JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1989, 52, 27-40 NUMBER 1 (JULY)

ALGORITHMIC SHAPING AND MISBEHAVIOR IN THE ACQUISITION OF TOKEN DEPOSIT BY RATS

MARIE MIDGLEY, STEPHEN E. G. LEA, AND RACHEL M. KIRBY

UNIVERSITY OF EXETER

In two experiments, rats were trained to deposit ball bearings down a hole in the floor, using an algorithmic version of shaping. The experimenter coded responses expected to be precursors of the target response, ball bearing deposit; a computer program reinforced these responses, or not, according to an algorithm that mimicked the processes thought to occur in conventional shaping. In the first experiment, 8 of 10 rats were successfully shaped; in the second, 5 of 5 were successfully shaped, and the median number of sessions required was the same as for a control group trained using conventional shaping. In both experiments, "misbehavior," that is, excessive handling and chewing of the ball bearings, was observed, and when the algorithmic shaping procedure was used, misbehavior could be shown to occur in spite of reduced reinforcement for the responses involved.

Key words: shaping, misbehavior, token reinforcement, fixed ratio, operant-respondent interactions, ball bearing deposit, rat

Shaping, or the systematic elaboration of complex response topographies by the reinforcement of successive approximations, is one of the best known and most influential discoveries of operant psychology. Skinner discussed it in some of the earliest systematic accounts of operant conditioning (e.g., Skinner, 1953, chap. 6; 1954). It has been used to establish many complex and even bizarre responses (e.g., the playing of a form of table tennis by pigeons, Skinner, 1962; or the use of cocktail sticks as tools by crows, Powell & Kelly, 1975). It was quickly applied in the clinical and educational fields (e.g., Hingtgen & Trost, 1966; Skinner, 1954). In its original form, the shaping of lever press responding in rats, it is probably used with thousands of animals every year as a preliminary to research and as a demonstration in teaching.

Scientifically speaking, however, shaping has scarcely been investigated, and its practice could fairly be said to be more an art than a science. Platt (1973) has pointed out that conventional procedures for shaping confound two variables: the *contact* between reinforcement contingencies and behavior (the proportion of emitted responses that are reinforced), and the *selectivity* of the procedure (the difference between reinforced and nonreinforced response classes). But in any case, it would be hard to point to published evidence that the selective reinforcement of successive approximations to the desired response actually occurs in practical shaping, let alone that it is responsible for the changes in response frequencies that are observed. And there are obvious competing accounts, for example the kinds of Pavlovian processes that are known to be capable of producing key pecking in pigeons (Brown & Jenkins, 1968) and lever pressing in rats (Myer & Hull, 1974).

Furthermore, there are serious doubts about the scope and power of shaping. These doubts have existed at least since Breland and Breland (1961) reported that shaped responses could be displaced by "instinctual drift" towards more "natural" responses of the animal concerned, and more systematic investigations (e.g., Shettleworth, 1975) showed that not all responses are affected by all reinforcers. Even Ferster and Skinner (1957, p. 32) questioned the efficiency of shaping as a means of getting pigeons to peck keys, suggesting that it is likely

Thanks are due to the UK Science and Engineering Research Council for a research studentship to Marie Midgley; to the Alexander Von Humboldt Foundation for support to Stephen Lea during the preparation of the manuscript; to Reg Dimon for apparatus design and construction; and to John Staddon and Clive Wynne for their comments on the manuscript. Marie Midgley is now at the Psychology Department, St. Andrew's Hospital, Billing Road, Northampton NN1 5DG, England. Correspondence and reprint requests may be sent to her or to Stephen Lea at the Department of Psychology, University of Exeter, Washington Singer Laboratories, Exeter EX4 4QG, England (electronic mail: LEA@UK.AC.EXETER).

to lead to superstitious persistence of reinforced precursor responses.

What is needed is to specify more clearly the set of rules that are being used to decide which responses to reinforce during shaping; we can then investigate experimentally how application of these rules affects behavior, and the circumstances in which they may fail. In this paper, any such set of rules for shaping is called a "shaping algorithm." The clearest theoretical treatment of such algorithms is that of Platt (1973). He points out that shaping has no meaning unless the different responses that an organism emits can be placed on at least an ordinal dimension. The object of shaping is then to alter the distribution of the subject's response on this dimension, usually in a particular direction or towards a particular target. Platt proposes a "percentile reinforcement" schedule, which reinforces responses lying within a particular distribution percentile of the target, based on a running sample. This is, in effect, a one-rule shaping algorithm: "Always reinforce the best 20% (or 50%, or whatever) of ongoing responding." Such a schedule controls the contact the reinforcement contingencies will make with behavior, but leaves their selectivity unspecified. The few experimental studies of shaping that have been reported, however, have tended to use algorithms that control selectivity rather than contact.

One such study is that of Eckerman, Hienz, Stern, and Kowlowitz (1980). They used a 10-in. (25 cm) wide response area, at which pigeons were pretrained to peck, and varied the part of the area where pecks would be reinforced. By varying the speed and size of changes in the criterial area, they sought to find the optimal parameters for producing rapid changes in the location of the pigeons' pecks. In this situation, the most rapid shaping occurred with large, fast changes in criterion (high selectivity, and, as a consequence, high contact, in the terms of Platt, 1973).

Eckerman et al.'s (1980) result is a very useful one if it can be generalized, but their procedure has three obvious disadvantages as a model of shaping in general. First, it studied only the variation of location of a single response, not changes in response topography (and furthermore, all response locations had some previous history of reinforcement). Second, responses at a given location were effectively subjected to only two schedules of reinforcement—either they were being reinforced or they were not. In normal shaping, responses that have recently emerged are usually given continuous reinforcement, which becomes more intermittent once they are well established. Finally, when a totally new response is being shaped, it is usually the case that earlier approximations to the desired response are in some sense logically necessary to later ones, or even part of them; for example, early in shaping rats to press a lever, touching the lever is commonly reinforced as an approximation; when the rat progresses beyond this stage, it must nonetheless continue to touch the lever—it is impossible to press it without doing so. In the Eckerman et al. study, however, pecking one part of the area was not a logical precursor to pecking another part.

The second of these objections is elegantly overcome by Platt's (1973) percentile reinforcement schedule. In principle, too, this schedule could also be used with the kind of heterogeneous sequence of responses used in normal shaping, because the required ordinal response dimension clearly exists in the form of nearness to the final target response. Unfortunately, however, although a number of experiments have been published using Platt's schedule, they have all involved quantitative variations in responses whose basic topography had already been established. For example, Alleman and Platt (1973) used percentile reinforcement to modify the interresponse times of pigeons' key pecking, and Davis and Platt (1983) used it to modify the direction in which rats pushed a suspended joystick. These experiments have established the effectiveness of the percentile reinforcement schedule, but throw light on the processes of normal shaping only by analogy.

A recent paper by Pear and Legris (1987) goes some way to overcoming this last limitation. Pigeons were observed using two videocameras connected to a microcomputer, which was programmed to detect the position of the bird's head within the test chamber (Pear & Eldridge, 1984). The reinforced response was moving the head into a "virtual sphere," that is, a spherical space that was defined only within the computer. Gradual reductions of this sphere were used to shape the pigeons to put their heads into a small (3-cm diameter) target area in one back corner of the chamber;

the operant level of this response was nil. The shaping algorithm was simple: After every reinforcement, the sphere of reinforceable locations was reduced in radius, whereas after every 10-s period without reinforcement, it was increased in radius. The procedure was effective in producing contact with the target within four sessions for each of the 3 pigeons tested. Two qualitatively distinct precursor responses (walking towards the target, and lowering the head to enter it) were observed during the course of shaping, but were not specifically reinforced. These results go some way towards demonstrating the reality of the processes usually supposed to occur during shaping, namely changes in the reward frequency for responses that are already in the animal's repertoire, followed by increases in the frequency of those responses. In other words, they demonstrate that shaping really does involve the selective reinforcement of successive approximations to the desired response.

The procedure of Pear and Legris (1987) points to a weakness of percentile reinforcement as a way of establishing truly novel responses. In order to maintain the schedule's contact with responding, it was necessary to relax its selectivity if too much time went by without a criterion response occurring. Contact, in other words, needs to be measured, and if possible controlled, in terms of reinforcements per unit time rather than responses per reinforcement. Platt (1973) proposed an extension of his schedule (interval-percentile reinforcement) that meets this objection, but it lacks the original schedule's elegance and has not led to any helpful experimentation.

The present experiments extend the experimental study of shaping by investigating rats' acquisition of a complex novel response, the deposit of ball bearings down a hole in the floor. This is a response in which misbehavior in the sense defined by Breland and Breland (1961) has been shown to occur (Boakes, Poli, Lockwood, & Goodall, 1978), and in which appeal has been made to Pavlovian processes to account for the observed behavior (Boakes & Jeffery, 1979; Timberlake, Wahl, & King, 1982). It is also complex enough for the process of shaping to be quite protracted, so that it is possible to study intermediate stages.

The general approach taken was as follows. An experimenter observed the rat throughout the shaping process and recorded on a microcomputer keyboard each of a number of responses that were thought of as approximations to the target response. But she was relieved of the task of reinforcing responses by the computer program, which implemented a shaping algorithm. This algorithm attempted to mimic the contingencies of reinforcement conventionally held to be used during shaping. Although the resulting schedule was considerably more complex than percentile reinforcement, it followed Platt (1973) in trying to control contact with the schedule rather than selectivity.

In summary, the experiments constituted an experimental synthesis of normal shaping, and the first question being asked was whether this kind of algorithmic shaping of a complex response is in fact possible (and acceptably efficient). If it is not, it suggests that we have seriously misspecified what happens during shaping. Our second aim was to find out whether any events during shaping were correlated with the emergence of misbehavior.

Table 1

Definitions of the behavior codes used during algorithmic shaping.

Response	Description
Approach	The rat orients towards and/or moves towards the ball bearing.
Touch	The rat orients towards and touches the ball bearing with nose or front paws.
Move	The rat moves the ball bearing by contacting it with nose or front paws.
Move towards hole	The rat pushes or pulls the ball bearing towards the vicinity of the floor hole.
Reach hole	The rat moves the ball bearing with nose or front paws, towards the floor hole and within 3 in (7.5 cm) of it.
Deposit ball bearing	The ball bearing falls down the floor hole.
Mouth	The rat holds the ball bearing in its open mouth.
Chew	The rat moves its teeth against the ball bearing.

EXPERIMENT 1

The main thrust of the first experiment simply was to demonstrate that algorithmic shaping of a novel response could work. In order to do this, we first had to find a suitable set of precursor responses (i.e., reinforceable approximations to ball bearing deposit) and a suitable algorithm. The precursor responses were derived from our previous experience of conventional shaping of this response (Midgley, 1986, chap. 1) and are listed in Table 1, in the order in which they were expected to emerge (Mouth and Chew, also included in Table 1, were not intended as precursors but were also observed in some conditions). Table 1 in effect specifies an ordinal dimension of response generalization, of the sort identified by Platt (1973) as essential if shaping is to be meaningful. The aim of shaping was to move the rats' responding along this dimension. The response order can also be thought of as a hierarchy, and responses earlier and later in the order are referred to below as "lower" or "higher," respectively.

The algorithm itself was as follows:

1. At any moment, a fixed-ratio (FR) schedule of reinforcement was in effect for each response.
2. An initial FR was defined for each response.
3. After a given number of reinforcements had been given for each response, its FR was increased by a constant.
4. After a given number of reinforcements had been given for each response, reinforcement for all responses earlier in the hierarchy was stopped (i.e., their FRs were raised to infinity).
5. If a given time elapsed without reinforcement (because no response for which the current FR was finite was occurring at a sufficient rate), then when a given number of any response lower in the hierarchy had been recorded, reinforcement of that response was reinstated at its *initial* FR; and the FRs associated with all higher responses were returned to their initial values.
6. The parameters associated with Steps 2 through 5 could be varied between (but not within) sessions.

Steps 2 through 5 of this algorithm obviously contain a large number of parameters: For each response there is the initial FR, the number of reinforcements given at each FR value, the number of reinforcements before cutting out reinforcement for lower responses, the time before backtracking, and the number of responses to be made before backtracking was invoked. With six responses in the hierarchy, this adds up to 30 parameters, so there was no question of exploring the parameter space systematically. Instead, a series of parameter combinations were defined before the start of the experiment, on the basis of previous experience with hand shaping the deposit response. They were arranged in a sequence, from the least demanding (in terms of having the lowest schedule parameters for the lower level responses) to the most. For the most part, the rats progressed through this sequence as sessions of shaping elapsed.

METHOD

Subjects

Ten experimentally naive female Lister hooded rats were obtained from OLAC, Bicester, Oxfordshire. They were approximately 5 months old at the start of the experiment, and had mean free-feeding weights of 273 g (range, 235 to 321 g). The rats were deprived of food for 12 hr before the start of the first session; thereafter they were given limited food, sufficient to maintain their body weights at or above 85% of free-feeding levels (actual body weights before testing averaged 86% of free-feeding weights over the experiment).

Apparatus

The test chamber used was made in the laboratory. Its general layout is shown in Figure 1. Its internal dimensions were 31.2 by 30.4 by 21.0 cm. The front and back walls (carrying the food dispensers and lever) were made of aluminum, the side walls and ceiling of clear perspex, and the floor of 0.9-cm plywood. The liquid dispenser, and the corresponding floor hole and wall light, were not used in this experiment, nor were the retractable lever and the light above it. The floor hole in front of the food dispenser, which was used, was 3 cm in diameter, tapering to 2 cm within the thickness of the floor. Movements of the flap in front of the food tray, and deposit of ball bearings down the floor hole, could be detected by microswitches. Ball bearings that fell down the floor hole were collected in a foam-lined plastic box. The food dispenser delivered 45-mg pellets of mixed diet (Campden Instruments Ltd.). A more detailed description of the chamber is given by Midgley (1986, pp. 56 ff.).

The ball bearings used were 21/32 in. (ap-

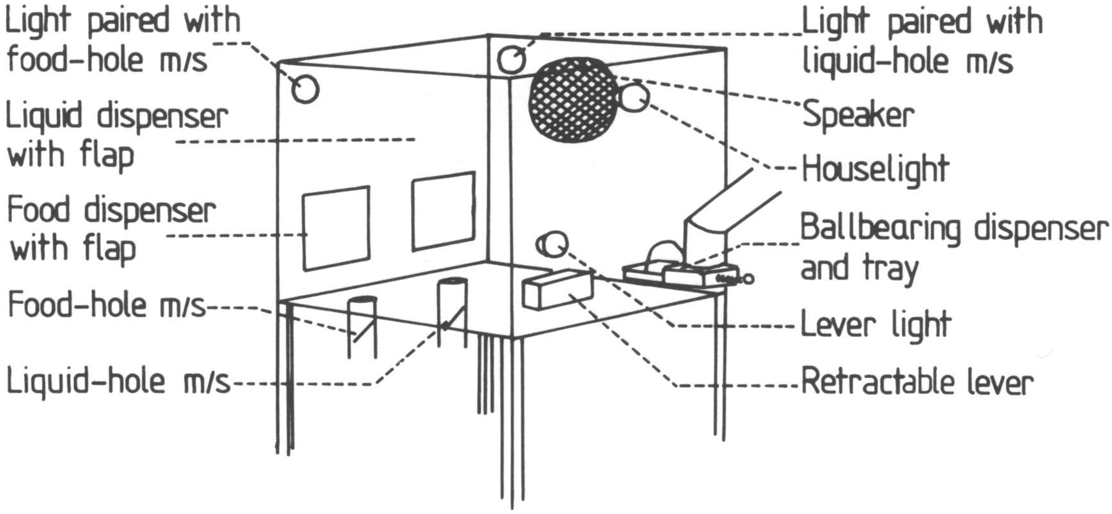

Fig. 1. Outline view of the test chamber, approximately to scale. M/S = microswitch.

proximately 17 mm) in diameter, were made of steel, and weighed 19 g. Before use, they were washed to remove industrial grease; after each experimental session they were washed in hot water and dried immediately to prevent rusting.

Throughout the experiment, low-volume white noise (68.5 dB, scale C) was played through the loudspeaker (35 ohm). The box was situated in a quiet room lit by a 40-W red light from a table lamp. Behavior was recorded, and experimental contingencies were implemented, using a microcomputer (Commodore Pet®) programmed in BASIC. The experimenter sat approximately 1 m from the test chamber, opposite the left-hand side (i.e., the side nearer to the food dispenser).

Procedure

All rats were first trained to push the food tray flap open and collect a pellet when a light that shone into the food tray was illuminated. This required four 15-min daily sessions. Subsequent sessions, which were given one per day, normally 7 days per week, and lasted for 15 min, involved the shaping algorithm outlined in the introduction to this experiment. An observer (always the first author) recorded all occurrences of the first five responses listed in Table 1; the sixth response, ball bearing deposit, was recorded automatically. This series of responses has the property that each higher level response requires the preceding occurrence of all lower level responses. However, it was possible for several levels to occur virtually simultaneously (e.g., a rat could touch a ball bearing, move it, and move it towards the hole in a single movement). In such cases, only the highest level response of the sequence was recorded. The reliability of this observer's coding of these and similar responses has been found to be satisfactory, both by repeated coding of the same session from a videotape and by comparison with the codings of an independent observer (correlations of response totals for individual sessions ranging from .92 to 1.00 for intraobserver tests, and .81 to .99 for interobserver tests: Midgley, 1986, Tables 3.2a and 3.1a).

Reinforcement was given according to the algorithm, with parameter values determined as follows. The amount by which FR schedules were increased, when sufficient reinforcers had been obtained for the corresponding response, was one for all responses throughout the experiment. For the remaining four parameters, the first nine combinations listed in Table 2 were usually used in sequence, with progression from one to the next usually every third session. However, this plan was varied for individual rats, depending on their performance. Occasionally a combination was skipped if a rat was learning quickly; in two cases it was necessary to give additional magazine training after a few sessions' exposure to Combination 1; in several cases it was necessary to return

Table 2

Parameter combinations used in algorithmic shaping. Parameters for each response are abbreviated as follows: Initial FR, the fixed ratio in force at the start of the session; Rfs/ratio, the number of reinforcements before the fixed ratio for that response was raised by 1; Rfs/ext, the number of reinforcements before the fixed ratios for all "lower" responses were raised to infinity; and Max IRfI, the maximum time (in minutes) without reinforcement before the fixed ratios for "lower" responses were restored to initial values. "MAX" stands for 32767 (i.e., effectively infinite).

		Response					
No.	Parameter	Approach	Touch	Move	Move towards hole	Reach hole	Deposit ball bearing
1	Initial FR	2	1	1	1	1	1
	Rfs/ratio	5	7	10	20	20	MAX
	Rfs/ext	MAX	10	15	30	50	50
	Max IRfI	3	3	2	2	2	2
2	Initial FR	4	3	2	1	1	1
	Rfs/ratio	3	3	7	15	20	MAX
	Rfs/ext	MAX	4	7	20	50	50
	Max IRfI	6	4	4	2	2	2
3	Initial FR	6	4	3	1	1	1
	Rfs/ratio	3	3	4	10	20	MAX
	Rfs/ext	MAX	3	5	15	30	50
	Max IRfI	8	6	5	3	2	2
4	Initial FR	10	5	4	1	1	1
	Rfs/ratio	2	3	4	7	20	MAX
	Rfs/ext	MAX	3	5	7	30	50
	Max IRfI	10	8	6	3	2	2
5	Initial FR	10	7	5	2	1	1
	Rfs/ratio	2	2	3	7	15	MAX
	Rfs/ext	MAX	3	3	6	25	50
	Max IRfI	12	9	7	4	2	2
6	Initial FR	10	10	7	3	2	1
	Rfs/ratio	2	2	2	7	15	MAX
	Rfs/ext	MAX	3	3	5	25	50
	Max IRfI	12	12	10	5	3	2
7	Initial FR	10	10	10	4	2	1
	Rfs/ratio	2	2	2	4	12	MAX
	Rfs/ext	MAX	2	2	3	10	50
	Max IRfI	12	12	12	7	3	2
8	Initial FR	10	10	10	10	4	1
	Rfs/ratio	2	2	2	2	6	MAX
	Rfs/ext	MAX	2	2	2	4	5
	Max IRfI	12	12	12	12	6	2
9	Initial FR	10	10	10	10	10	1
	Rfs/ratio	1	1	1	1	1	MAX
	Rfs/ext	MAX	1	1	1	1	1
	Max IRfI	12	12	12	12	12	2
10	Initial FR	10	10	10	2	1	1
	Rfs/ratio	2	2	2	20	20	MAX
	Rfs/ext	MAX	2	2	4	30	50
	Max IRfI	12	12	8	2	2	2
11	Initial FR	10	10	10	2	1	1
	Rfs/ratio	2	2	2	4	15	MAX
	Rfs/ext	MAX	2	2	3	25	50
	Max IRfI	12	12	12	7	3	MAX

Table 3

Experiment 1. Numbers of sessions in which each parameter combination was used for each rat. "Mag. training" indicates additional magazine training sessions given after the beginning of shaping. "Backtracking" indicates the number of departures from the normal sequence of combinations (1 to 9).

Combination	Rat 3	5	6	7	8	9	10	11	12	14
Mag. training	—	3	1	—	—	—	—	—	—	—
1	2	8	11	2	2	2	4	2	2	2
2	3	5	5	3	3	3	4	3	3	3
3	3	3	3	—	3	2	3	3	3	3
4	3	3	3	1	—	—	3	3	3	3
5	2	2	3	1	2	1	3	3	3	2
6	—	—	3	1	—	1	2	—	3	3
7	1	2	7	1	—	3	11	1	3	5
8	—	3	9	1	1	2	3	—	7	4
9	2	—	—	3	4	6	3	3	7	—
10	—	—	—	—	—	—	—	—	—	10
11	—	—	—	—	—	—	—	—	—	10
Total	16	29	45	13	15	20	36	18	34	45
Backtracking	—	2	5	—	—	—	1	—	—	1

to an earlier combination because the reinforcement rate dropped severely; and two additional combinations (the last two in Table 2) were used with Rat 14, which was not progressing. The combinations used for each rat are shown in Table 3.

During shaping, the houselight and the light on the side wall above the open floor hole were normally lit, and the food tray light was off. When a pellet was delivered, the floor hole light was turned off and the food tray light was turned on. The lights were returned to their normal state when the food tray flap next closed, indicating that the food pellet had been collected. At this stage, if there were no ball bearings left in the chamber, one more was delivered to the ball bearing tray at the back of the chamber. Sessions started with six ball bearings in the chamber, but this number was reduced to one once a rat was reliably depositing them. Shaping was continued for a maximum of 45 sessions.

RESULTS

Eight of the 10 rats learned to deposit the ball bearings reliably within 45 sessions. These rats required a median of 14.5 sessions (range, 10 to 34) for shaping; if the two failures are included, the median rises to 21. Table 3, which reports the rats' progress through the parameter combinations listed in Table 2, gives an idea of the relative speeds of learning. Figure 2 gives more details of the rats' behavior during shaping and of the reinforcement frequencies for the precursor responses.

Of the 2 rats that did not acquire reliable ball bearing deposit within 45 sessions, Rat 6 was still making progress at the end of this time, and the changes in its behavior were qualitatively similar to those seen in the 8 successfully shaped rats. The other (Rat 14) showed a quite different pattern, with high frequencies of Move and Move Towards Hole behavior from Session 4 onwards, and the Reach Hole behavior never increasing. The rat would bring the ball bearing near to the hole and then move it backwards and forwards there, as though constrained not to let it go. This pattern of responding corresponds to what Boakes et al. (1978) described as misbehavior in a token deposit situation. On a smaller scale, similar behavior was shown by other rats, with temporary increases in Move and Move Towards Hole but not Reach Hole (or Move Towards Hole and Reach, but not Deposit) behavior without corresponding changes in reinforcement frequencies. The last two panels of data shown for Rat 12 in Figure 2 are an example: Move Towards Hole and Reach both increase, despite negligible reinforcement. Except in the case of Rat 14, however, the misbehavior effect was not strong enough to prevent progression to reliable ball bearing deposit.

In Figure 2, the data are arranged in the

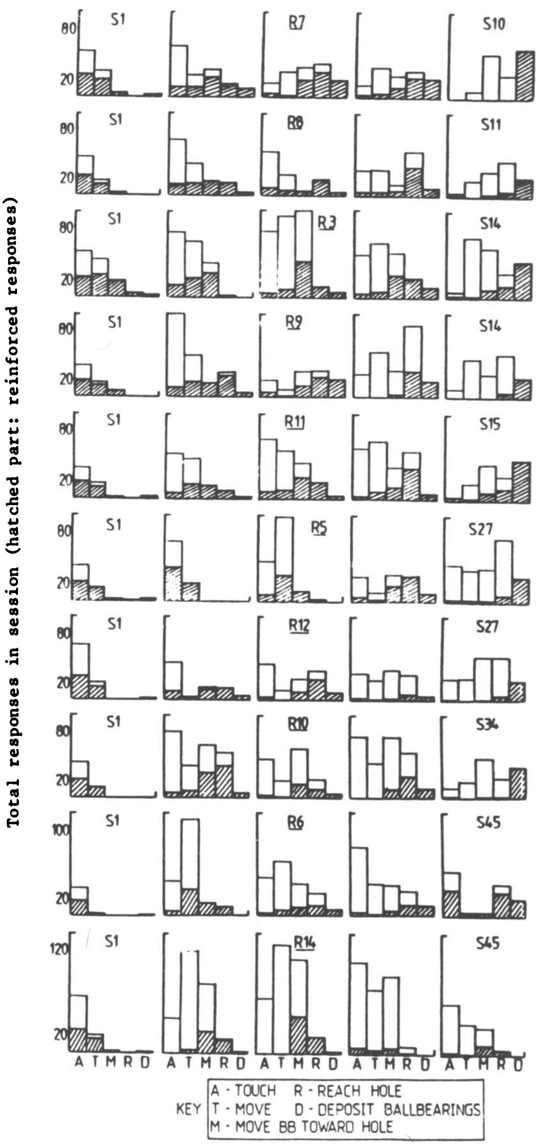

Fig. 2. Experiment 1. Response and reinforcement rates for ball bearing deposit and four precursor responses during the first session of algorithmic shaping, the last session in which parameter Combination 8 or below was used (see Table 2), and three approximately equally spaced intermediate sessions. Data are shown separately for each rat. Labels such as "R7" identify individual rats, labels such as "S10" identify session numbers. Open bars show response rates, hatched bars reinforcement rates. The five responses on the x axis are: A—touch ball bearing; M—move ball bearing; MB—move towards hole; R—reach hole (i.e., move the ball bearing in the vicinity of the hole); D—deposit ball bearing. For more detailed definitions of these responses, see Table 1. Note that where several responses occurred in an integrated sequence, only the final response of the sequence was scored.

order of the number of sessions required to achieve stable ball bearing deposit. An attempt was made to distinguish between fast and slow learners on the basis of the contingencies of reinforcement imposed by the shaping algorithm. However, there was no obvious pattern: Slow learners received as many reinforcements as fast learners in early sessions, for example. The more successfully shaped rats did emit more high-level responses early in training (Spearman's rho between sessions required and Move Towards Hole in Session 1 = -0.75, $p < .01$, two-tailed test), and it could be argued that they thus engaged with the reinforcement contingencies more quickly. But it could be said equally that they *required* less shaping because their preshaping performance was closer to the desired response.

Discussion

This first attempt at algorithmic shaping was reasonably successful, in that all but 2 of the subjects achieved reliable ball bearing deposit, and one of the failures showed substantial relevant behavior change. Furthermore the procedure made it possible to get an objective record of misbehavior that occurred during acquisition of a complex operant.

An important question, however, is whether algorithmic shaping is as effective as hand shaping, that is, the ordinary procedure in which the experimenter decides at the moment of observing behavior whether a particular response should be reinforced or not. Only if the algorithm works as well as an experienced experimenter can it reasonably be claimed to be a synthesis of what he or she does. To investigate this question, the time taken to train the rats to the point of token deposit was compared with a previous experiment (Midgley, 1986, Experiment 1) in which 6 rats were hand shaped to deposit ball bearings in the same apparatus. The median number of 10-min sessions required was 28 (range, 13 to 48) for the 5 rats that were successfully shaped; 1 rat was still untrained after 75 sessions. The results of the present experiment (eight successes out of 10, taking a median of 13.5 15-min sessions) are similar. Algorithmic shaping is certainly not noticeably worse than hand shaping for this task.

It should be noted that the present procedure was not completely algorithmic. First, the observer's coding of the various responses could have been affected by her knowledge of the experiment's aims and the rat's progress through the training procedure. This is an unavoidable limitation in shaping a complex operant. The experimenter's judgment was also involved in the progression from one parameter combination to the next, and because this was the first experiment of its kind, it was not practicable to establish a firm criterion for this in advance.

The algorithmic shaping procedure made it possible to observe the development of misbehavior, in the sense defined by Breland and Breland (1961) and Boakes et al. (1978). The data shown in Figure 2, though only an extract from the protocols recorded, are sufficient to show that intermediate types of behavior were not maintained by adventitious reinforcement, but rather increased in spite of decreasing reinforcement frequency, and increasing reinforcement frequencies for higher level types of behavior. But the present data cannot establish whether such misbehavior is due to the intrusion of unconditioned responses (as the Brelands argued), to Pavlovian contingencies (as Boakes et al. proposed), to motivational interactions (see Timberlake, 1983; Timberlake et al., 1982), or simply to reinforcement of the final response in the chain.

In terms of the categories introduced by Platt (1973), Figure 2 shows that successful shaping was accompanied by a gradual increase in the algorithm's selectivity: The range of reinforcement responses became steadily narrower. Contact with the reinforcement contingencies also tended to increase in the course of shaping; the early sessions for some subjects (e.g., Rat 7) show little differential reinforcement, whereas in the later sessions virtually all reinforcers were given following Deposit or Move Ball Bearing Towards Hole responses, and lower responses went unreinforced. Both these trends, however, seem to be more reflections of what the rats' changing repertoires made possible, rather than causes of behavior change.

EXPERIMENT 2

The first experiment demonstrated that algorithmic shaping could work, but the evidence that it is as successful as hand shaping depends on a comparison between experiments, which is not very satisfactory. The following experiment therefore involved a direct comparison between the two procedures. The shaping algorithm was slightly refined in the light of the results of Experiment 1: To reduce dependence on the experimenter's judgment, fewer parameter combinations were used and the criteria for progression between them were specified more exactly. In the hope of speeding up the shaping process, these criteria were also less demanding than in Experiment 1. Finally, to throw more light on misbehavior, the observer rated the rats' performance for apparent misbehavior at the end of each session. The intention was to see whether these ratings corresponded to any of the types of behavior being coded systematically. In so far as they do, it should be possible to study misbehavior by scoring those specific types of behavior, without doing violence to the more global idea of what constitutes misbehavior that is expressed in the observer's ratings.

METHOD

Subjects

Ten experimentally naive female Lister hooded rats were divided at random into two groups of 5. They were approximately 10 months old at the start of the experiment and their mean free-feeding body weights were 275 g (range, 254 to 306 g). They were maintained under the same conditions as the subjects of the first experiment; actual body weights during the experiment averaged 89% of free-feeding weights.

Apparatus

The experiment used the same apparatus as Experiment 1, although it was located in a different room in the same laboratory suite.

Procedure

All rats were trained to approach and eat from the food tray as in Experiment 1. Thereafter they were given one 20-min session per day, usually 7 days per week, but the procedure for the two groups differed as follows:

Algorithmic shaping group. The observer (always the first experimenter) recorded all the types of behavior listed in Table 1, except ball bearing deposit, which was recorded automatically; Mouth and Chew were recorded in this experiment, although no contingencies of reinforcement were applied to them. The shaping algorithm described in Experiment 1

Table 4

Experiment 2. Codes used to rate the subjective impression of the amount of misbehavior shown in each session.

Code	Definition
1	No misbehavior: The rat handled the ball bearings without displaying any misbehavior.
2	Efficient handling: At the beginning of training, used for rats that very occasionally handled the ball bearings more than was necessary for reinforcement; at the end of training, used for rats that handled the ball bearings in a stereotyped manner.
3	Excessive handling of no more than 25% of the ball bearings.
4	Misbehavior towards between 25% and 50% of ball bearings (e.g., carrying the ball bearing towards the hole and then away again, apparent conflict in letting the ball bearing drop down the hole, or chewing, mouthing, and excessive rolling of ball bearings).
5	Similar behavior as in (4) but shown towards over half the ball bearings.
6	Similar behavior as in (4) but to all or virtually all the ball bearings.
7	Very excessive handling of all ball bearings, combined with sitting for long periods of time holding or mouthing ball bearings, or rolling them to and fro.

was used. The parameter combinations used were those listed in Table 2, with the exception of Combination 1. The original intention was to use each combination for 2 days for each rat; however, because the rats in the present experiment made more rapid progress than those in Experiment 1, exposure to Combinations 5, 6, 7, and 9 was reduced to 1 day, and Combination 8 was omitted for 4 of the rats and Combination 7 for 3 of these. In no case was it necessary to give extra magazine training or to return to an earlier parameter combination.

Hand shaping group. The observer (always the first experimenter) reinforced approximations to ball bearing deposit directly, using one of the keys on the microcomputer keyboard. Ball bearing deposit was reinforced immediately by the computer program. As in algorithmic shaping, the session started with six ball bearings present in the test chamber, and an additional ball bearing was delivered whenever the last one present was deposited. No record of behavior, other than ball bearing deposit and food tray entry, was made.

For both groups, signal lights and food reinforcement were used in the same way as in Experiment 1. Following successful training to deposit ball bearings, all rats were given two sessions in which ball bearing deposit was the only reinforced response, and the observer recorded all the other responses listed in Table 1. At the end of each session, the observer recorded her subjective impression of the amount of misbehavior shown, using the scale given in Table 4.

Results

All rats were shaped successfully. Algorithmic shaping took a median of 8 sessions (range, 8 to 11), and used a median of 827 reinforcers (range, 738 to 839); hand shaping took about as long on average, but was slightly more variable (median of 7 sessions, range, 6 to 14; median of 534 reinforcers, range, 367 to 1,103). Mean reinforcers per session were somewhat higher for algorithmically shaped than for hand shaped rats (medians of 92.3 and 78.8 reinforcers per session, Mann–Whitney $U_{5,5} = 2$, $p < .05$).

Table 5 examines the effects of two crucial steps in the shaping algorithm. Both were concerned with the rat's progress through the response hierarchy. Step 4 (see the Introduction to Experiment 1) specified that once a given response had been reinforced a given number of times, all responses lower in the hierarchy were subjected to extinction; Step 5 specified that if no response was reinforced within a fixed time, reinforcement of the next lowest response that was still occurring at a high rate was reinstated. The table shows, for each shaping session of each rat in the algorithmic shaping group, the number of times upward and downward steps in the response hierarchy were taken. Both the upward and downward steps could take the rat through more than one level of the hierarchy at a time, so the table also shows the total number of levels upward and downward through which the rats moved. Note that although there were only six responses in the hierarchy, a session in which there were several steps upwards and downwards could involve moving though more than six levels in total (e.g., Rat 14, Sessions 6, 10, and 11).

Figure 3 shows response and reinforcement frequencies of the five precursor responses in five sessions for each rat in the algorithmic shaping group. It can be seen that frequencies of two precursors (Touch and Move Towards

Table 5

Experiment 2, algorithmic shaping group. Effect of the shaping algorithm on the minimum response required for reinforcement, session by session. Entries are the number of changes in the minimum response level required for reinforcement, upward (i.e., in the direction of closer approximation to the ball bearing deposit response) and downward. The second figure in each column gives the total number of response levels through which the algorithm moved in each direction during the session.

	Rat									
	4		6		9		12		14	
Session	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
1	2, 2	1, 1	2, 2	1, 1	2, 2	—	1, 1	—	2, 2	1, 1
2	2, 3	1, 2	2, 2	1, 1	2, 3	1, 2	2, 3	—	1, 1	—
3	2, 3	—	2, 3	1, 2	2, 3	1, 2	3, 3	1, 2	2, 2	1, 1
4	2, 3	1, 2	2, 5	1, 2	2, 3	1, 2	2, 3	—	2, 3	1, 2
5	2, 5	1, 2	2, 5	1, 2	1, 3	1, 2	2, 4	1, 3	2, 5	2, 4
6	2, 5	2, 4	2, 5	1, 2	1, 3	1, 2	3, 4	—	3, 7	3, 6
7	1, 3	—	2, 5	1, 2	1, 4	1, 1	2, 4	1, 1	2, 4	1, 1
8	2, 5	1, 2	2, 5	1, 2	2, 5	—	2, 4	—	1, 3	1, 2
9	2, 4	—							4, 6	2, 3
10									3, 9	2, 8
11									4, 16	5, 13

Hole) remained high despite low or zero reinforcement rates.

The observer's ratings of misbehavior were examined for correlation, across sessions for each individual rat, with the raw frequencies of each reinforced precursor and with their frequencies divided by the number of reinforcers delivered (for all responses) in the session. As regards raw frequencies, the most consistent correlation was obtained with the number of Mouth behaviors recorded per ball bearing deposited (mean Spearman's rho value = 0.69, $p < .05$ for every rat). When the frequencies per reinforcement were considered, the precursor behavior most consistently correlated with the observer's ratings of misbehavior was Touch (mean Spearman's rho between Touches per reinforcer delivered and misbehavior rating = 0.68, $p < .01$ for 3 rats, nonsignificant for the other 2). Ratings of misbehavior in the hand shaped group were not significantly different from those in the algorithmic shaping group. They could not, of course, be tested for correlation with rates of precursor behavior, because those types of behavior could not be recorded during hand shaping.

Discussion

Experiment 2 confirmed that algorithmic shaping is possible, and that it can be as efficient as hand shaping. The results of algorithmic shaping were, indeed, somewhat more consistent. This may be a reflection of the slightly higher reinforcement rates it tended to produce; perhaps algorithmically trained rats are less likely than hand shaped rats to be subjected to periods of extinction.

Both algorithmic and hand shaping were more efficient in the present experiment than previously (Experiment 1 of the present paper and Experiment 1 of Midgley, 1986). This may simply reflect a difference between the subject samples (although the rats in the two experiments came from the same strain and supplier, they differed in age), or it may be an effect of greater skill in the observer's recording of behavior. The latter explanation is suggested by the fact that more unreinforced precursor responses were recorded in the present experiment than in Experiment 1, suggesting that the observer was following rapid sequences of behavior more closely. In interobserver reliability trials it was found that less experienced observers coded many fewer responses than the present observer, although their patterns of response frequencies were very similar.

Table 5 shows that the minimum response required for reinforcement frequently varied during algorithmic shaping sessions. All sessions involved at least some upward movement in the response hierarchy as higher level responses increased in rate; the great majority also involved some downward movement as reinforcement rate fell. The table also shows

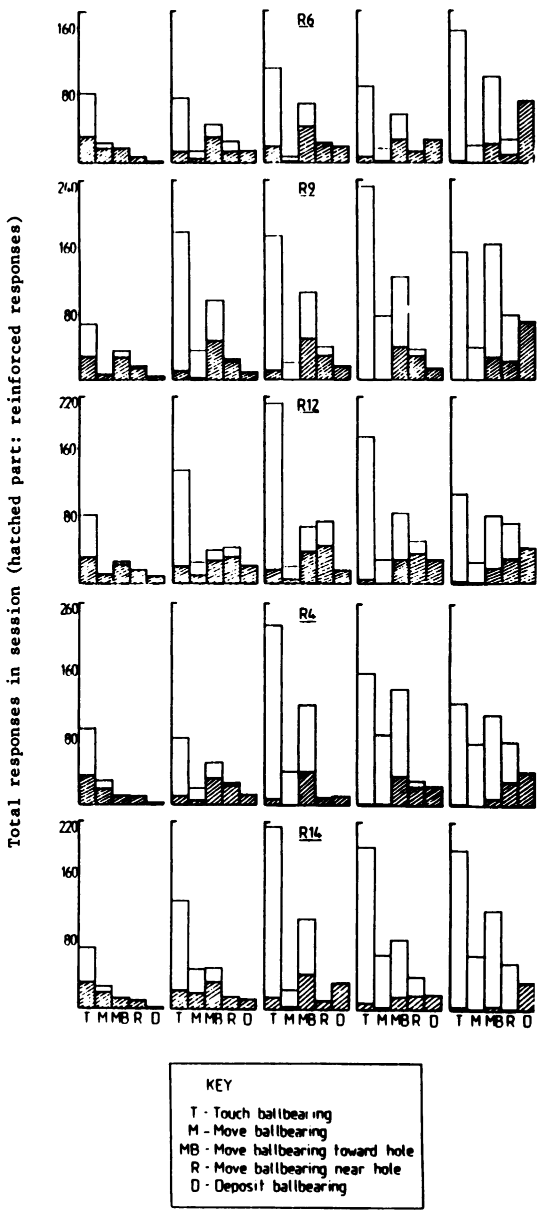

Fig. 3. Experiment 2. Response and reinforcement rates for ball bearing deposit and four precursor responses during the first and last sessions of shaping and three equally spaced intermediate sessions; Data are shown for each rat in the algorithmic shaping group. T—touch ball bearing; M—move ball bearing; MB—move towards hole; R—reach hole (i.e., move the ball bearing in the vicinity of the hole); D—deposit ball bearing. For more detailed definitions of these responses, see Table 1. Other details as in Figure 2.

that many of the movements involved more than one level of the response hierarchy, because in the majority of cases the number of movements made (up or down) is less than the number of response levels moved through. These results confirm that the rats' behavior was making contact with the rules of the algorithm. As in Experiment 1, contact and selectivity as defined by Platt (1973) both increased in the course of shaping (see Figure 3).

The data from the algorithmic shaping group show the development of misbehavior very clearly. All these rats showed increases in some precursor behavior during shaping, despite falling reinforcement rates for them. The correlational data show that the observer's ratings of misbehavior were most reliably correlated with the Mouth behavior, and this was never directly reinforced because it was not defined as a precursor. It could of course form part of reinforced behavior such as Move Towards Hole, and some types of precursor behavior also correlated with the observer's misbehavior ratings. Methodologically speaking, these two results show that it should be possible to investigate the origins of misbehavior, during shaping, by using objectively defined behavior categories.

GENERAL DISCUSSION

The results of these two experiments confirm that, if the experimenter's role is confined to passively recording the frequencies of responses and successively closer approximations to a desired response are selectively reinforced, the desired response does indeed emerge. This is in accordance with the results of Eckerman et al. (1980), Platt and his colleagues (e.g., Alleman & Platt, 1973; Davis & Platt, 1983), and Pear and Legris (1987); it extends their work to the case of a complex operant emerging slowly through a chain of observable precursors that are specifically reinforced. As Pear and Legris put it, such experiments "begin to move shaping from the realm of art to that of science" (p. 241).

The present data also show that the target response will not necessarily dominate behavior; precursor responses, and responses never directly reinforced, remain at high levels. At least some of these are correlated with an experienced observer's subjective ratings of misbehavior. Clearly, though, the occurrence of such misbehavior does not necessarily make shaping impossible.

These experiments extend substantially the scientific base on which the practice of shaping is founded. They do so not because they are on an impressive scale themselves, but because there has been so little systematic work on the subject previously. However, they leave a number of important questions unanswered.

The first is the genesis of the precursor responses. We chose a particular set of precursors for use in these experiments from our experience of hand shaping the target response. Could precursors be selected in some less arbitrary way? In this respect the use of a response continuum, as in the experiments of Pear and Legris (1987), has advantages over a heterogeneous sequence: Deciding to shrink the reinforced virtual sphere towards a target location involves much less experimenter judgment than does deciding which of many discrete responses to include in a precursor hierarchy.

Second, the shaping algorithm we used is, of course, not the only one possible. We attempted to include in it all the reinforcement contingencies and variations in contingencies that we felt we had used during previous experiments involving hand shaping of ball bearing deposit. It included large numbers of parameters, so that we could try to reproduce the conditions of hand shaping quite precisely. But from the point of view of exploring the determinants of shaping, a much simpler algorithm might well be better—and there is no reason to suppose that it would not work. Table 5 shows that the key features of the algorithm were all invoked in typical shaping sessions, but it should be possible to produce simpler rules that incorporate those features while having far fewer parameters. The interval percentile schedule proposed by Platt (1973) would be one possibility. The data in Table 5 do suggest that straightforward percentile reinforcement would not be wholly satisfactory, because of the frequency with which the time-based backtracking rule was invoked (percentile reinforcement ignores the distribution of responses in time).

Finally, the present experiments are only a demonstration that algorithmic shaping can work. If the algorithm had failed, we should have had to rethink our account of shaping in terms of the reinforcement of successive approximations to the desired response. But much stronger tests of that account are possible by simple variations in the shaping algorithm. Suppose, for example, that we had arranged the responses in the hierarchy in reverse order, or made reinforcement probabilities increase instead of decreasing as response rates increased. Such manipulations ought to have dramatic effects on the success of shaping. The possibilities for further experimentation (which may not need to use such difficult responses as ball bearing deposit) are considerable.

REFERENCES

Alleman, H. D., & Platt, J. R. (1973). Differential reinforcement of interresponse times with controlled probability of reinforcement per response. *Learning and Motivation*, **4**, 40–73.

Boakes, R. A., & Jeffery, G. (1979). Automodellaggio e malcomportamento. *Ricerche di Psicologia*, **10**, 53–68.

Boakes, R. A., Poli, M., Lockwood, M. J., & Goodall, G. (1978). A study of misbehavior: Token reinforcement in the rat. *Journal of the Experimental Analysis of Behavior*, **29**, 115–134.

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, **16**, 681–684.

Brown, P. L., & Jenkins, H. M. (1968). Autoshaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, **11**, 1–8.

Davis, E. R., & Platt, J. R. (1983). Contiguity and contingency in the acquisition and maintenance of an operant. *Learning and Motivation*, **14**, 487–512.

Eckerman, D. A., Hienz, R. D., Stern, S., & Kowlowitz, V. (1980). Shaping the location of a pigeon's peck: Effect of rate and size of shaping steps. *Journal of the Experimental Analysis of Behavior*, **33**, 299–310.

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

Hingtgen, J. N., & Trost, F. C., Jr. (1966). Shaping cooperative responses in early childhood schizophrenics: II. Reinforcement of mutual physical contact and vocal responses. In R. Ulrich, T. Stachnik, & J. Mabry (Eds.), *Control of human behavior* (pp. 110–113). Glenview, IL: Scott Foresman.

Midgley, M. (1986). *Token reinforcement and misbehaviour in the rat*. Unpublished doctoral dissertation, University of Exeter.

Myer, J. S., & Hull, J. H. (1974). Autoshaping and instrumental learning in the rat. *Journal of Comparative and Physiological Psychology*, **86**, 724–729.

Pear, J. J., & Eldridge, G. D. (1984). The operant-respondent distinction: Future directions. *Journal of the Experimental Analysis of Behavior*, **42**, 453–467.

Pear, J. J., & Legris, J. A. (1987). Shaping by automated tracking of an arbitrary operant response. *Journal of the Experimental Analysis of Behavior*, **47**, 241–247.

Platt, J. R. (1973). Percentile reinforcement: Paradigms for experimental analysis of response shaping. In G.

Bower (Ed.), *The psychology of learning and motivation* (Vol. 7, pp. 271–296). New York: Academic Press.

Powell, R. W., & Kelly, W. (1975). A method for the objective study of tool-using behavior. *Journal of the Experimental Analysis of Behavior,* **24,** 249–253.

Shettleworth, S. J. (1975). Reinforcement and the organization of behavior in golden hamsters: Hunger, environment, and food reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes,* **1,** 56–87.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review,* **24,** 86–97.

Skinner, B. F. (1962). Two "synthetic social relations." *Journal of the Experimental Analysis of Behavior,* **5,** 531–533.

Timberlake, W. (1983). The functional organization of appetitive behavior: Behavior systems and learning. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behaviour: Vol. 3. Biological factors in learning* (pp. 177–221). Chichester, England: Wiley.

Timberlake, W., Wahl, G., & King, D. (1982). Stimulus and response contingencies in the misbehavior of rats. *Journal of Experimental Psychology: Animal Behavior Processes,* **8,** 62–85.

Received July 8, 1988
Final acceptance January 19, 1989