Psychological Review
1979, Vol. 86, No. 5, 496-500

Matching and Maximizing in Concurrent Schedules

Gene M. Heyman
Harvard University

In a recent article, Staddon and Motheral derived a mathematical model of responding in concurrent variable-interval – variable-interval schedules. According to this model, maximization of overall reinforcement rate predicts the operant matching law. However, Staddon and Motheral's derivation ignores a fundamental aspect of the concurrent schedule contingency. There are two simultaneously available schedules. A reinforcer, therefore, can occur in two ways, (a) following two consecutive responses on the same schedule, and (b) following a switch between the two schedules. Staddon and Motheral's model does not distinguish between these two possibilities, and this omission leads to an incorrect estimate of the expected reinforcement rate. In addition, the exact density function for obtained interreinforcement times in a concurrent interval schedule shows that matching and maximizing are different.

Research in operant psychology suggests that choice conforms to a simple algebraic symmetry. In experimental procedures that arrange the same type of schedule at each alternative, for example, concurrent variable interval – variable interval (VI–VI), the proportionate distribution of behavior between the two schedules generally approximates the distribution of reinforcements between the two schedules; whereas in procedures that arrange a different type of schedule at each alternative, for example, concurrent variable interval – variable ratio (VI–VR), the ratios of behavior between the two schedules systematically vary with the ratios of reinforcement. (See Baum, 1974, and de Villiers, 1977, for reviews.) This set of relations is referred to as the matching law (Baum, 1974; Herrnstein, 1970), and experiments show that matching has wide generality over species and procedures. For example, every species tested to date shows the predicted symmetry. However, despite extensive research, there is no consensus as to why subjects match responding to reinforcement (Staddon, 1977). In a recent article, Staddon and Motheral (1978) suggest a possible answer. They claim that in the concurrent VI–VI situation, "matching . . . can . . . be derived from reinforcement maximization . . ." (p. 436). The following comments review their argument.

Staddon and Motheral base their conclusion on a mathematical model of concurrent VI–VI performance. They derive an equation that gives the expected overall reinforcement rate as a function of the subject's choices between the two VI schedules. According to this equation, the maximum expected reinforcement rate is obtained at the choice proportions predicted by the matching law, so that in principle, the matching law can be derived from maximization. In a rather similar analysis, which Staddon and Motheral refer to, Heyman and Luce (1979) also calculated the maximizing choice proportions for concurrent VI–VI schedules. However, they concluded that maximizing overall reinforcement rate does not yield the matching law. The different findings result from Staddon and Motheral's assertion that in the concurrent situation, it is possible to represent responding on a given schedule by a single exponential distribution of interresponse times. A description of performance in concurrent VI–VI schedules shows that this assertion is not correct.

Figure 1 displays the sequence of events in concurrent VI–VI schedule performance.[1] Sides 1 and 2 refer to the experimentally arranged reinforcement sources: two VI schedules, with

I thank R. D. Luce, B. A. Williams, and H. J. Wilkinson for their helpful remarks on an earlier draft of this article.

Requests for reprints should be sent to Gene M. Heyman, Department of Psychology and Social Relations, William James Hall, Harvard University, Cambridge, Massachusetts 02138.

---

[1] For the reasons given by Staddon and Motheral, the changeover delay (Herrnstein, 1961) is not included in this account.

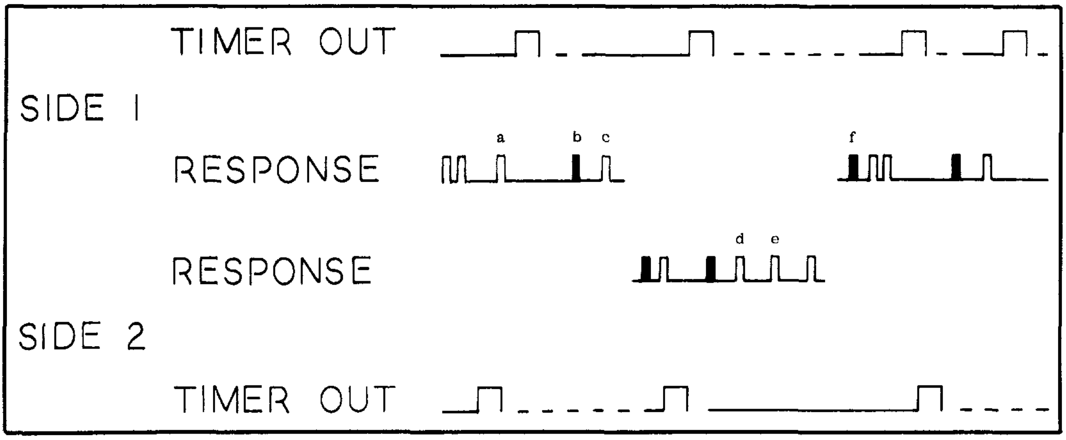

Figure 1. The sequence of events in a concurrent variable-interval–variable-interval schedule. (The second and third lines show that the subject switched from side 1 to side 2 and then back to 1. The spaces between the deflections indicate local interresponse times, *a* to *b* and *d* to *e*. The filled deflections are reinforced responses. The first reinforcement on side 1 followed a local interresponse time, because the timer set up while the subject was on side 1. The second reinforcement on side 1 followed an interchangeover time, because the timer set up while the subject was at side 2. In lines 1 and 4, a broken line indicates that a reinforcer is available and that the timer is not running. The point at which a broken line becomes continuous indicates that a response restarted the timer.)

associated discriminative stimuli and manipulanda for responses. Lines 2 and 3, labeled *response,* show that the subject switches back and forth between the sides. For each side, two classes of interresponse times can be distinguished. First, there are those that occur between consecutive responses on the same side, local interresponse times, such as *a* to *b* and *d* to *e*. Second, there are those that occur when a changeover to the other side intervenes between the responses, such as *c* to *f*. Next, the first and fourth lines show how the timers operate. They are independent of one another and are typically provided with an exponential distribution of intervals. For the duration of the interval, the timer operates independently of which side the subject is responding on, but when the interval is completed, the timer stops (indicated by the broken lines) and a reinforcer is set up and held until the next response on the associated side. The "correct" response produces a reinforcer (the filled deflections) and restarts the timer with a new random time. Depending on the side the subject is attending when the interval elapses, the reinforced response either follows a response on the same side or follows a changeover from the other side. In the first case, the delay from setup to reinforcement is related to the local response rate; in the second case, the delay is related to the changeover rate.

A word on the distinction between local response rate and overall response rate would be helpful here, for Staddon and Motheral do not refer to the former measure, whereas Figure 1 indicates that local response rate is a determinant of reinforcement rate. Local response rate is calculated by dividing the amount of time spent on a side into the number of responses there. In contrast, overall response rate is obtained by dividing the entire session time into the number of responses on a side. Both interchangeover times and local interresponse times, then, are averaged together to obtain the overall response rate.

Given this background, Staddon and Motheral's equation can be described and assessed. The equation involves two quantities: the programmed VI timer rates and the subject's overall response rates. First, the programmed interreinforcement times for a VI schedule are often approximately exponentially distributed. Staddon and Motheral assume this for their model. (Note that the scheduled times are the minimum possible interreinforcement intervals; the subject increases these times by not responding immediately to a schedule that is set up.) Next, Staddon and Motheral state that the overall rate interresponse times are exponentially distributed. But as described above, the overall rate is a mixture of interchangeover times between the schedules and local interresponse times on a given schedule. Therefore, if the assumption of random responding is maintained, the distribution of overall interresponse times is the addition of two exponentials. This mixture, it is easy to show, is not itself exponential. Indeed, it seems rather unlikely that the mixture of any two probability distributions could be exponential.

For a single distribution of interresponse times, Staddon and Motheral's assumption, the expected time between reinforcements for two VI schedules is

$$D_1 = 1/a + 1/x \text{ and } D_2 = 1/b + 1/y, \quad (1)$$

where D_i is the expected delay, a and b are the average respective programmed VI timer rates, and x and y are the average respective overall response rates.

Staddon and Motheral claim that Equation 1 calculates the expected reinforcement rate for a concurrent VI–VI schedule. The equation expresses the expected time to the next reinforcement on a side as the average timer interval plus the respective average overall interresponse time. But note that since the distinction between interchangeover times and local interresponse times has been obscured, Equation 1 asserts that time to reinforcement is independent of the side the subject is responding on when a timer sets up. In other words, Equation 1 does not distinguish between delays to reinforcement due to interchangeover times and delays due to local interresponse times. The following examples and mathematical model of Figure 1 demonstrate that this distinction is essential to calculating the expected reinforcement rate for a concurrent VI–VI schedule.

First, matching occurs in both continuous and discrete response procedures. For a continuous response requirement, say, standing on one or the other side of the experimental chamber (Baum & Rachlin, 1969), reinforcers are delivered immediately when they set up at the attended side but are held until the next changeover when they set up at the vacated side. For attended setups, then, there is virtually no delays, but for unattended setups, the delay until reinforcement depends on the interchangeover times, and the ratio of the average interchangeover times is identical to the ratio of the overall amounts of time spent on each side. For the discrete response procedure (e.g., a lever press), the above discussion implies that the properties of local responding need to be considered. At both sides the local response rates are approximately equal (Baum & Rachlin, 1969; Catania, 1966; de Villiers, 1977) so that delays to reinforcement that occur within a run of responses are virtually identical for each side and independent of the allocation of time and responses between the sides. In contrast, delays due to switching are the same as in the continuous procedure. Second, as suggested by Figure 1, the absolute duration of interchangeover and interresponse times differ. Local response rates are generally on the order of 60–120 per minute, whereas the switching rate is invariably less, generally on the order of about 3–15 per minute. These examples emphasize the consequences of the subject's location vis-à-vis a primed timer. The following mathematical model calculates these effects.

A model of Figure 1 can be derived directly from the mean delays (Heyman, 1977) or from the probability density function of interreinforcement times (Heyman & Luce, 1979). Both naturally observe the differences between interchangeover and local interresponse times, but the direct argument seems somewhat easier to condense.

Let r_i equal the mean local interresponse time on side i for exponentially distributed times. (For a continuous response requirement, r_i is virtually zero, and for a discrete response procedure, it is reasonable to assume that $r_1 = r_2$.) Let x_i equal the conditional probability of a switch from side i at time t since the last switch for exponentially distributed interchangeover times. This implies that the expected interchangeover time is $1/x_1$ for side 1 and $1/x_2$ for side 2. Given these definitions, it follows that the overall proportion of time spent on side 1 is $x_2/(x_1 + x_2) = p$. Or, in other words, the proportion of time spent on side 1 is equal to the relative rate of switching to that side. Next, let I equal $1/(x_1 + x_2)$, one half the harmonic mean of the interchangeover times (Heyman, 1977). The data show that this quantity is approximately constant (e.g., Heyman, 1979; Hunter & Davison, 1978), and it is assumed so here. Since p, the overall distribution of time between the schedules, is the behavioral variable of interest for matching and maximizing, it is convenient to make the following substitutions for the mean interchangeover times:

$$1/x_1 = \frac{1/(x_1 + x_2)}{x_1/(x_1 + x_2)} = I/(1 - p)$$

and

$$1/x_2 = \frac{1/(x_1 + x_2)}{x_2/(x_1 + x_2)} = I/p. \quad (2)$$

The quantity I, therefore, scales the overall tendency to switch. For example, at $p = .5$ and I equal to 10 sec, the average interchangeover times are both 20 sec, but for the same p and I equal to 25 sec, they are both 50 sec.

From Figure 1 and the preceding definitions, it follows (Heyman & Luce, 1979) that the expected reinforcement rate in a concurrent VI–VI schedule is

$$R_T = \frac{p}{V_1 + r_1} + \frac{1-p}{V_2 + r_2} + \frac{1-p}{V_1 + I/p} + \frac{p}{V_2 + I/(1-p)}, \quad (3)$$

where R_T is the overall expected reinforcement rate, the sum from the two available schedules, and in the denominators, V_i gives the average programmed timer intervals.

The four terms of Equation 3 represent the different ways reinforcement can occur. With probability p, the subject is at side 1 when its timer sets up, which occurs at a mean rate of $1/V_1$, and the response requirement, if there is one, adds a mean delay of r_1. With probability $1 - p$, the subject is at side 2 when the timer sets up, and there is a mean delay of I/p, until the subject switches back and receives the reinforcer. These are the two ways reinforcement can occur at side 1. The two other terms correspond to reinforcement at side 2.

The maximum reinforcement rate can be determined from Equation 3 by varying p. It is obtained at a value of p that closely approximates, but is not equal to, matching. This result eliminates reinforcement rate maximization as a logical condition for the matching law (given the assumptions that led to Equation 3), but because of the similarity of the two predictions, it does not eliminate maximization as a possible psychological mechanism in concurrent schedule performance. To test for psychological relevance here, it is necessary to consider procedures in which matching and maximizing predict measurably different outcomes. Two such examples have been examined.

In a variant of the standard concurrent VI–VI schedule described in Figure 1, the two timers are made interdependent. When one timer sets up, the other is stopped. An expected reinforcement rate equation for this procedure was derived (Heyman, 1977; Heyman & Luce, 1979), and it shows that the maximum reinforcement rate is obtained by a distribution of behavior that is about midway between equality (50%) and perfect matching. For example, in an interdependent concurrent VI 60-sec – VI 180-sec schedule, the maximum reinforcement rate is produced by behavior proportions of about 52–63%, depending on the absolute response rates. Yet subjects frequently approach perfect matching in this type of schedule (e.g., Stubbs & Pliskoff, 1969). An expected reinforcement rate equation was also derived for a procedure that arranges a variable-ratio schedule at one side and a variable-interval schedule at the other side (Herrnstein & Heyman, 1979). In this situation, matching and maximizing differ even more widely than in the interdependent schedule, yet subjects closely approximate response ratios to behavior ratios (Herrnstein & Heyman, 1979; and see below for some of the details).

To summarize the discussion so far, then, Staddon and Motheral's (1978) equation (Equation 1) does not correctly calculate the expected reinforcement rate for performance in a concurrent VI–VI schedule, and models of concurrent performance that assume an exponential distribution for changeover times and an exponential distribution for interresponse times show that matching and maximizing yield different predictions, with the obtained choice proportions more closely approximating matching.

Staddon and Motheral also misinterpret some previous theoretical accounts of maximizing in concurrent schedules. In summarizing Rachlin, Green, Kagel, and Battalio's (1976) computer simulation results, Staddon and Motheral (1978) commented that "matching can also be derived from molar maximizing without making any assumptions about switching" (p. 437). Rachlin et al., however, did define switching. They set it at 6 per minute and assumed that switching was independent of the subject's overall allocation of time between the sides. This model, incidentally, does not correspond to the data. Switching varies in an orderly way with time and response proportions, as described by Hunter and Davison (1978), Stubbs and Pliskoff (1969), and others. In referring to momentary maximizing theories (e.g., Shimp, 1969), Staddon and Motheral (1978) wrote: "Heyman and Luce have recently shown that a more realistic switching model, based on the observed pattern of random switching, fails to predict matching" (p. 437). This is misleading. We failed to predict that matching is a consequence of maximizing, and we were not testing a momentary maximizing theory, but, like Staddon and Motheral, we were concerned with the relationship between the overall allocation of behavior and reinforcement rate.

Staddon and Motheral refer to the Heyman and Luce article without attempting to reconcile the differences. The two approaches (Equations 1 and 3) predict different overall reinforcement rates in all but one situation. The exception is when the average interchangeover times and local interresponse times are equal. That is, if $r_1 = r_2$

and $r_1 = I/p$ and $r_2 = I/(1-p)$, then

$$[p/(V_1 + r_1) + (1-p)/(V_1 + I/p)],$$

the Heyman and Luce model, is equal to $1/(V_1 + x)$, the Staddon and Motheral model. However, this equality requires that the subject spend equal proportions of time on each side and switch after each response. This pattern of switching is not observed, and, in effect, Figure 1 and Equation 3 imply that Staddon and Motheral's theory does not describe concurrent performance, except when the subject's distribution of time between the two sides is approximately equal.

Staddon and Motheral conclude with a discussion of asymmetrical concurrent interval-ratio schedules. They observed that performance in this environment is not "congenial" to either matching or maximizing. However, in one set of studies (unpublished experiments by Herrnstein described in Baum, 1974; de Villiers, 1977), the slopes relating the ratios of responding to the ratios of reinforcement did not systematically differ from the matching prediction of 1.0, and the pooled data yield slopes of 1.04 (Herrnstein & Heyman, 1979). Moreover, the subjects (pigeons) spent most of their time on the VI schedule side, whereas maximizing predicts a qualitatively different outcome, preference for the ratio side; and a model similar to Equation 3 estimates that the pigeons gave up about 60 reinforcements per hour by matching rather than maximizing (Herrnstein & Heyman, 1979). Staddon and Motheral acknowledge these discrepancies, but suggest that more general optimality or regulatory analyses are not ruled out. Perhaps not, but their conjecture receives little support from experiments in which subjects at 80% of their normal body weight persist in matching, although higher reinforcement rates are available by simply redistributing the same responses so as not to match. Consequently, it seems at least equally plausible to conjecture that choice in concurrent schedules is constrained by factors that are independent of the overall reinforcement rate. For example, given the recent striking success of molecular level conditioning models in classical conditioning and discrimination learning (Rescorla & Wagner, 1972), it would not be surprising if the asymptotic choice proportions in free operant procedures were the consequence of molecular level response strengthening.

References

Baum, W. M. On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 1974, *22*, 231–242.

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

Catania, A. C. Concurrent operants. In W. K. Honig (Ed.), *Operant behavior*. New York: Appleton-Century-Crofts, 1966.

de Villiers, P. A. Choice in concurrent schedules and a quantitative formulation of the law of effect. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior*. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

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

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

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

Heyman, G. M. *Matching and maximizing in concurrent schedules.* Unpublished doctoral dissertation, Harvard University, 1977.

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

Heyman, G. M., & Luce, R. D. Operant matching is not a logical consequence of maximizing reinforcement rate. *Animal Learning and Behavior*, 1979, *7*, 133–140.

Hunter, I. W., & Davison, M. C. Response rate and changeover performance on concurrent variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, 1978, *29*, 535–556.

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

Rescorla, R. A., & Wagner, A. R. 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.* New York: Appleton-Century-Crofts, 1972.

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

Staddon, J. E. R. On Herrnstein's equation and related forms. *Journal of the Experimental Analysis of Behavior*, 1977, *28*, 163–170.

Staddon, J. E. R., & Motheral, S. On matching and maximizing in operant choice experiments. *Psychological Review*, 1978, *85*, 436–444.

Stubbs, D. A., & Pliskoff, S. S. Concurrent responding with fixed relative rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1969, *12*, 887–895.

Received March 20, 1979 ■