference should shift toward the better terminal-link side; with sufficiently rich schedules, exclusive preference should be exhibited. A similar argument may be made here: Given unequal durations of reinforcement, as equal VI schedules become richer, preference should shift toward the longer reinforcer. As shown in Figure 16 (which is similar in form to Fantino's Figure 1), the present model implies such a shift, but the relevant experiment remains to be done. In terms of the present model, unequal durations of reinforcement are equivalent to unequal terminal links. (This may be seen by noting that Equation 10 relies on V_{2L} and V_{2R}; those quantities may be derived from the presentation of immediate reinforcement, or from the operation of a terminal link.) The present model is thus consistent with Fantino's position.

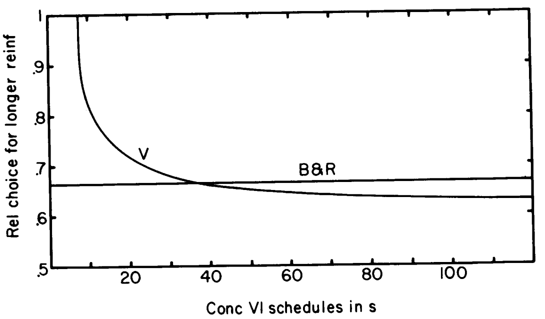

Fig. 16. Predictions of two models for a simple concurrent procedure in which equal VI schedules, which are varied, lead to reinforcer durations of 2 and 4 s. Relative time for the longer reinforcer is shown as a function of the VI schedules.

The property of the present model shown in Figure 16 also relates to part of an experiment reported by Fantino and Dunn (1983). Birds were presented with two concurrently available VI 90-s schedules, one leading to a VI 60-s terminal link and one to a VI 20-s terminal link. A third VI 90-s initial link, leading to a VI 9-s terminal link, was either present or absent, in alternate conditions. The main finding was that preference for the VI 20-s terminal link, relative to the VI 60-s terminal link, was greater in the presence of the third schedule than in its absence. This outcome is consistent with the delay-reduction hypothesis, but inconsistent with Luce's (1959) choice axiom.

The same qualitative results may be deduced from the present model: As above, unequal terminal links are treated as equivalent to unequal durations of reinforcement. Consider two concurrent VI 90-s schedules leading to different durations of reinforcement. Some preference will be exhibited for the longer reinforcer. Suppose, now, that less time is available for responding on those schedules. If, for example, only half of the time previously available is now available, the local rates of forcement on the concurrent schedules will rise to those normally produced by schedules twice as rich (i.e., concurrent VI 45-s schedules). As shown above, making equal concurrent schedules leading to different durations of reinforcement richer produces a greater preference for the longer reinforcer. The effect of adding the third schedule, given that it is sufficiently valuable to sustain responding, is simply that of removing some of the time available for the other two schedules. The results of Fantino and Dunn are thus consistent with the present model.

Baum and Rachlin assumed that the value of a reinforcer was proportional to its duration (and hence unbounded); the present analysis assumes that the value of a reinforcer is bounded (by its unconditioned value). One implication of boundedness is that, given concurrent VI 60-s schedules leading to reinforcer durations in a ratio of 2 to 1, as both durations increase preference should shift toward indifference. Figure 17 shows preference for the longer reinforcer under these conditions as both reinforcers increase in duration. Although the Baum and Rachlin model predicts matching to relative duration at all durations, the present model predicts such matching at left and right durations of 4 and 2 s, respectively.

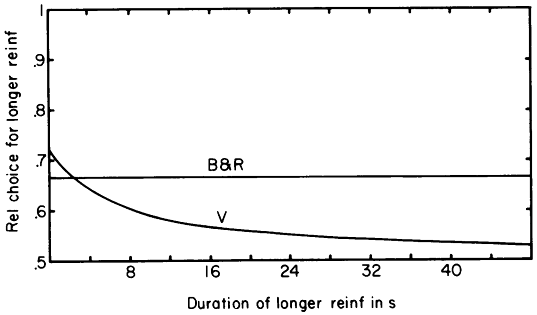

Fig. 17. Predictions of two models for a simple concurrent procedure in which concurrent VI 60-s schedules lead to reinforcer durations that are varied but kept at a ratio of 2 to 1. Relative time for the longer reinforcer is shown as a function of the duration of that reinforcer.

From what has been said so far, it can be seen that in a number of cases in which both the current model and that of Squires and Fantino make predictions, the two models are qualitatively similar. One implication of this similarity is that many of the experiments originally reported as confirming the delay-reduction hypothesis also serve to confirm the present model. It would, however, also be useful to find an experimental situation in which the two approaches made qualitatively different predictions.

Consider an experiment in which the two terminal links are VI 60-s schedules, and the right initial link is a VI 60-s as well. Figure 18 shows, for both the current model and that of Squires and Fantino, what happens as the left initial-link requirement is decreased from VI 60 s toward VI 0 s. Both models predict indifference, given concurrent VI 60 VI 60-s initial links. The current model predicts that choice responses will shift toward exclusive preference for the rich initial link; the Squires and Fantino model predicts an asymptotic distribution of behavior at 0.67. That is, two thirds of the initial-link responses would be directed to the side that provided immediate access to one terminal link, and one third would be directed to the side that provided terminal-link access only once a minute, overall. A more recent version of that model (Fantino & Davison, 1983) predicts that choice responses will shift

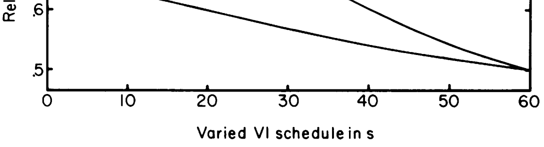

Fig. 18. Predictions of two models for a concurrent-chains procedure in which the right initial link is VI 60 s, the left initial link is varied, and both lead to VI 60-s terminal links. Relative time for the varied initial link is shown as a function of its value in seconds.

only to 0.59. The relevant experiment remains to be done.

SUMMARY

The generalization of melioration discussed here is in qualitative agreement with trends exhibited in a number of simple concurrent, as well as concurrent-chains, experiments. It is also a straightforward exercise to generate new predictions that discriminate between the present model and others. The presentation of reinforcement has no special status—the same framework applies to both neutral stimuli and positive reinforcers (as well as punishers, though this aspect has not been elaborated here; see Vaughan, in press). Finally, the values of initial-link stimuli are continuous functions of the schedule parameters, a presumably desirable mathematical constraint that is sometimes ignored.

The present model is not so much one model as it is one of the simplest of a class of models (see Appendix 2), each consisting of two equations, one relating to conditioned reinforcement (e.g., Equation 9) and one relating to the distribution of behavior (e.g., Equation 8). This class of models, in turn, is perhaps the simplest of a larger class employing more than two equations. For example, time in the presence of food, shock, and a keylight might require somewhat different equations. Such an eventuality would not necessarily challenge the core of the present model, but only the more specific assumptions. As these considerations suggest, the present paper is not meant to advocate a particular model of choice, but rather a particular framework. The formalism of that framework derives from the model of Rescorla and Wagner (1972); the basic tenets regarding pairing are consistent with the approach taken by Skinner in *The Behavior of Organisms*.

REFERENCES

Ainslie, G. W. (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior,* **21,** 485-489.

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin,* **82,** 463-496.

Ainslie, G., & Herrnstein, R. J. (1981). Preference reversal and delayed reinforcement. *Animal Learning & Behavior,* **9,** 476-482.

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

Boelens, H. (1984). Melioration and maximization of reinforcement minus costs of behavior. *Journal of the Experimental Analysis of Behavior,* **42,** 113-126.

Brownstein, A. J., & Pliskoff, S. S. (1968). Some effects of relative reinforcement rate and changeover delay in response-independent concurrent schedules of reinforcement. *Journal of the Experimental Analysis of Behavior,* **11,** 683-688.

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

Catania, A. C. (1966). Concurrent operants. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 213-270). New York: Appleton-Century-Crofts.

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

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. (1972). The psychological distance to reward. *Journal of the Experimental Analysis of Behavior,* **18,** 23-34.

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

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

Fantino, E. (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. (1983). On the cause of preference for unsegmented over segmented reinforcement schedules. *Behaviour Analysis Letters,* **3,** 27-33.

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

Fantino, E., & Duncan, B. (1972). Some effects of interreinforcement time upon choice. *Journal of the Experimental Analysis of Behavior,* **17,** 3-14.

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

Fantino, E., & Herrnstein, R. J. (1968). Secondary reinforcement and number of primary reinforcements. *Journal of the Experimental Analysis of Behavior,* **11,** 9-14.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

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

Green, L., Fisher, E. B., Jr., Perlow, S., & Sherman, L. (1981). Preference reversal and self control: Choice as a function of reward amount and delay. *Behaviour Analysis Letters,* **1,** 43-51.

Green, L., & Snyderman, M. (1980). Choice between rewards differing in amount and delay: Toward a choice model of self control. *Journal of the Experimental Analysis of Behavior,* **34,** 135-147.

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

Herrnstein, R. J. (1964). Aperiodicity as a factor in choice. *Journal of the Experimental Analysis of Behavior,* **7,** 179-182.

Herrnstein, R. J., & Heyman, G. M. (1979). Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *Journal of the Experimental Analysis of Behavior,* **31,** 209-223.

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

Herrnstein, R. J., & 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.

Hinson, J. M., & Staddon, J. E. R. (1983). Matching, maximizing, and hill-climbing. *Journal of the Experimental Analysis of Behavior,* **40,** 321-331.

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

Lea, S. E. G. (1976). Titration of schedule parameters by pigeons. *Journal of the Experimental Analysis of Behavior,* **25,** 43-54.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* New York: Wiley.

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. E. (1981). Optimization theory fails to predict performance of pigeons in a two-response situation. *Science,* **214,** 823-825.

Moore, J. (1979). Choice and number of reinforcers. *Journal of the Experimental Analysis of Behavior,* **32,** 51-63.

Morse, W. H. (1966). Intermittent reinforcement. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 52-108). New York: Appleton-Century-Crofts.

Myerson, J., & Hale, S. (1984, May). *A transition-state comparison of melioration and the kinetic model.* Poster presented at the meeting of the Association for Behavior Analysis, Nashville, TN.

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

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

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior,* **17,** 15-22.

Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. (1976). Economic demand theory and psychological studies of choice. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 129-154). New York: Academic Press.

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

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

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

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

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., & Motheral, S. (1978). On matching and maximizing in operant choice experiments. *Psychological Review,* **85,** 436-444.

Vaughan, W., Jr. (1976). *Optimization and reinforcement.* Unpublished doctoral thesis, Harvard University.

Vaughan, W., Jr. (1981). Melioration, matching, and maximization. *Journal of the Experimental Analysis of Behavior,* **36,** 141-149.

Vaughan, W., Jr. (1982). Choice and the Rescorla-Wagner model. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts* (pp. 263-279). Cambridge, MA: Ballinger.

Vaughan, W., Jr. (in press). Choice and punishment: A local analysis. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.). *Quantitative analyses of behavior: Vol. 5. Reinforcement value: The effect of delay and intervening events.* Hillsdale, NJ: Erlbaum.

Vaughan, W., Jr., & Herrnstein, R. J. (in press). Stability, melioration, and natural selection. In L. Green & J. H. Kagel (Eds.), *Advances in behavioral economics* (Vol. 1). Norwood, NJ: Ablex.

Vaughan, W., Jr., Kardish, T. A., & Wilson, M. (1982). Correlation versus contiguity in choice. *Behaviour Analysis Letters,* **2,** 153-160.

Wallace, R. F. (1973). *Conditioned reinforcement and choice.* Unpublished doctoral thesis, University of California, San Diego.

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.

Received November 9, 1982
Final acceptance February 19, 1985

Appendix 1

In Vaughan (1982) the presentation of reinforcement (a strengthening process) was assumed to change the value of the current stimulus by an amount proportional to the difference between the value of reinforcement and the value of the current stimulus, along the lines of Rescorla and Wagner (1972). In the absence of reinforcement the value of the current stimulus (i.e., a key currently being pecked by a pigeon) was assumed to change toward its unconditioned value at a rate proportional to the difference between its current value and its unconditioned value (a weakening process). From these two assumptions it could be inferred that at equilibrium the value of a stimulus was a strictly monotonically increasing function of the local rate of reinforcement (see Vaughan, 1982, Figure 12-2). The specific mathematical form of that function, however, was not discussed.

Here, a specific mathematical form (hyperbolic) is taken as the starting point. This is done on the assumption that some strengthening and weakening processes (though not necessarily strict proportionalities) could be found that would produce such a form at equilibrium.

Appendix 2

Although the hyperbolic form is used here, other functional forms should not be dismissed. In fact, the actual function could well prove not to be analytic. The hyperbola is one member of the following class: assuming $V_2 > V'_1$, and $0 < r < \infty$, the function is strictly monotonically increasing; $f(0) = V'_1$; $f(r) \rightarrow V_2$ as $r \rightarrow \infty$; the first derivative is not a function solely of the height of the function; the second derivative does not change sign.

In fact, there are two experiments (Green & Snyderman, 1980; Navarick & Fantino, 1976) that are inconsistent with Equation 9. Navarick and Fantino (Experiment 2) ran a concurrent-chains procedure with equal FI terminal links, one ending in 1.5 s of reinforcement and one in 4.5 s of reinforcement. As the two terminal links increased (from 5 to 40 s), preference for the larger reinforcement increased. Equation 9 predicts a constant preference. Green and Snyderman ran a concurrent-chains experiment with FI terminal links kept at a fixed ratio (6:1, 3:1, or 3:2), with the longer FI ending in 6 s of reinforcement and the shorter ending in 2 s of reinforcement. As the terminal links were varied (keeping their ratio constant), preference for the longer reinforcement increased in some cases and decreased in others. Equation 9 predicts that preference for the longer reinforcement will always decrease.

It turns out that Equation 9 embodies a certain invarience of scale that leads to the erroneous predictions. Let $d = V_2 - V'_1$, and suppose d is some positive value. Then for some value of r, V_1 will lie halfway between V'_1 and V_2. Now double the value of d. The same value of r will result in V_1 lying halfway between V'_1 and V_2. What is necessary in order to account for the above two experiments is that the value of r leading to $V_1 = (V_2 - V'_1)/2$ increase as d decreases. Because larger values of a lead to lower rates of increase of V_1, one solution would be for a to increase as d decreases. This may be accomplished by substituting for a in Equation 9 the quantity $a/[1-\exp(-|d|)]$. This change complicates the mathematics without changing the qualitative trends discussed below; therefore, it is not incorporated in what follows.

Appendix 3

Let $V_{1\mathrm{L}}$ and $V_{1\mathrm{R}}$ represent the values of the left and right terminal links, t_L and $1 - t_\mathrm{L}$ the relative times on the left and right initial links, and $V_{2\mathrm{L}}$ and $V_{2\mathrm{R}}$ the values upon entering the left and right terminal links.

From Equation 9:

$$V_{1\mathrm{L}} = \frac{(R_\mathrm{L}/t_\mathrm{L})V_{2\mathrm{L}}}{(R_\mathrm{L}/t_\mathrm{L}) + a};$$

$$V_{1\mathrm{R}} = \frac{[R_\mathrm{R}/(1 - t_\mathrm{L})]V_{2\mathrm{R}}}{[R_\mathrm{R}/(1 - t_\mathrm{L})] + a}.$$

At equilibrium the two sides of the initial links must be equal in value ($V_{1L} = V_{1R}$), so we have:

$$\frac{(R_L/t_L)V_{2L}}{(R_L/t_L) + a} = \frac{[R_R/(1 - t_L)]V_{2R}}{[R_R/(1 - t_L)] + a}.$$

By cross multiplying:

$$(R_L/t_L)V_{2L}[(R_R/(1 - t_L)) + a] = (R_R/(1 - t_L))V_{2R}[(R_L/t_L) + a].$$

This gives:

$$\frac{R_L R_R V_{2L}}{t_L(1 - t_L)} + \frac{aR_L V_{2L}}{t_L} = \frac{R_L R_R V_{2R}}{t_L(1 - t_L)} + \frac{aR_R V_{2R}}{1 - t_L}.$$

Multiplying all terms by $t_L(1 - t_L)$:

$$R_L R_R V_{2L} + aR_L V_{2L}(1 - t_L) = R_L R_R V_{2R} + aR_R V_{2R} t_L .$$

Or:

$$R_L R_R V_{2L} + aR_L V_{2L} - aR_L V_{2L} t_L = R_L R_R V_{2R} + aR_R V_{2R} t_L .$$

Collecting all terms with t_L on the right, and exchanging left and right:

$$aR_R V_{2R} t_L + aR_L V_{2L} t_L = R_L R_R V_{2L} - R_L R_R V_{2R} + aR_L V_{2L}.$$

Or:

$$t_L(aR_R V_{2R} + aR_L V_{2L}) = R_L R_R(V_{2L} - V_{2R}) + aR_L V_{2L}.$$

So:

$$t_L = \frac{R_L R_R(V_{2L} - V_{2R}) + aR_L V_{2L}}{a(R_L V_{2L} + R_R V_{2R})},$$

which is equivalent to Equation 10. Because the initial links are VI schedules, a will equal 0.1 in this equation.

Appendix 4

Before it is possible to take account of the values of stimuli following food, it is necessary to make an assumption regarding how the value of a stimulus is affected when followed by two or more concurrent stimuli with different values. The assumption that will be made here is that it is the stimulus with maximum value that affects the values of earlier stimuli. At equilibrium, given that the keys are equal in value, that maximum value is then the value of any of the keys.

It turns out there are two distinct cases that must be considered—that in which the durations of food are the same and that in which they are not the same. To begin with the first case, consider a concurrent-chains experiment with equal durations of food, whose values may be represented by V_3. Their values will always be equal, because their durations are equal, and the same stimulus (the stimulus with maximum value) is employed in calculating how their values change given transitions to the initial link. Thus, the value of the left terminal link (V_{2L}) is given by $r_{2L}V_3/(r_{2L} + a_2)$, and the value of the right terminal link (V_{2R}) is given by $r_{2R}V_3/(r_{2R} + a_2)$. Substituting these quantities into Equation 10 gives the following equation:

$$\frac{T_L}{T_L + T_R} = \frac{R_L R_R\left(\frac{r_{2L}V_3}{r_{2L}+a_2} + \frac{r_{2R}V_3}{r_{2R}+a_2}\right) + aR_L\left(\frac{r_{2L}V_3}{r_{2L}+a_2}\right)}{a\left[R_L\left(\frac{r_{2L}V_3}{r_{2L}+a_2}\right) + R_R\left(\frac{r_{2R}V_3}{r_{2R}+a_2}\right)\right]}.$$

This equation, in turn, is equivalent to the following:

$$\frac{T_L}{T_L + T_R} = \frac{R_L R_R V_3\left(\frac{r_{2L}}{r_{2L}+a_2} + \frac{r_{2R}}{r_{2R}+a_2}\right) + aR_L V_3\left(\frac{r_{2L}}{r_{2L}+a_2}\right)}{aV_3\left[R_L\left(\frac{r_{2L}}{r_{2L}+a_2}\right) + R_R\left(\frac{r_{2R}}{r_{2R}+a_2}\right)\right]}.$$

In this latter equation, the quantity V_3 can be eliminated altogether, provided it is nonzero, which implies that under these conditions t_L is independent of V_3. The general conclusion, then, is that given equal durations of reinforcement, neglecting the values of stimuli following food has no effect on the predicted equilibrium distribution of time.

In the case of unequal durations of reinforcement, attempts at an analytical expression for the equilibrium distribution of time have so far failed. The following approach appears to converge quickly (within about five iterations) at a satisfactory approximation. Using Equations 9 and 10, determine t_L on the assumption that the values of the stimuli following reinforcement are zero. Calculate the values of the initial-link at that distribution, and let that value be the values of the stimuli following food. Again use Equations 9 and 10 to redetermine the distribution of behavior, and so on.

In general, it appears that the effect of taking into account the stimuli following food, given unequal durations of food, is to increase the preference for the shorter of the two durations of food (i.e., mitigating the preference produced directly by the longer duration of food), with the effect being greater with shorter terminal links (and hence greatest with zero terminal links, or simple concurrent schedules). The general shapes of the predicted functions are not affected; where they approach asymptote or where they cross indifference does tend to be affected to some extent.