

BEPROC 00549

Learning new response sequences

Alliston K. Reid

Department of Psychology, Eastern Oregon State College, La Grande, OR 97850, USA

(Accepted 21 January 1994)

Abstract

Four rats were required to press either a right or left lever to complete various three-response sequences. After extended exposure to a training sequence, subjects were shifted to a new target sequence. The new target sequences always differed from the previous sequence by the response required in the first or last position of the sequence. Subjects were repeatedly exposed to all possible combinations of training and new target sequences. Learning of new sequences occurred more rapidly when the change in the new target sequence was in the last position. Errors persisted longer in new sequences in which the change was in the first position. Extinction of the training sequence occurred faster when the change was in the last position. Responses in the last position were considerably more sensitive to the shift to new target sequences than were responses in the first position. Even though response sequences may form new behavioral units from the training sequence, reinforcement and extinction acted differentially on the individual lever presses within new target sequences rather than on the sequences as a whole. These findings support the hypothesis that response strength is determined by contiguity to reinforcement.

Key words: Acquisition; Animal learning; Reinforcement theory; Response strength; Sequence learning; Rats

Introduction

What is the process of response-sequence learning in animals? Two separate lines of research have provided partial answers to this question.

One line of research assumes that reinforcement acts on the individual responses within a sequence in such a way that their contiguity to reinforcement is the main determinant of their response strength (cf. Nevin, 1974; Skinner, 1938). That is, reinforcement acts on

* Corresponding author. Tel: (503) 962 3580.

individual responses within a sequence (rather than on the sequence as a whole), and those responses that are immediately followed by reinforcement will typically have greater strength than do responses less contiguous with reinforcement (Nevin, 1974, 1979; Nevin et al., 1981; Wasserman et al., 1986; Williams et al., 1992). I will call this assumption the contiguity hypothesis.

Several theories of behavior are in accordance with the contiguity hypothesis. Perhaps the most well known is chaining theory (Keller and Schoenfeld, 1950; Skinner, 1938, 1950; Straub and Terrace, 1981; Terrace, 1986; Wasserman et al., 1986; Williams et al., 1992) which proposes that each response in the sequence receives its strength from the conditioned reinforcing properties of the resulting stimuli, except the last response which receives its strength from primary reinforcement. In contrast, Staddon (1983) argues that the temporal distance to food delivery determines the strength of the response. Similarly, several recent models of habituation (Staddon, 1993), recurrent choice (Davis et al., 1993), temporal learning (Staddon and Higa, 1991), and most connectionist models of learning and behavior (e.g. Commons et al., 1991; Gabriel and Moore, 1990) rely heavily on the concept of response strength and the contiguity hypothesis.

It is important to point out that the contiguity hypothesis does not specify the nature of the associations that are learned. Nor does it specify the mechanism by which responses receive their value. It is a simple statement of the functional relation between the strength of responses and their contiguity with reinforcement. The various theories that are in accordance with the contiguity hypothesis frequently differ in nature of the assumed associations, the means of acquiring response strength, and the shape of the function relating response strength to temporal distance from reinforcement.

A second line of research emphasizes that contingent reinforcement of response sequences often results in the formation of integrated behavioral units comprised of the individual responses in the reinforced sequence (e.g. Adams, 1984; Pisacreta, 1982; Reed, Schachtman & Hall, 1991; Schneider and Morris, 1992; Schwartz, 1980, 1986; Shimp, 1979). That is, reinforcement organizes the individual responses into highly stereotyped sequences which often function as single responses. With continued exposure to a schedule which reinforces a particular sequence of behavior, reinforcement comes to control the complete sequence just as it normally controls individual responses. The main criterion for assessing whether a response sequence functions as a behavioral unit is to reinforce the putative unit according to an additional contingency and see if its rate or probability changes in ways sufficient to demonstrate conditioning (Zeiler, 1977; Shimp, 1979). The putative unit should remain unaltered on the new schedule, maintaining its integrity, timing, and accuracy (Reed et al., 1991; Thompson and Lubinski, 1986; Zeiler, 1977, 1986).

For example, Schwartz (1981) trained pigeons to complete an eight-response sequence of key pecks on left and right keys. Subsequently, he repeatedly exposed the birds to extinction. Sequence variability and sequence completion time were not affected. Similarly, Schwartz (1986) shifted similarly trained pigeons to multiple variable-interval (VI) schedules and concurrent VI schedules. Highly stereotyped response sequences developed. As typically observed with individual responses, he observed behavioral contrast of the stereotyped sequences on the multiple schedules and the matching relation on the concurrent schedules. Therefore, these stereotyped sequences functioned as behavioral units.

This general finding has been widely replicated and extended to other schedule manipulations. For example, with sufficient training response sequences function as behav-

ioral units on discrete trials procedures (Grayson and Wasserman, 1979; Pisacreta, 1982; Schwartz, 1980, 1981, 1985; Schwartz and Reilly, 1985; Wasserman et al., 1986), free operant procedures (Arbuckle and Lattal, 1988; Reed et al., 1991; Schwartz, 1982; Shimp, 1979), and multiple and concurrent schedules (Fetterman and Stubbs, 1982; Schneider and Morris, 1992; Schwartz, 1986; Stubbs et al., 1987). In these studies behavioral units were composed of sequences of two or more simple responses such as left and right lever presses or key pecks. Other researchers have extended the notion of the behavioral unit to sequences of other elements, such as interresponse times (IRIs) or particular temporal patterns of responding (Adams, 1984; Reed et al., 1991; Shimp, 1979; Stubbs et al., 1987; Tarpy and Roberts, 1985; Zeiler, 1977).

The characterization of a sequence of behavior as a behavioral unit is descriptive of the way responding becomes organized with extended exposure to a reinforcement schedule and the way the sequence as a whole is influenced by reinforcement or manipulations of the schedule of reinforcement. If reinforcement appears to act on the individual responses within a sequence, rather than on the sequence as a whole, then one would conclude that the sequence was not acting as a behavioral unit. This conclusion would be true even though the sequence may continue to be highly structured and repetitive.

This paper is primarily concerned with the way in which new response sequences are learned after extended training on other sequences. Do changes in reinforcement act on the individual responses within the sequence or on the sequence as a whole? Even though the classification of a behavioral unit is descriptive rather than explanatory, I will call the behavioral unit hypothesis the prediction that changes in reinforcement will act on sequences of responses as a whole without causing differential effects on the individual responses within the sequence (such as facilitating some responses but not others).

It is useful to compare the roles played by the behavioral unit hypothesis and the contiguity hypothesis when animals are required to complete new response sequences. Consider the predictions made by the two hypotheses in the following procedure designed for this study.

Suppose rats are trained to complete a three-response sequence composed of left and right lever presses. Food is delivered immediately after the successful completion of the training sequence, and timeout occurs after the completion of all other sequences. No feedback is provided within trials to indicate the accuracy of individual responses. After accuracy is well established and stable on this training sequence, the contingency is changed in the middle of a session: a different response sequence becomes the new target sequence. The previous training sequence is no longer reinforced. The rats are exposed to this new target sequence for several days until accuracy is stable. Over a period of weeks, each rat is repeatedly exposed to each possible transition from training sequence to new target sequence. This repeated exposure allows within-subject comparisons of the changes in response probability in each position in the sequence as a function of the previous training sequence. The primary question is, how does learning the new sequence depend upon characteristics of the previous training sequence?

Allow the original training sequence to be represented by the following sequence: $R_1 R_2 R_3$. Suppose that the possible transitions to new target sequences are limited to only two types: those in which the new target sequence differs from the training sequence by the required response in the first position ($R_1 R_2 R_3 \rightarrow R_1^* R_2 R_3$) or in the last position ($R_1 R_2 R_3 \rightarrow R_1 R_2 R_3^*$) of the sequence.

What are the specific predictions made by the two hypotheses that can be tested with this procedure? The behavioral unit hypothesis was developed to explain the way behavior

becomes organized during steady-state exposure to certain reinforcement schedules. It was not proposed as a model of the dynamics of acquisition of response sequences. Nevertheless, it can be extended to make several clear predictions about the way learning of new sequences should occur. It proposes that ample exposure to the training sequence will cause the response sequence to function as a single response unit. When the contingency is changed, the sequence as a unit should extinguish. Extinction of the training sequence should occur at the same rate whether the two sequences differ in the required response in the first or the last position. Therefore, the extinction curves for both types of transitions should not differ. If they do differ, then it would be evidence that reinforcement was acting on the individual responses within the unit, rather than on the unit as a whole. In this case, the behavioral unit hypothesis would not adequately account for the changes observed when learning new sequences, although it may have accurately described responding on the training sequence. In addition, the speed of learning to complete the new target sequence should not be a function of the type of transition. Finally, the different response positions in a sequence should not show differential sensitivity to new target sequences. Rather, the changes in responding in the first and last positions should be the same.

In contrast, the contiguity hypothesis predicts differences in the speed of learning in the two conditions. It assumes that the different responses comprising the sequence will vary in response strength depending upon their contiguity with reinforcement. The last response in the sequence should have the greatest response strength, and its strength should change most rapidly after the transition to a new target sequence because of the inhibitory effects of nonreinforcement. More detailed predictions require a more precise, quantitative model of how response strength is related to temporal distance to reinforcement and how it changes as a function of extinction. Assuming the relation between response strength and temporal contiguity to food (after repeated reinforcement of a single sequence) is positively accelerated (Staddon, 1983), if the changes in response strength due to nonreinforcement are steeper than this relation, extinction of the training sequence should be faster when the required response in the last position differs from the new target. In this case, response strength and resistance-to-change (Nevin, 1974; Nevin et al., 1981) should be inversely related after the transition. In contrast, if the changes in response strength due to nonreinforcement are shallower, extinction should occur faster when the required response in the first position differs from the new target. In this case response strength and resistance-to-change should be directly related.

Materials and Methods

Subjects

Four female rats, reared in this laboratory, served as subjects. All animals were approximately four months old at the beginning of the experiment. Subjects were maintained at 80% ad lib body weight and housed individually in home cages with free access to water. The animal room was maintained at an approximately constant temperature and humidity with a 12:12 h light/dark cycle. Supplemental feeding was provided after experimental sessions using Purina Rat Chow.

Apparatus

Four identical BRS two-lever rat chambers (30 × 23 × 24 cm) were used. A Gerbrands Feeder dispensed 45 mg Noyes pellets (Improved Formula A). The two response levers were located 8.5 cm over the floor and separated by 10.5 cm. Centered between these two levers was the feeder tray, located 2.5 cm above the floor. Three 28-V white stimulus lamps (Sylvania 28ESB) were centered 6.0 cm over the two levers and over the feeder. Two 28-V house lights (GE-1819) were located at the top of the rear wall. Individual microcomputers controlled the experiment and collected data. Every bar press and stimulus change was recorded along with their times of occurrence and later transferred to diskettes for permanent storage and subsequent data analysis.

Procedure

Pretraining

Subjects were shaped to press levers by a successive-approximations procedure after their body weights were reduced to 80% ad lib. Subjects were then exposed to continuous reinforcement and a fixed-ratio (FR) 4 schedule for two or three days. They were then shifted to a "changeover" condition in which pellet delivery was contingent upon pressing both levers at least once in any order. This condition ensured that subjects received approximately equal exposure to both levers. After seven sessions of this condition, pellet delivery became contingent upon two changeovers in any order. Subjects were exposed to this modified changeover condition for approximately seven sessions.

Experimental procedure

Sessions lasted until 150 pellets were delivered, or 90 min had elapsed, whichever came first. The experimental procedure involved training a three-response sequence in which each response consisted of either a right or left lever press, resulting in eight possible sequences. Subjects were exposed to a single training sequence for several days until session duration and accuracy were stable. The accuracy criteria were that for three consecutive days, subjects had to complete more sequences correctly than the sum of all the incorrect sequences, and the percentage of correct sequences could not increase or decrease systematically for the last three days. They were then shifted to a new target sequence in which the required response in either the first or last position of the original training sequence was altered. After behavior on this new target sequence stabilized (over several sessions), this sequence was treated as a training sequence leading to a new target sequence. The required response in the middle position always remained the same in the training and new target sequences. Therefore, the eight possible sequences were divided into two groups of sequences, differing only in the required middle response.

Table 1 depicts these two groups of sequences and shows how changes in the response requirement produced new target sequences. Changes in the response required in the first or last position of a sequence produced the new target sequence that is depicted in the same column or row, respectively. The order of the target sequences for each subject was determined by clockwise or counter-clockwise rotations around the groups of sequences depicted in Table 1. Therefore, as new target sequences were selected, the required response in the first or last position changed in an alternating fashion.

The experimental procedure was divided into four similar phases, depicted in Table 2. In each phase each subject was exposed to each of the four target sequences in one group,

TABLE 1

Procedure for generating the series of target sequences in each phase.

Group 1	Group 2
Middle Response = Left	Middle Response = Right
LLL ↔ LLR	LRL ↔ LRR
‡ ‡	‡ ‡
RLL ↔ RLR	RRL ↔ RRR

The eight possible sequences were divided into two groups based on their common middle response. Changes in the response required in the first or last position of a sequence produced the new target sequence depicted in the same column or row, respectively. The order of sequences in each phase was produced by a clockwise or counter-clockwise rotation around Group 1 or Group 2.

beginning and ending with the same sequence. Two subjects began with the sequences requiring a left lever press in the middle position, and two subjects began with the sequences requiring a right lever press. Both subjects exposed to each group of sequences began with the same sequence, but differed in the subsequent target sequence. That is, the direction of rotation in Table 1 was different for both subjects. As each subject was shifted to the next phase, the required middle response was different (i.e. subjects were shifted to the other group of sequences), and the direction of rotation was changed. The reason for changing directions was to ensure that subjects could not learn an alternating pattern of required changes in the target sequences.

Therefore, all subjects were exposed to all target sequences twice. For each subject, the training sequences leading to a given target sequence differed by the required response in the first position once, and by the last position once. Thus, the total number of transitions

TABLE 2

Order of sequences for each subject.

	Rat	Group	1	2	3	4	5
Phase 1	Rat 1	1	RLL	LLL	LLR	RLR	RLL
	Rat 2	2	LRL	LRR	RRR	RRL	LRL
	Rat 3	1	RLL	RLR	LLR	LLL	RLL
	Rat 4	2	LRL	RRL	RRR	LRR	LRL
Phase 2	Rat 1	2	LRR	LRL	RRL	RRR	LRR
	Rat 2	1	RLR	LLR	LLL	RLL	RLR
	Rat 3	2	LRL	LRR	RRR	RRL	LRL
	Rat 4	1	RLR	RLL	LLL	LLR	RLR
Phase 3	Rat 1	1	LLR	LLL	RLL	RLR	LLR
	Rat 2	2	RRL	RRR	LRR	LRL	RRL
	Rat 3	1	LLR	RLR	RLL	LLL	LLR
	Rat 4	2	RRL	LRL	LRR	RRR	RRL
Phase 4	Rat 1	2	RRR	RRL	LRL	LRR	RRR
	Rat 2	1	LLL	LLR	RLR	RLL	LLL
	Rat 3	2	RRR	LRR	LRL	RRL	RRR
	Rat 4	1	LLL	RLL	RLR	LLR	LLL

from training to new target sequences for all animals in all phases was 64. Of these transitions, 32 differed by the response required in the first position, and 32 differed by the last position.

Changes from a training sequence to a new target sequence always occurred on the 51st trial within a session. This allowed time for the subjects to stabilize on the training sequence before the shift to the new target sequence. Shifts to the new sequence were not signaled. After shifting to a new target sequence, this sequence remained in effect for several days until accuracy and session duration were stable. This sequence was then treated as a training sequence.

Correct sequences resulted in a pellet delivery after a 0.3-s delay. The next trial began immediately after pellet delivery. Incorrect sequences resulted in an 5-s timeout after the third response. During timeout the house lights were turned off, and responses had no programmed consequences. The first two responses in all trials produced a brief 0.3-s flash on all three identical stimulus lights located over the two levers and the hopper. The third, final response in the trial did not affect these lights, but produced timeout or a pellet delivery. No feedback was provided within a trial to indicate correct or incorrect responses.

Results

The procedure produced two types of transition from training to new target sequences. The target sequence could differ by the lever press required only in the first or last positions. The two panels of Fig. 1 show the errors on each trial summed across all four subjects in the 64 sessions in which the target sequence was changed. In the top panel, the two curves represent the errors committed when the training and new target sequences differed by the required response in the first position (unfilled circles) or the last position (filled circles) in the sequence. The target sequences contributing to both curves were the same: each of the eight possible target sequences is represented in both curves four times, with the difference in curves resulting only from the type of transition that produced the target sequence. Thus, the maximum on the ordinate would be 32, corresponding to an error at that trial number in every session containing that type of transition from training to target sequence.

A value of 300 on the abscissa does not necessarily mean that all subjects completed 300 trials in each session. Rather, sessions ended after 150 pellets were delivered or after 90 min, whichever came first. The number of completed trials in the sessions used for this figure ranged from 96 to 600 ($\bar{x} = 329.4$). Therefore, the number of errors represented in the top panel of Fig. 1 reflect only completed trials. If a subject received the maximum number of pellets or failed to respond any longer in the session, the number of completed trials in that session may not equal 300. This panel ignores uncompleted trials since if trials did not occur, an error (or a correct sequence) could not occur. In order to identify trends more clearly, a three-term moving average was used to smooth the curves in all figures.

When the rats were shifted to a new target sequence on the 51st trial, errors immediately increased. A side effect of using a moving average is to smooth the abrupt change in errors that actually occurred.

The top panel of Fig. 1 shows that when the sequences differed by the response required in the first position, errors persisted longer than when the difference was in the last position. This difference between curves appeared after approximately 190 trials in the session.

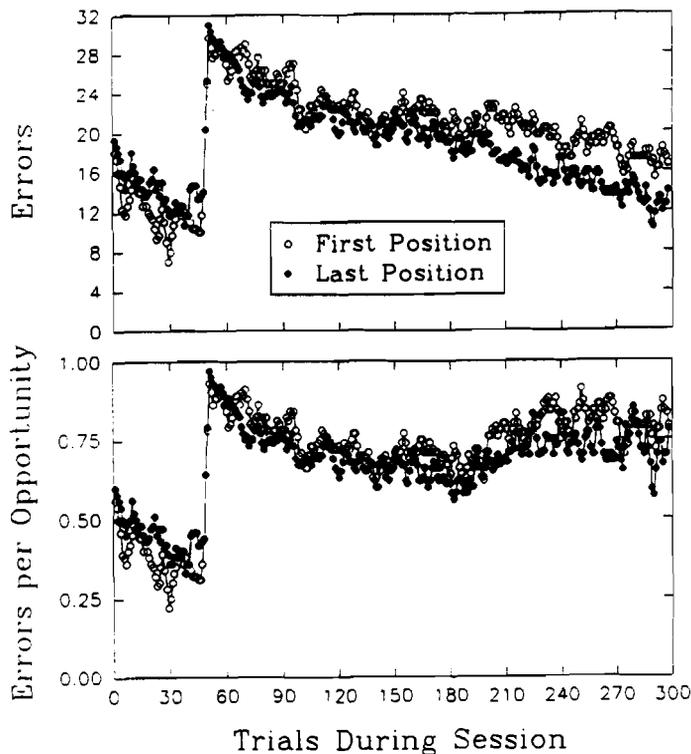


Fig. 1. Top panel: curves represent errors committed when the new target sequence differed from the previous training sequence by the response required in the first position (unfilled circles) or the last position (filled circles) in the sequence. On the 51st trial, subjects were switched to the new target sequence. Bottom panel: errors were divided by the number of sequences actually completed to produce errors per opportunity.

The difference between curves could be due to a difference in the number of trials per session actually completed. Although the same target sequences contributed to both curves equally often, subjects could vary the number of sequences actually completed from session to session. In order to determine if the number of completed sequences contributed to the difference between conditions, the bottom panel depicts errors per opportunity. Errors per opportunity were calculated by dividing the sum of the errors in each trial (from the top panel) by the number of times that trial was actually completed in the condition.

In the bottom panel of Fig. 1, the systematic difference between conditions observed in the top panel was substantially reduced by dividing by the number of completed trials. Errors per opportunity decreased on the new target sequences, then increased after approximately 190 trials. When sessions contained a large number of trials, the relative frequency of errors in completed trials was high late in the session in both conditions. This higher frequency is due to some sessions not being represented in the total because the subject had already received all 150 pellets before that trial number could occur. This increase occurred at approximately the same point within sessions where the curves in the top panel separated. When the new target sequence required a different response in the

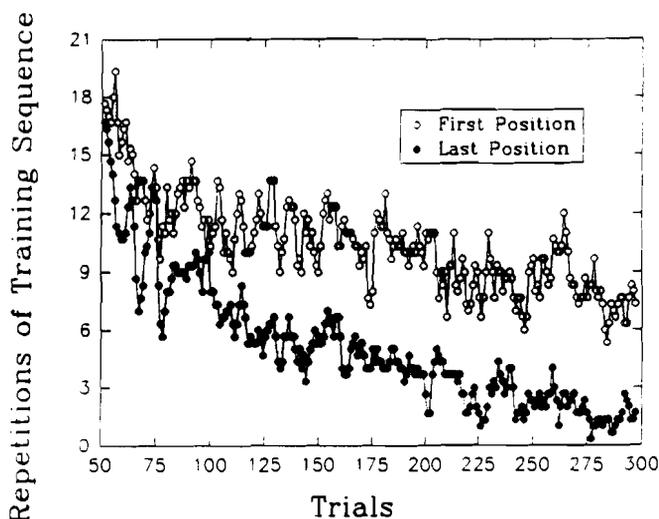


Fig. 2. Number of repetitions of the previous training sequence observed after the switch to the new target sequence. Points represent trials in which the training and target sequences differed by the required response in the first position (unfilled circles) or the last position (filled circles) in the sequence.

first position (unfilled circles), sessions contained slightly more completed trials ($\bar{x} = 335$) than when the difference was in the last position ($\bar{x} = 324$). Therefore, much of the difference between the conditions observed in the top panel was due to the small, but systematic difference in the number of trials completed in the two conditions, rather than due to a substantial difference in sequence accuracy.

Figure 2 depicts the number of times the original training sequence was repeated after shifting to a new target sequence (trials 51–300). Immediately after the transition to the new target sequence, the previous training sequence was frequently repeated in both conditions. Repetitions of the previous training sequence gradually diminished over the session. Extinction occurred substantially more rapidly when the two sequences differed by the response required in the last position (filled circles) than in the first position (unfilled circles).

The previous figures emphasized complete sequences. Also of interest are the changes that occurred in the individual responses comprising the sequence. The central issue concerns the sensitivity of responding in each position of the sequence to the shift from the training sequence to the new target sequence. The three panels of Fig. 3 show the persistence of responding in the first, middle, and last positions in the sequence, respectively. The vertical line depicts the point in which the transition from training to new target sequence occurred.

The individual points in Fig. 3 were calculated by the following method: for every trial that was completed, a counter was incremented if the observed response was the same as the response that had been required in the training sequence: This counter was then divided by the number of completed trials, producing a per-opportunity measure with a range of 0 to 1. This calculation was carried out for each of the three positions in the sequence, producing the three panels of Fig. 3. Because the procedure produced two types

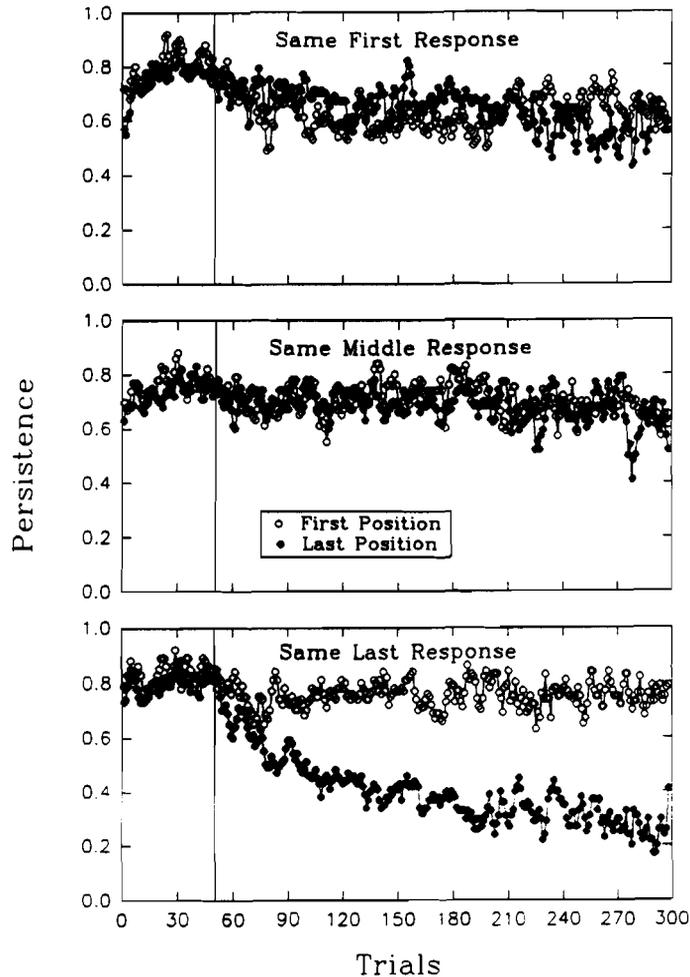


Fig. 3. Persistence in responding in each position in the sequence as a function of the changes in contingency. For each curve a counter was incremented if the observed response was the same as the response that had been required in that position in the training sequence. This counter was then divided by the number of completed trials, producing a per-opportunity measure with a range of 0 to 1. Points represent trials in which the training and target sequences differed by the required response in the first position (unfilled circles) or the last position (filled circles) in the sequence.

of transition from training to new target sequences, this calculation was carried out for the two transition types separately. Therefore, one curve in each panel shows the persistence of responding in a particular position occurring when the training and new target sequences differed by the response required in the first position (identified by open circles). Similarly, the other curve (identified by filled circles) shows the persistence when the difference in sequences was due to the last position.

An example may be useful. The bottom panel depicts changes in the last response position of the sequence. During the first 50 trials (left of the vertical line) subjects were reinforced for completing the training sequence, as in prior sessions. Approximately 80% of

the completed sequences correctly ended with the required response, independent of the upcoming target sequence. After the transition one group of target sequences required a different response in this position (filled circles) and one group of sequences required the same response as in the previous training sequence. Therefore, as the subjects learned to complete the new target sequences correctly, the persistence of responding in the last position should be a function of the new contingency. If the new sequence required a different response than before (filled circles), persistence should decrease as observed – the other lever should be chosen. If the new sequence required the same response as before (open circles), persistence should remain high – the subjects should continue to select that lever at that position in the sequence.

When shifted to new target sequences, the last response in the sequences (bottom panel) quickly shifted to match the required response in that position for both transition types. When the required response was different from the response in the previous training sequence (filled circles), extinction progressed systematically, beginning immediately after the switch to the new sequence. When the required response was the same as in the previous training sequence (open circles), its occurrence remained relatively constant across the session.

In contrast, responding in the first position of the sequence (top panel) appeared relatively unaffected by the shift to a new target sequence. The curve with the open circles in the top panel represents the transition type requiring a different response in the first position than had been required in the training sequence. Sensitivity to the new target sequence would, therefore, require a decrease in this curve, just as the observed decrease in the bottom panel (the curve with filled circles). However, persistence of responding in the first position was relatively constant, regardless of whether the new target sequence required a different response (open circles) or not (filled circles).

Responses in the last position (bottom panel) were substantially more sensitive to changes to the new target sequences than were responses in the first position (top panel).

The response required in the middle position of the training and new target sequences was always the same. If the shift to new target sequences is very demanding, one might expect an increase in variability in responding in the middle of the sequence until the new sequence is learned. Nevertheless, the middle panel shows that persistence in responding in the middle position of the sequence was approximately constant over the session, and did not vary as a function of the differences between training and target sequences. Response selection in this position was not random: a randomly chosen response would have resulted in a value of 0.5 on the ordinate.

Some new target sequences could have been easier to learn than other sequences, independent of the previous training sequence. It is likely, for example, that LLL would be learned faster than LLR. The procedure used in this study does not allow a systematic investigation of all variables that might contribute to the difficulty of learning a response sequence. But it does allow a comparison of sequences containing response repetitions located in different positions of the sequence: at the beginning of a sequence (RRL and LLR) versus the end (RLL and LRR).

In Fig. 4 the two curves represent the errors (per opportunity) committed when responding to these types of target sequences: those in which a response repetition occurs at the beginning of the sequence (RRL and LLR; open circles), and those with one at the end (RLL and LRR; filled circles). Because subjects did not complete as many trials if the target sequence began with a repetition, the number of errors was divided by the number of completed sequences to produce errors per opportunity. Errors persisted substantially

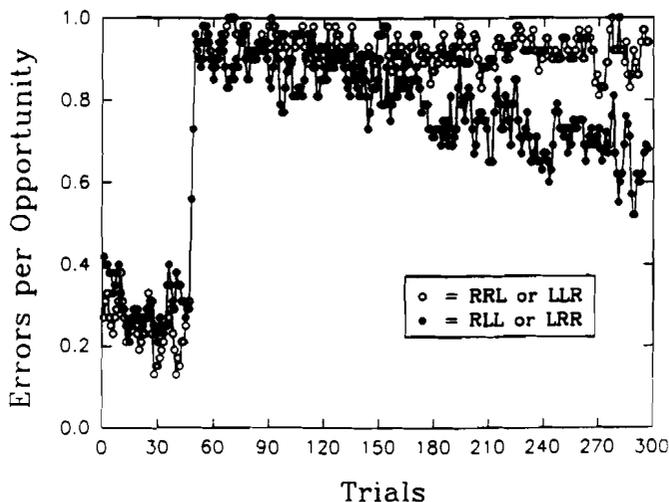


Fig. 4. Number of errors committed when subjects were switched to new target sequences that contained a required repetition of responses at the beginning (RRL or LLR) or end (RLL or LRR) of a sequence. The errors have been divided by the number of trials actually completed on each type of sequence to produce a per-opportunity measure.

longer when the target sequences required repetitions at the beginning of the sequence rather than at the end.

Discussion

Each of these results is in accordance with the contiguity hypothesis. As new sequences were learned, reinforcement acted differentially on the individual responses within the sequence (cf. Fig. 3). While I cannot rule out the existence of reinforcement effects on complete sequences, the behavioral unit hypothesis cannot account for the differential effects on individual responses (cf. Fig. 3) or for the different rates of extinction of the training sequence for the two transition types (cf. Fig. 2). These results do not support an extension of the behavioral unit hypothesis to the acquisition of new response sequences, but do not reject the existence of complex behavioral units.

The contiguity hypothesis assumes that the different responses comprising the sequence would vary in response strength depending upon their contiguity with reinforcement. Therefore, the different responses should be differentially affected by the two types of transition. The last response in the sequence demonstrated the fastest changes after the transition to new target sequences (cf. Fig. 3). Assuming the relation between response strength and temporal contiguity to food (after extended reinforcement of the training sequence) is positively accelerated (Staddon, 1983), this finding supports the notion that changes in response strength due to nonreinforcement are steeper than this relation. Also, extinction of the training sequence was faster when the required response in the last position differed from the new target sequence (e.g. Fig. 2). Response strength and resistance-to-change (Nevin, 1974; Nevin et al., 1981) appeared to be inversely related after the transition.

While several theories of learning and behavior incorporate the contiguity hypothesis (cf. Commons et al., 1991; Davis et al., 1993; Gabriel and Moore, 1990; Staddon, 1983; Staddon and Higa, 1991), the procedure used in this study does not provide a means of distinguishing between them or testing them individually.

In contrast, the behavioral unit hypothesis predicted that after extended exposure to a training sequence, the sequence would function as a single response unit. When the contingency is changed, the sequence as a unit would be extinguished. Extinction of the training sequence should have occurred at the same rate whether the required response in the first or the last position was different in the two sequences. However, Fig. 2 showed a clear difference in the rate of extinction for the two transition types. This difference is evidence that reinforcement was acting on the individual responses within the sequence, rather than on the sequence as a whole.

Fig. 3 provided the most compelling demonstration that the behavioral unit hypothesis does not adequately describe the process of learning new sequences. The different response positions clearly showed differential sensitivity to the new target sequences. Responses in the last position of the sequence were more sensitive to changes in the target sequence than were responses in the other positions. After a shift to a new target sequence, the response in the last position quickly matched the new required response in that position (cf. Fig. 3, bottom panel). Responding in the first position was remarkably insensitive to changes in the contingency (cf. top panel). The response unit on which reinforcement operated was at the level of the individual lever press, rather than on complete response sequences.

These results do not refute the formation of complex behavioral units. Rather, they constrain the application of the behavioral unit hypothesis to particular new learning situations. Behavioral units are descriptive of the way responding becomes organized with extended exposure to a reinforcement schedule. Reinforcement often organizes individual responses into highly stereotyped sequences which may function as single responses (Adams, 1984; Piscreta, 1982; Reed et al., 1991; Schwartz, 1980, 1981, 1986; Shimp, 1979). The main criterion for assessing whether a response sequence functions as a behavioral unit is to reinforce the same putative unit according to an additional contingency. The putative unit should remain unaltered on the new schedule, maintaining its integrity, timing, and accuracy (Reed et al., 1991; Thompson and Lubinski, 1986; Zeiler, 1977, 1986).

The procedure used in this study did not reinforce the same putative unit that may have formed from the training sequence. It would be surprising if any behavioral units that may have formed would not break apart or decompose following the shift to the new target sequence since it always required a different response pattern. The behavioral unit hypothesis may be applicable to new learning situations that continue to reinforce the same unit (cf. multiple and concurrent schedules: Fetterman and Stubbs, 1982, Schneider and Morris, 1992, Schwartz, 1986, Stubbs et al., 1987). It would be of considerable interest to examine the dynamics of learning in these situations as the new contingencies are imposed to see if units maintain their integrity.

To some extent, when behavioral units are formed, they should hinder the learning of new contingencies that target different patterns of behavior. Subjects should be less sensitive to the new contingency because variability, timing, and accuracy of the unit should not change (Reed et al., 1991; Schwartz, 1981, 1986; Thompson and Lubinski, 1986). Using Staddon's (1983) terminology, when exposed to a reinforcement schedule, there is typically a tradeoff between exploiting the schedule (get as much food as possible)

and sampling for new contingencies, better ways of obtaining food, or other food sources. Characteristics that promote exploitation should also promote the formation of behavioral units. The most important of these characteristics is extended exposure to an unchanging schedule (Schwartz, 1981). Characteristics that promote sampling should interfere with the formation of behavioral units. One of these characteristics would be a dynamic environment containing changes in the reinforcement contingency, particularly those that target new response units.

Acknowledgements

Special thanks to Allen Neuringer for lively discussions that led to this study and to Kelli Long for her considerable help running the animals. Armando Machado, Chip Ettinger, John Staddon, Allen Neuringer, and Charles Lyons graciously provided comments on earlier versions of this paper.

References

- Adams, J.A., 1984. Learning of movement sequences. *Psychol. Bull.*, 96: 3–28.
- Arbuckle, J.L. and Lattal, K.A., 1988. Changes in functional response units with briefly delayed reinforcement. *J. Exp. Anal. Behav.*, 49: 249–263.
- Commons, M.L., Grossberg, S. and Staddon, J.E.R., 1991. *Neural Network Models of Conditioning and Action*. Lawrence Erlbaum, Hillsdale, NJ.
- Davis, D.G.S., Staddon, J.E.R., Machado, A. and Palmer, R.G., 1993. The process of recurrent choice. *Psychol. Rev.*, 100: 320–341.
- Fetterman, J.G. and Stubbs, D.A., 1982. Matching, maximizing, and the behavioral unit: Concurrent reinforcement of response sequences. *J. Exp. Anal. Behav.*, 37: 97–114.
- Gabriel, M. and Moore, J., 1990. *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. The MIT Press, Cambridge, MA.
- Grayson, R.J. and Wasserman, E.A., 1979. Conditioning of two-response patterns of key pecking in pigeons. *J. Exp. Anal. Behav.*, 31: 23–29.
- Keller, F.S. and Schoenfeld, W.N., 1950. *Principles of Psychology*. Appleton-Century-Crofts, New York.
- Nevin, J.A., 1974. On the form of the relation between response rates in a multiple schedule. *J. Exp. Anal. Behav.*, 21: 237–248.
- Nevin, J.A., 1979. Reinforcement schedules and response strength. In: M.D. Zeiler and P. Harzem (Editors), *Reinforcement and the Organization of Behavior*, Wiley, New York, pp. 117–158.
- Nevin, J.A., Mandell, C. and Yarensky, P., 1981. Response rate and resistance to change in chained schedules. *J. Exp. Psychol.: Anim. Behav. Process.*, 7: 278–294.
- Papini, M.R. and Silingardi, C.A., 1989. Mechanisms of acquisition of behavioral sequences in rats. *Learn. Motiv.*, 20: 73–86.
- Pisacreta, R., 1982. Some factors that influence the acquisition of complex, stereotyped response sequences in pigeons. *J. Exp. Anal. Behav.*, 37: 359–369.
- Reed, P., Schachtman, T.R. and Hall, G., 1991. Effect of signaled reinforcement on the formation of behavioral units. *J. Exp. Psychol.: Anim. Behav. Process.*, 17: 475–485.
- Schneider, S.M. and Morris, E.K., 1992. Sequences of spaced responses: Behavioral units and the role of contiguity. *J. Exp. Anal. Behav.*, 58: 537–555.
- Schwartz, B., 1980. Development of complex, stereotyped behavior in pigeons. *J. Exp. Anal. Behav.*, 33: 153–166.

- Schwartz, B., 1981. Reinforcement creates behavioral units. *Behav. Anal. Lett.*, 1: 33-41.
- Schwartz, B., 1982. Interval and ratio reinforcement of a complex sequential operant in pigeons. *J. Exp. Anal. Behav.*, 37: 349-357.
- Schwartz, B., 1985. On the organization of stereotyped response sequences. *Anim. Learn. Behav.*, 13: 261-268.
- Schwartz, B., 1986. Allocation of complex, sequential operants on multiple and concurrent schedules of reinforcement. *J. Exp. Anal. Behav.*, 45: 283-295.
- Shimp, C.P., 1979. The local organization of behavior: method and theory. In: M.D. Zeiler and P. Harzem (Editors), *Reinforcement and the Organization of Behavior*, Wiley, New York, pp. 261-298.
- Skinner, B.F., 1938. *The Behavior of Organisms*. Appleton-Century-Crofts, New York.
- Skinner, B.F., 1950. Are theories of learning necessary? *Psychol. Rev.*, 57: 193-216.
- Staddon, J.E.R., 1983. *Adaptive Behavior and Learning*. Cambridge University Press, Cambridge.
- Staddon, J.E.R., 1993. On rate-sensitive habituation. *Adapt. Behav.*, 1: 421-436.
- Staddon, J.E.R. and Higa, J.J., 1991. Temporal learning. *The Psychol. Learn. Motiv.*, 27: 265-294.
- Stubbs, D.A., Fetterman, J.G. and Dreyfus, L.R., 1987. Concurrent reinforcement of response sequences. In: M.L. Commons, J.E. Mazur, J.A. Nevin and H. Rachlin (Editors) *Quantitative Analysis of Behavior Vol V: The Effect of Delay and of Intervening Events on Reinforcement Value*, Lawrence Erlbaum, Hillsdale, NJ, pp. 205-224.
- Tarpy, R.M. and Roberts, J.E., 1985. Effects of signaled reward in instrumental conditioning: enhanced learning on DRL and DRH schedules of reinforcement. *Anim. Learn. Behav.*, 13: 6-12.
- Terrace, H.S., 1984. Simultaneous chaining: The problem it poses for traditional chaining theory. In: M.L. Commons, R.J. Herrnstein, and A.R. Wagner (Editors), *Quantitative Analysis of Behavior, Vol IV: Discrimination Processes*, Ballinger, Cambridge, MA, pp. 99-113.
- Thompson, T. and Lubinski, D., 1986. Units of analysis and kinetic structure of behavioral repertoires. *J. Exp. Anal. Behav.*, 46: 219-242.
- Wasserman, E.A., Deich, J.D. and Cox, K.E., 1984. The learning and memory of response sequences. In: M.L. Commons, R.J. Herrnstein, and A.R. Wagner (Editors), *Quantitative Analysis of Behavior, Vol IV: Discrimination Processes*, Ballinger, Cambridge, MA, pp. 99-113.
- Williams, B.A., Ploog, B.O. and Bell, M.C., 1992, November. Extinction of chain schedule performance. Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis, Missouri.
- Zeiler, M.D., 1977. Schedules of reinforcement. In: W.K. Honig and J.E.R. Staddon (Editors), *Handbook of Operant Behavior*, Prentice Hall, Englewood Cliffs, NJ, pp. 201-232.
- Zeiler, M.D., 1986. Behavioral units: A historical introduction. In: T. Thompson and M.D. Zeiler (Editors), *Analysis and Integration of Behavioral Units*, Erlbaum, Hillsdale, NJ, pp. 1-12.