MAXIMIZING PRESENT VALUE: A MODEL TO EXPLAIN
WHY MODERATE RESPONSE RATES OBTAIN ON
VARIABLE-INTERVAL SCHEDULES

ALAN SILBERBERG, FREDERICK R. WARREN-BOULTON, AND
TOSHIO ASANO

THE AMERICAN UNIVERSITY AND KYOTO UNIVERSITY

In Phases 1 and 3, two Japanese monkeys responded on a multiple variable-ratio 80 variable-interval X schedule, where the value of X was adjusted to ensure equal between-schedule reinforcement rates. Components strictly alternated following the delivery of a food pellet, and each session ended following 50 components. Phase 2 differed from the others only in that the 50 pellets previously earned during the session were delivered together at session's end. Variable-ratio response rates did not decrease across phases, but variable-interval response rates decreased substantially during the Phase 2 procedure. This rate decrease was attributed to the food-at-session's-end manipulation removing the greater immediacy of reinforcement provided by short interresponse times relative to long interresponse times. Without this time preference for short interresponse times, the variable-interval interresponse-time reinforcement feedback function largely controlled response emission, dictating a response-rate reduction. This result was explained in terms of the economic notion of "maximizing present value."

Key words: maximizing, multiple variable ratio variable interval, yoked-control procedure, key press, monkeys

Peele, Casey, and Silberberg (1984) found that when each component of a multiple variable-ratio (VR) variable-interval (VI) schedule provided the same reinforcement rate, the VR component supported higher response rates than the VI component. They attributed this finding to differences in each schedule's molecular reinforcement feedback function: For the VI, probability of reinforcement grows as a function of time between successive responses (interresponse times or IRTs), whereas for the VR schedule, reinforcement probability does not vary with IRT size. They argued that response rates are lower on VI than VR because the VI's feedback function differentially reinforces long IRTs whereas the VR's does not.

This molecular account attributes behavioral output on VR and VI schedules to an optimization calculation made by the subject on the IRT feedback function. It can be readily compared with alternative approaches that claim VR and VI response rates are a conse-

quence of a maximizing calculation, made at a molar level, on the feedback function between response rate and reinforcement rate (e.g., Rachlin, 1978; Staddon, 1979). Because there is a one-to-one correspondence between the rate of responding and the rate of reinforcement on VRs, molar-maximizing accounts predict high rates of responding on these schedules. However, for VIs, the function for reinforcement rate over response rate is concave, resulting in diminishing returns to marginal increases in response rate. Hence, molar maximizing predicts lower response rates on VIs.

As a consequence of between-schedule differences in the form of the feedback function, molar accounts can predict the VR-VI rate difference. Nevertheless, Prelec (1982) has challenged these molar models because the rates maintained by VIs, although lower than VRs, still seem too high to be consistent with optimization principles. The problem is that animals can produce virtually the same VI reinforcement rate with lower response rates. The fact that VIs maintain moderate rates, despite the near uselessness of responding above a low rate, argues against attributing behavioral output to a molar-maximizing process.

Because response rates and mean IRT size are highly correlated, Prelec's (1982) complaint with molar-maximizing models natu-

This research, supported by a grant from the National Science Foundation, was completed during A. Silberberg's sabbatical stay at the Primate Research Institute of Kyoto University. He gratefully acknowledges the courtesies extended by the Chair, Professor Murofushi, and the technical and programming assistance of Mr. Nagumo. Address correspondence to Alan Silberberg, Department of Psychology, 321 Asbury Building, The American University, Washington, D.C. 20016.

rally extends to molecular accounts. As applied to the results of Peele *et al.* (1984), Prelec questions how an optimization account can explain the high frequency of short-IRT emission on VI schedules given that short IRTs have very low reinforcement likelihoods.

In this report, we address this question by modifying Peele *et al.*'s (1984) model so that the emission of short IRTs on VI schedules is consistent with maximizing principles. To do this, we posit that IRT selection is governed by two factors: (a) the rate-decreasing effects of the IRT feedback function, and (b) the rate-increasing effects of a preference for the greater immediacy of reinforcement provided by short IRTs. Behavioral output on VIs is a blend of these opposing processes, resulting in VI response rates lower than those seen on VR but well above those necessary to ensure a near-maximal reinforcement rate.

Evidence supporting the importance of immediacy of reinforcement in choice comes from Rachlin and Green (1972; see also Shimp, 1969). In their study, pigeons chose between a small, immediately available reinforcer and a large delayed reinforcer. Pigeons preferred immediate reinforcement, even though it resulted in a drop in the amount of reinforcement obtained. We believe the same impulsivity occurs in the relative frequency of IRT emission on VIs: When choosing between the high-probability, long-delay reinforcement provided by long IRTs and the low-probability, short-delay reinforcement provided by short IRTs, animals often behave impulsively, favoring the greater immediacy of reinforcement that attends short-IRT emission.

Peele *et al.* (1984) showed how differences in VI and VR IRT feedback functions could explain why VI response rates are lower than VR rates. Now we wish to show that time preferences for the greater immediacy of reinforcement provided by short IRTs elevate VI response rates above levels needed to ensure near-maximal reinforcement rates. We propose to demonstrate this by altering whether or not the opportunity to exercise time preferences is allowed to affect behavior in a VI schedule. To permit time preferences in VI schedules, one need only reinforce responding according to a VI schedule. To remove time preferences requires a different arrangement, in which short IRTs do not produce primary reinforcement sooner than long IRTs. Such a

VI schedule could be created by replacing each food pellet earned during a session with a light flash and delivering at session's end a sum of food pellets equal to the number of light flashes during the session. In this circumstance, the animal could "count" reinforcers during the session, but could not obtain them immediately. Now the only reason for emitting short IRTs is that they provide more immediate access to a light flash. On the assumption that impulsivity for this conditioned reinforcer will be weaker than that for food, we predict that VI IRT emission will be more strongly controlled by the rate-decreasing contingencies of the IRT feedback function. Thus, we predict that this manipulation will lower VI response rates from levels obtained when food pellets are earned throughout the session.

One problem with this design is that the predicted rate reduction may occur, caused not by a reduction in impulsivity in IRT selection but by a reduction in the value of the schedule. To the extent that response rates correlate with the value of a schedule, and value is depreciated by moving primary reinforcement to session's end, then the experiment presented above is confounded. Assuming rate decreases are observed, did they occur because of reducing time preferences among reinforced IRTs, or did they occur because moving reinforcement to session's end diminished the value of the schedule?

One solution to this problem is to make the proposed manipulation not only to a VI schedule, but also to a VR as a comparison schedule. As was true with VIs, moving VR food reinforcement to session's end should weaken reinforcer immediacy as a factor in IRT selection and strengthen control by the IRT feedback function. However, the strengthened operation of the IRT feedback function does not lead to a prediction of a rate change on VRs because this function defines no dependency between IRT size and probability of reinforcement. Given the absence of pressure from the IRT feedback function for a rate change, weakening the role of reinforcer immediacy in IRT selection should leave VR response rates unchanged.

If providing reinforcement at session's end reduces schedule value, both VR and VI response rates should diminish from levels obtained when food was given throughout the session. On the other hand, if this manipula-

tion only affects impulsivity in IRT selection on these schedules, we predict a differential outcome: no change in VR response rates and a decrease in VI response rates. We test these predictions in the experiment described below.

METHOD

Subjects

Two adult male Japanese macaques (*Macaca fuscata*, 2-T320 and 5-J519) that were part of the subject pool at the Primate Research Institute, Kyoto University, Japan, served as subjects. These monkeys had varied and extensive histories of experience on operant procedures. They began the experiment at approximately 90% of their free-feeding weights and were supplementally fed on those occasions when their body weights dropped 5%.

Apparatus

A 70-cm cube, with surfaces composed of Plexiglas (ceiling), steel bars (one side wall and floor), two steel side walls, and an aluminum intelligence panel (side wall), served as the experimental space. On the intelligence panel, the distances from the floor to the lower edge of a 10- by 5-cm (horizontal by vertical dimensions) food-cup aperture and the center of two 4-cm diameter response keys were, respectively, 20 cm and 40 cm. The food aperture was in the center of the vertical midline of the intelligence panel, and each response key (which required a force of 15 N to operate) was displaced 10 cm on each side of this midline. An exterior white-noise generator masked extraneous sounds, and a 40-W lamp over the Plexiglas ceiling provided general chamber illumination. Food pellets of Japanese manufacture, similar in composition to Purina monkey chow, were provided by one of two brush-operated rotary feeders (Davis UF-100). The weight of each pellet ranged from 3.2 to 4.0 g, and averaged 3.6 g. A water spout, located opposite to the intelligence panel, provided unrestricted access to a 500-mL water bottle. A microcomputer recorded all data and controlled all experimental events.

Procedure

In Phase 1, both subjects were exposed to a multiple VR 80 VI X schedule in which the value of X equaled the time it took to complete

the VR interreinforcement ratio that had occurred five components earlier. Because components strictly alternated, with VR accompanied by a right red key and VI by a right green key, this arrangement meant that the current VI component was yoked to the completion time of the VR component with a lag of three. Each daily session began with a VR component and ended after 50 components. The interreinforcement intervals for the first two VI components of each session were determined by the duration of the last two VR components of the prior session. During the first session, the first two VI components were arbitrarily set at 1 min.

In Phase 1, completion of a component resulted in the operation of the feeder brush, illumination of the food cup, delivery of a single food pellet, and a change in the multiple-schedule component in effect. In Phase 2, the first 49 completions of VR and VI components had the same consequences except that no pellet was delivered. Following completion of the 50th component, a different feeder operated, dropping pellets at the rate of one per second into the food cup. The session ended after the delivery of 50 pellets. In Phase 3, the original Phase 1 contingencies were reinstated.

Experimental phases remained in effect until response rates were stable, based on experimenter judgment. During Phase 2, subjects were shaped over a period of weeks to respond through 50 components before primary reinforcement was available. The shaping regime involved two manipulations. One manipulation combined the Phase 1 and Phase 2 procedures. For example, during the first shaping session, the Phase 2 contingencies were in effect for the first 26 components. Following completion of the 26th component, 26 pellets were delivered, with the remaining 24 components following the procedures of the first phase. Over sessions, the number of components employing Phase 2 contingencies was increased until both subjects responded through a 50-component delay to 50 pellets of food. The size of the VR was also manipulated during the shaping regime.

RESULTS

Figure 1 presents the VR (open symbols) and VI (filled symbols) response rates of each monkey during each session of the three phases

Fig. 1. Responses per minute during the VR (open symbols) and VI (filled symbols) components as a function of sessions during each of the three phases of the experiment.

of the experiment. Performances during Phases 1 and 3 are denoted by circles and inverted triangles, respectively, whereas response rates during and after shaping in Phase 2 are defined by triangles and squares, respectively. These data show considerable across-session variability, especially during shaping. Nevertheless, in comparing terminal performances in each condition, both monkeys tended to have larger VR-VI rate differences when food came at the end of the session than when it was delivered throughout the session.

To make this claim explicit, Figure 2 presents for each subject the VR-VI response-rate difference, averaged across the last five sessions of a phase, for each of the three phases of the experiment. These data show that the VR-VI response-rate difference increased substantially when food was moved to session's end, and that the Phase 1 baseline rate difference was re-established in Phase 3 when the Phase 1 conditions were reintroduced even though the absolute response rates in Phase 3 were higher than in Phase 1 for both animals (see Figure 1).

Figure 3 presents VR and VI response rates for each subject during each of the 25 VR and 25 VI components defining a session during Phases 2 and 3. Each data point is an average of the last 10 sessions of each condition. This graph shows that (a) Phase 2 response rates were lower during the first few components than they were throughout the rest of the session, (b) this effect was greater for the VR schedule than for the VI, (c) initial-component VR and VI rates were not lower than rates in subsequent components in Phase 3, and (d) the VR-VI rate difference reported in Figure 2 is also evident when VR and VI rates are compared on a component-by-component basis.

DISCUSSION

Marginal increases in response rate on VI schedules produce diminishing returns in increased reinforcement rates. For this reason, Prelec (1982) argued that reinforcement-optimization accounts of VI responding should predict low rates, a prediction inconsistent with the moderately higher response rates actually found on these schedules.

In this report, we modify a model by Peele et al. (1984) so that it accommodates Prelec's

Fig. 2. VR-VI rate difference in responses per minute for each subject during each phase of the experiment. Vertical lines signify the standard error of the mean.

(1982) criticism. In this modified account, VI responding is viewed as a blend of opposed response tendencies. One tendency, based on the IRT feedback function, is to select long IRTs because they have the highest reinforcement probability. The second tendency is to emit short IRTs because of their greater immediacy of reinforcement. To the extent that this time preference is present, many short IRTs will occur, raising VI response rates to levels at which marginal changes in response rate have little influence on overall reinforcement levels.

Because Peele et al. (1984) demonstrated the role of the IRT feedback function in controlling VI responding, the present experiment tested only for time preferences in IRT emission. Toward this end, an experiment was designed that varied the strength of IRT time preferences by giving food reinforcement either throughout a session or only at session's end. In the former case, time preferences should be robust because different IRTs differ in their delay to food reinforcement; however, in the

Fig. 3. Response rate in responses per minute during each of the 25 VR and VI components of the last 10 sessions of Phase 2 (left side) and Phase 3 (right side). Solid and dashed lines, respectively, denote the VI and VR schedules.

latter case, time preferences should be weak because IRT size is unrelated to delay to food reinforcement. Because this condition removes the rationale for impulsivity in IRT emission, we predicted that responding would be more strongly controlled by the rate-diminishing IRT feedback function.

The results of the experiment were consistent with our expectations: Moving food to the end of the session reduced VI response rates. Further, it did so without reducing response rates on the VR comparison schedule. The VR result is important because it argues against explaining the VI rate reduction as a strength- or value-reducing consequence of moving food to the end of the session. Were this so, why should VR rates not decrease as well?

One answer might be that VR rates did not diminish because moving food to session's end has only a small impact on a high-strength schedule such as VR. We offer two criticisms of this explanation. First, Figure 1 presents no evidence of any decrease in VR responding produced by moving food to session's end. To the degree that moving food to session's end weakens schedule value, we would expect a rate decrease in both schedules, with the VI decrease greater than the VR decrease. Second, in Figure 3, where a rate decrease can be discerned in both schedules during the first few components of Phase 2 (the "food at session's end" condition), the results are inconsistent with this strength explanation because the VI rate decrease is less than the VR rate decrease. Interpreting this result in terms of a strength hypothesis argues that VI schedules have greater reinforcing value than VR schedules. Such an interpretation obviously fails to support a strength interpretation of the results of the present study.

Although we interpret our results in terms of a maximizing model, it also is possible to explain our findings based on the idea that IRT selection is jointly governed by the immediacy of reinforcement provided by each IRT and the tendency of animals to reproduce their historic distribution of reinforced IRTs (e.g., Anger, 1956). Because the VI schedule's IRT feedback function tends to reinforce longer IRTs with higher likelihoods than shorter IRTs, the average reinforced IRT is of longer duration than the average emitted IRT. Hence, the idea that IRT emission is jointly determined by a tendency of animals to reproduce

their reinforced IRT distribution and to prefer short- over long-IRT reinforcement is no less adequate than our maximizing account in accommodating the results of this study.

The two-element molecular-maximizing model offered in the present report is compatible with the economic notion that in intertemporal choice animals maximize present value. To maximize present value, an animal first discounts the value of each good as a function of the time to its availability. With each good's value depreciated by this delay, all goods are compared as if they were concurrently available, and the good with the highest "present value" is selected.

An advantage of the "present-value" approach is that it subsumes molar- and molecular-maximizing accounts, relegating the labels "molar" and "molecular" to end ranges on a discount-rate continuum. For example, if the discount rate is low or *de minimis*, the animal is maximizing at a molar level—"molar" in that performance is based on integrating reinforcers over a long period of time without depreciating, to any substantial degree, temporally distant reinforcers relative to temporally proximal reinforcers. Such behavioral output might be properly characterized as governed by the molar feedback function between response rate and reinforcement rate. On the other hand, a high discount rate implies an animal's sensitivity to the behavioral and reinforcement stream is limited to the next IRT and the reinforcer it produces. Such reinforcement-optimization calculations might be well described by a molecular-maximizing rule.

One advantage of the present-value maximizing model is that it makes the complementarity between molar and molecular maximizing explicit. A second advantage is that it accommodates the large phyletic differences that are apparent in discount rates. At one extreme are humans, who discount money amounts at roughly 3% per year (see also Logue, Peña-Correal, Rodriguez, & Kabela, 1986). At the other extreme are pigeons, whose discount rates are so high that there is little evidence that they are controlled by reinforcers beyond the next one in the reinforcement stream (Mazur, Snyderman, & Coe, 1985).

In many experimental settings, a pigeon's impulsivity ensures its failure to maximize the rate or amount of reinforcement. Nevertheless, its behavior may remain compatible with the

idea that it is maximizing the present value of its reinforcers. We illustrate this claim in terms of a study by Vaughan and Miller (1984). In their first experiment, pigeons' key pecks produced reinforcement according to a pair of interacting schedules. One schedule was a VI that continued to operate and cumulate reinforcer assignments even if one had already been assigned for a response. The second schedule was an FR. Every time its ratio requirement was met, a single reinforcer was subtracted from the store of cumulated VI assignments. The interesting property of this schedule is that the reinforcement-maximizing strategy was to respond slowly. In fact, because VI assignments cumulated and fulfilling the ratio requirement debited the reinforcement store, near-zero response rates were optimal. Instead, pigeons responded at rates so high that they often depleted all stored VI reinforcers and received near-zero reinforcement rates.

Although these data are obviously incompatible with a simple reinforcement-maximization strategy, they can nevertheless be explained in terms of maximizing present value. To see this, notice that Vaughan and Miller's (1984) procedure imposes a tradeoff between reinforcer likelihood and immediacy. In their study, if animals opt to maximize the probability of reinforcement by emitting long IRTs, they must accept long delays between responses and reinforcement. Conversely, if they maximize reinforcer immediacy by emitting short IRTs, they reduce the probability of reinforcement. If one assumes that Vaughan and Miller's procedure involves intertemporal choice among IRTs, and that pigeons discount delayed reinforcers at very high rates, then their results can be explained in terms of a present-value account. From this perspective, Vaughan and Miller's findings are a powerful demonstration that for pigeons, maximizing reinforcer immediacy can be more important than maximizing the rate of reinforcement. But for a present-value model, it matters little whether immediacy of reinforcement is more important than rate or whether rate is more important than immediacy. The primacy of either variable can be readily accommodated.

The maximizing-present-value model and the results of the present report invite evaluation of all schedules for the role played by time preferences among reinforced IRTs in

producing a particular schedule effect. For example, why don't pigeons wait until a fixed-interval schedule has timed out or nearly timed out before they begin responding? And why don't pigeons on differential-reinforcement-of-low-rate schedules wait until the IRT-reinforcement criterion is well past before responding? The prematurity of responding on these schedules may well reflect the same trade-off noted in explaining VI performances. Responding may occur not because reinforcement is perceived as probable, but because being reinforced *now* is so much preferred to being reinforced *later*.

REFERENCES

- Anger, D. (1956). The dependence of interresponse times upon the relative reinforcement of different interresponse times. *Journal of Experimental Psychology*, *52*, 145-161.
- Logue, A. W., Peña-Correal, T. E., Rodriguez, M. L., & Kabela, E. (1986). Self-control in adult humans: Variation in positive reinforcer amount and delay. *Journal of the Experimental Analysis of Behavior*, *46*, 159-173.
- Mazur, J. E., Snyderman, M., & Coe, D. (1985). Influences of delay and rate of reinforcement on discrete-trial choice. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 565-575.
- Peele, D. B., Casey, J., & Silberberg, A. (1984). Primacy of interresponse-time reinforcement in accounting for rate differences under variable-ratio and variable-interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 149-167.
- Prelec, D. (1982). Matching, maximizing, and the hyperbolic reinforcement feedback function. *Psychological Review*, *89*, 189-230.
- 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.
- Shimp, C. P. (1969). The concurrent reinforcement of two interresponse times: The relative frequency of an interresponse time equals its relative harmonic length. *Journal of the Experimental Analysis of Behavior*, *12*, 403-411.
- Staddon, J. E. R. (1979). Operant behavior as adaptation to constraint. *Journal of Experimental Psychology: General*, *108*, 48-67.
- Vaughan, W., Jr., & Miller, H. L., Jr. (1984). Optimization versus response-strength accounts of behavior. *Journal of the Experimental Analysis of Behavior*, *42*, 337-348.

Received July 23, 1987

Final acceptance January 11, 1988