GRADED DIFFERENTIAL REINFORCEMENT: RESPONSE-DEPENDENT REINFORCER AMOUNT

G. DAVID GENTRY AND RHEA T. ESKEW, JR.

UNIVERSITY OF ROCHESTER SCHOOL OF MEDICINE AND DENTISTRY
AND
GEORGIA INSTITUTE OF TECHNOLOGY

After key pecking had been autoshaped, six pigeons were exposed to a condition in which the duration of grain availability at the end of an 8-second trial depended on the number of responses emitted during the trial (0.25-second access to grain per response). This procedure, called correlated reinforcement, alternated across conditions with the automaintenance baseline in which the 8-second trial terminated with a constant 2.5-second access to grain. Two control procedures were run; in both, the reinforcer durations were yoked to those obtained in the last correlated session. In the yoked control no responses were required, but in the single-response yoked control at least one response was required to receive the yoked duration. The correlated condition maintained response rates above those produced by the two control conditions. These results may be accounted for by differential reinforcement.

Key words: automaintenance, correlated reinforcement, terms function, reinforcer amount, differential reinforcement, yoked procedures, maximization, key pecking, pigeons

An important topic in the analysis of behavior is the study of the effects of varying the parameters of the reinforcer. The amount of reinforcer made contingent upon emitting some behavior might intuitively seem to be quite an influential parameter in controlling behavior; yet, with some important exceptions, the literature suggests that variation in the amount of the reinforcer is not a powerful technique for altering behavior (see Morse, 1966).

In single-key, free-operant procedures, providing various amounts of the reinforcer seems to have little or transient effect on behavior (Catania, 1963; Keesey & Kling, 1961), although chained schedules may enhance the ef-

fect (Lendenmann, Myers, & Fantino, 1982). Varying the amount of the reinforcer within a session may have the effect of decreasing rates immediately following delivery of increased amount (Staddon, 1970). Variations in the amount of the reinforcer produce no effect on automaintained responding, but with omission procedures and multiple automaintenance arrangements, increased amounts lead to increased responding (Balsam, Brownstein, & Shull, 1978). When an animal is permitted to choose between two responses which are reinforced with different amounts, a strong effect of amount is seen (e.g., Catania, 1963; Neuringer, 1967; Picker & Polling, 1982). Those procedures which show strong effects of variations in reinforcer amount all seem to involve either changes in stimulus conditions or some form of differential reinforcement (such as "choice" procedures).

In the experiment described here a different sort of differential reinforcement using reinforcer amount was employed. The amount of reinforcer delivered to the pigeons at the end of an 8-s trial was determined by the number of responses emitted during the trial; Logan (1960) called this sort of arrangement "correlated reinforcement." Although such correlated arrangements have been explored using runways (e.g., Brown & Horsfall, 1965; Logan, 1960), these procedures have not been studied

This work was supported in part by Grant ES-05106 to G. David Gentry from the NIEHS and in part by Contract No. DE-AC02-76EV03490 with the U.S. Department of Energy at the University of Rochester Department of Radiation Biology and Biophysics and has been assigned Report No. UR-3490-2207. The authors would like to thank Chris Newland, Vic Laties, and Jack Marr for their comments on the paper. Reprints may be obtained from G. David Gentry, now at Psychology Department, Mount Union College, Alliance, Ohio 44601. The present address for Rhea T. Eskew, Jr., is Department of Psychology, C-009, University of California at San Diego, La Jolla, California 92093. Portions of these data were presented at meetings of the Southeastern Psychological Association, Atlanta, Georgia, 1981, and the American Psychological Association, Washington, DC, 1982.

using simpler responses in a more controlled environment (however, see Baum, 1973, for a related approach).

METHOD

Subjects

The subjects were six male White Carneaux pigeons maintained at approximately 80% of their free-feeding weights throughout the experiment. All subjects were experimentally naive.

Apparatus

The Coulbourn test chamber (Model E10-10) had two keys, mounted 16 cm apart, 23 cm above a grid floor. A white houselight (G.E. 1829), mounted 5 cm above and midway between the keys, was illuminated throughout the session. A minimum force of .2 N was required to activate the key. The test chamber was enclosed in an acoustical isolation chamber (Industrial Acoustics Co.) in which there was continuous masking noise. The sequence of events was controlled by a PDP 8/e mini-computer using the SUPERSKED operating system (Snapper & Inglis, 1979). All interevent times were recorded with 10-ms resolution for later analysis.

Procedure

Throughout the study, each session consisted of 75 trials. The bird's key pecking was first autoshaped using 8-s trials separated by an average intertrial interval of 20 s, with a 10 to 30 s range. At the beginning of each trial the left key was transilluminated by a red light. After 8 s, the keylight was turned off, the feeder magazine was lit, and grain was presented for 2.5 s. Under this procedure, key pecking had no scheduled consequence. This autoshaping procedure was used throughout the study as a baseline and will henceforth be referred to as the automaintenance procedure.

In the correlated procedure, a direct relation was imposed between the number of responses emitted during the 8-s trial and the duration of operation of the feeder at the end of the trial. If no responses were emitted, there was a 0.25-s operation of the feeder light and hopper. For each response made during the trial, an additional 0.25 s was added to the light and feeder operation (e.g., at the end of a trial containing 10 responses the feeder was

activated for 2.75 s). Two control procedures were examined: Under the yoked procedure, the sequence of feeder durations presented to each bird was the same as the sequence that bird had obtained in the last session of the preceding correlated condition; the only difference was that no response was required for reinforcer delivery. The second control procedure, the single-response yoke, was identical except that at least one response was required to produce the yoked duration of feeder operation. If no responses were emitted in this condition, there was a 0.25-s feeder flash. In addition, two birds (P10 and P25) were exposed to a condition in which a single response was required to obtain constant 4.5-s access to grain (the single-response procedure); otherwise there was a 0.25-s feeder flash. The sequence of procedures and the minimum numbers of sessions under each are shown in Table 1.

RESULTS

The results are based on means and ranges of the session averages over the last five sessions under each condition. The last eight conditions, run for all birds, will be described first.

Figure 1 shows the mean number of responses per trial and the range over the replications of procedures. Here the data have been averaged over the three replications of the automaintenance and correlated procedures; the data from the individual conditions show the same results except as discussed below.

There were no consistent differences across subjects between the automaintenance and correlated procedures. There were also no con-

Table 1
Order of conditions, minimum number of sessions, and birds run under each.

Pigeons	Condition	Session
P10 and P25	I Automaintenance	40
	II Correlated Reinforcement	7
	III Automaintenance	9
	IV Yoked Reinforcement	9
	V Automaintenance	4
	VI Single Response	8
All Birds	I Automaintenance (A ₁)	60
	II Correlated Reinforcement (C ₁)	21
	III Automaintenance (A ₂)	22
	IV Correlated Reinforcement (C ₂)	12
	V Automaintenance (A ₃)	16
	VI Yoked Reinforcement (Y)	10
	VII Single-Response Yoked (R)	16
	VIII Correlated Reinforcement (C ₃)	24

Fig. 1. Each panel shows the mean responses per trial for each experimental procedure averaged over the last five sessions per condition. The error bars for the auto-maintenance (A) and correlated (C) conditions are the ranges of condition means for the three replications. Y and R are the yoked and single-response yoked procedures, respectively.

sistent differences across subjects between the two yoked procedures. However, the two yoked conditions produced response rates lower than those of the correlated condition for all birds except P203, for which the yoked rate was low but the single-response yoke rate was as high as the baseline and correlated rates.

With the exceptions of P10 and P28, the birds responded at least once on almost every trial in all conditions of the study, and therefore an almost identical pattern of results is seen when response-containing trials were examined alone. The same general pattern—that is, the automaintenance and correlated procedures about equal and both greater than the two yoked procedures—was also found for the running rates.

Figure 2 shows the mean number of responses emitted in the last three conditions

Fig. 2. Each panel shows the responses per trial as a proportion of the second correlated condition (C₂). The two hatched bars are the yoked procedures (Y—yoked, and R—single-response yoked) for which C₂ was the lead. The third correlated procedure (C₃) is shown to illustrate recovery. The error bars on C₂ show the range of session means expressed as proportion of condition mean.

expressed as a proportion of the number emitted in the second correlated condition, from which the yoked reinforcer amounts were obtained. As the figure shows, the yoked conditions maintained lower response rates than the correlated conditions. Recovery from the two yoked conditions was not complete for P16 and P28 but was complete for the other birds.

Figure 3 shows the transition from the single-response yoke to the last correlated condition. As shown in Figures 1 and 2, P25 and P203 did not show elevated rates. The other four subjects showed a gradual increase in rate over a number of sessions. Notice that P28 went seven sessions without a change in responding. It would appear that P28 was not responding enough to make contact with the dependency.

The first six conditions, run only for P10 and P25, produced results which were almost identical to those of the last set of conditions. The single-response condition, which required at least one response for a constant duration of

Fig. 3. Transition over sessions in responding (responses per trial) from single-response yoked (R—squares) to the correlated reinforcement (C₃—triangles). The last five session for the R condition, numbers 12 to 16, are shown in the left of each panel. All 24 sessions of C₃ condition are in the right portion of each panel.

access to grain, did not elevate rates to correlated condition rates.

Delay is defined as the time elapsing after the last response until reinforcer delivery. The

individual condition data indicate that only P10 and P28 showed regular variation in this delay with condition over the course of the experiment and that all birds tended to have

slightly less variable delays under the correlated procedure than under the automaintenance procedure; however, the mean delay under the correlated conditions was shorter than that under the automaintenance conditions for five birds.

It is interesting to note that the difference across birds seems replicable. P10 and P25 responded in much the same way under the two sets of experimental conditions to which they were exposed. P10 had relatively few response-containing trials, fairly short delays, and long pauses before beginning to respond (under all except the correlated procedure), throughout the experiment. P25 had fairly long delays and short pauses that did not vary consistently with procedure, and these were also seen throughout the experiment.

DISCUSSION

The present correlated-reinforcer arrangement produced higher response numbers, longer times spent responding, and shorter delays (for five birds) than did the two control procedures. The yoked procedure was identical to the correlated procedure except that there was no response requirement; the single-response yoke was identical except for a minimal response requirement. The automaintenance procedure, in which there was a constant amount of reinforcer, also elevated these same measures of responding above the yoked-control procedures; however, the automaintenance procedure differed from the correlated one in several ways, particularly in that a constant smaller amount of reinforcer was provided on each trial. Providing a constant, but larger, amount of food for at least one response, as was done for P10 and P25 in the single-response condition, did not increase responding to the correlated level.

Essentially there are two findings requiring explanation. First, there was the failure to find a significant difference in the results of the correlated and automaintenance procedures. The numerous differences between the two procedures do not allow for a clear isolation of the important factor(s). Dodd (1980) attempted to increase pigeons' autoshaped response rates using a DRH procedure. However, it was difficult because the birds were already responding at a high rate. A similar sort of "ceiling effect" could be operating here

to prevent correlated rates from increasing above automaintained rates. However, this explanation seems untenable for P16 and P25.

The second, and major, finding is that the correlated procedure, in which the amount of reinforcer was proportional to the number of responses emitted during the trial, produced greater response numbers than when the same amounts of reinforcer were given independently of responding. That differential reinforcement works to change responding is not surprising; but that differential reinforcement can operate through variations in reinforcer amount when there is no choice response available, and without any requirement of contiguity between response and reinforcer delivery, is perhaps surprising. Note that none of these birds had ever experienced an experimental procedure in which the reinforcer was presented as an immediate scheduled consequence of a response.

Logan (1960) described the sort of arrangement used in this study as "correlated reinforcement." It should be emphasized that reinforcer magnitude is not just varied, but rather it is systematically varied according to the performance of the subject. Logan (1960) used what he called "terms functions" to describe the relationships between reinforcers and performances.

The advantage of using terms functions to describe experimental procedures is that a large number of schedules can be placed in a single framework of differential reinforcement. These terms functions can be illustrated graphically as shown in Figure 4. In this figure some property of the reinforcer is plotted as a function of some property of responding. The lines on the graph depict the terms functions. The standard all-or-none differential reinforcement is shown as a step function (solid line). The straight line (dashed) from lower left to upper right indicates graded differential reinforcement. The straight line (dotted) from upper left to lower right represents graded differential reinforcement of other behavior. A terms function can be represented by any set of points in the space defined by the response and reinforcer parameters; these functions can include smooth functions, discontinuities, and even probability density functions.

If variables other than magnitude are included in the terms function, then ratio sched-

TERMS FUNCTIONS

Fig. 4. Potential terms functions that can be used to describe schedules of reinforcement. The y axis can be any parameter of the reinforcer and the x axis any measure of performance. The step function (solid line) illustrates the standard differential-reinforcement procedure. The positive-slope line (dashed line) shows graded differential reinforcement. The negative-slope line (dotted line) shows graded differential reinforcement of other behavior.

ules also fit under the heading of correlated reinforcement. On a ratio schedule, the faster the responses are completed the sooner the reinforcer is delivered. Thus, there is a correlation between rate of responding and immediacy of reinforcement. Ratio schedules have a built-in positive feedback loop: High-rate responding produces high reinforcement rate, high reinforcement rate produces high-rate responding. However, as Zeiler (1979) pointed out, this positive feedback loop cannot alone explain performance in ratio schedules because it is just as likely that a lower response rate would occur and produce a low reinforcement rate. Zeiler (1979) suggested that animals "learn the *relation* [*italics added*] between their behavior and the demands of the environment" (p. 97). In the case of ratio schedules, the relation is between response rate and reinforcement rate. This account appears to go beyond saying that the animal comes under control of the contingency, an arrangement which could produce either high or low rates. It would appear that there is some sort of maxi-

mizing occurring (see Rachlin, Battalio, Kagel, & Green, 1981). Logan (1960) described this as behavior occurring such that maximum reinforcement is obtained with minimum effort, a view similar to that of Rachlin et al. (1981), who consider maximization to be a motivational concept.

The differences between the fixed-ratio schedule and the present correlated schedule are interesting. With the former, a constant magnitude of reinforcer is given with a variable interreinforcer interval that is correlated with response rate. With the latter, a variable magnitude that is correlated with response rate is given with a fixed interreinforcer interval. Given this similarity, it is not too surprising that both procedures maintain high rates of responding. However, if one considers contiguity of responses and reinforcers as essential, the similarity is not so clear.

Under ratio schedules, short IRTs are differentially reinforced (Morse, 1966; Zeiler, 1977) with the more immediate presentation of the reinforcers. With the present correlated condition, contiguity between responses and reinforcer presentations was not guaranteed and in fact the delay between the last response and the delivery of the reinforcer was quite variable across conditions and subjects. Furthermore, the opportunity for adventitious contiguity was equal for the correlated and yoked conditions. Thus it would seem that any differences between number of responses under the correlated and yoked conditions must be attributed to the differences in the relation specified by the terms function.

In the case of humans, it would be tempting to attribute the power of the correlated arrangement to the use of rules (see Skinner, 1969). However, in the case of nonhuman subjects we cannot attribute the control of such a relation to the use of rules; in the absence of linguistic competence, it must be assumed that the behavior engendered under the correlated arrangement was contingency-shaped without the benefit of guaranteed contiguity. Shimp (1975) has noted that under standard schedules of reinforcement, contingency and contiguity are confounded. In Shimp's terms, if one allows for an animal to remember recent behavior, then the reinforcer can strengthen behavior over a longer time span (Hawkes & Shimp, 1975; Shimp, 1975). Hawkes and Shimp (1975) demonstrated that under a short fixed

interval (FI 5-s), a behavioral pattern of successively shorter IRTs could be shaped. Hawkes and Shimp (1975) chose to discuss their results in terms of short-term memory for recent behavior, but it is equally descriptive and perhaps more accurate to say that the reinforcer acts on more than the response that immediately preceded it (see Catania, 1971; Dews, 1969). That is, the reinforcer may act back over some period to strengthen a series of responses or a pattern of responses. A time span of action for reinforcement is also in agreement with the development of mediating behavior with IRT schedules (e.g., Laties, Weiss, Clark, & Reynolds, 1965). In the present study the trial was 8 s, a time span which is reasonably short. If reinforcement can act over such a span, then there is a mechanism for differential reinforcement or a contingency-shaped high response rate.

Differential reinforcement is a term which has referred primarily to schedules in which a reinforcer is delivered if, and only if, the required response or response pattern is emitted. This arrangement is an all-or-none terms function. However, as Logan (1960) implied, other forms of differential reinforcement can be arranged based on a graded terms function. Rather than all-or-none reinforcement, the reinforcer can be determined in some manner proportional to behavior. Ratio schedules and the present correlated arrangement are examples of graded differential reinforcement. With ratio schedules, the rate of reinforcement is graded according to the rate of responding. With the present correlated schedule, the magnitude of reinforcement was graded with rate of responding. Although proper experimental data are still lacking, it is possible that the limitation on differential reinforcement is the time span over which a reinforcer might act. The time-span aspect can of course be overcome by use of chaining of behavior or exteroceptive stimuli.

There is formal similarity between Figure 4 and the temporal-constraint functions of Rachlin et al. (1981). The terms functions drawn in Figure 4 are the relationships between some property of behavior (rate, topography, force, pattern, etc.) and some property of the reinforcer (rate, magnitude, quality, etc.); these relationships do not necessarily correspond to the temporal constraints involved in making the response or the time taken to

consume the reinforcer. The terms functions do limit the obtained outcome to a point on the line. The maximization theory of Rachlin et al. (1981) represents an attempt to predict where that point will fall on such a line. The present data are not sufficient to favor a molar-maximization or a molecular-differential-reinforcement theory; however, prior to invoking a second principle, maximization, perhaps differential reinforcement should be more fully explored. Investigating the range of potential terms functions (such as shown in Figure 4) for a variety of response and reinforcer parameters could potentially delimit the usefulness of the two approaches.

REFERENCES

- Balsam, P. D., Brownstein, A. J., & Shull, R. L. (1978). Effects of varying the duration of grain presentation on automaintenance. *Journal of the Experimental Analysis of Behavior*, 29, 27-36.
- Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, 20, 137-153.
- Brown, J. S., & Horsfall, R. B. (1965). Speed-contingent reinforcement. *Psychonomic Science*, 3, 305-306.
- 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. (1971). Reinforcement schedules: The role of responses preceding the one that produces the reinforcer. *Journal of the Experimental Analysis of Behavior*, 15, 271-287.
- Dews, P. B. (1969). Studies on responding under fixed-interval schedules of reinforcement: The effects on the pattern of responding of changes in requirements at reinforcement. *Journal of the Experimental Analysis of Behavior*, 12, 191-199.
- Dodd, P. W. D. (1980). Discrimination and differentiation of response number in stimulus directed pecking of pigeons. *Journal of the Experimental Analysis of Behavior*, 33, 253-264.
- Hawkes, L., & Shimp, C. P. (1975). Reinforcement of behavior patterns: Shaping a scallop. *Journal of the Experimental Analysis of Behavior*, 23, 3-16.
- Keesey, R. E., & Kling, J. W. (1961). Amount of reinforcement and free-operant responding. *Journal of the Experimental Analysis of Behavior*, 4, 125-132.
- Laties, V. G., Weiss, B., Clark, R. L., & Reynolds, M. D. (1965). Overt "mediating" behavior during temporally spaced responding. *Journal of the Experimental Analysis of Behavior*, 8, 107-116.
- Lendenmann, K. W., Myers, D. L., & Fantino, E. (1982). Effects of reinforcer duration on responding in two-link chained interval schedules. *Journal of the Experimental Analysis of Behavior*, 37, 217-222.
- Logan, F. A. (1960). *Incentive: How the conditions of reinforcement affect the performance of rats*. New Haven: Yale University Press.

- 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.
- Neuringer, A. J. (1967). Effects of reinforcement magnitude on choice and rate of responding. *Journal of the Experimental Analysis of Behavior*, 10, 417-424.
- Picker, M., & Poling, A. (1982). Choice as a dependent measure in autoshaping: Sensitivity to frequency and duration of food presentation. *Journal of the Experimental Analysis of Behavior*, 37, 393-406.
- Rachlin, H., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences*, 4, 371-388.
- Shimp, C. P. (1975). Perspectives on the behavioral unit: Choice behavior in animals. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 2. Conditioning and behavior theory* (pp. 225-268). Hillsdale, NJ: Erlbaum.
- Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. New York: Appleton-Century-Crofts.
- Snapper, A. G., & Inglis, G. B. (1979). *The SKED software system: Time shared SUPERSKED, Manual 3*, Rev. E. Kalamazoo, MI: State Systems.
- Staddon, J. E. R. (1970). Effect of reinforcement duration on fixed-interval responding. *Journal of the Experimental Analysis of Behavior*, 13, 9-11.
- Zeiler, M. D. (1977). Schedules of reinforcement: The controlling variables. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 201-232). Englewood Cliffs, NJ: Prentice-Hall.
- Zeiler, M. D. (1979). Output dynamics. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behaviour: Vol. 1. Reinforcement and the organization of behaviour* (pp. 79-115). New York: Wiley.

Received August 27, 1982

Final acceptance August 16, 1983