

*THE DEVELOPMENT OF FUNCTIONAL RESPONSE
UNITS: THE ROLE OF DEMARCATING STIMULI*

ALLISTON K. REID, CYNTHIA Z. CHADWICK,
MYLA DUNHAM, AND ANGELA MILLER

WOFFORD COLLEGE

An experiment with rats examined the roles of demarcating stimuli and differential reinforcement probability on the development of functional response units. It examined the development of units in a probabilistic, free-operant situation in which the presence of demarcating stimuli was manipulated. In all conditions, behavior became organized into two-response sequences framed by changes in local reinforcement probability. A tone demarcating the beginning and end of contingent response sequences facilitated the development of functional response units, as in chunking, but the same units developed slowly in the absence of the tone. Complex functional response units developed even though reinforcement contingencies remained constant. These findings demonstrate that models of operant learning must include a mechanism for changing the response unit as a function of reinforcement history. Markov models may seem to be a natural technique for modeling response sequences because of their ability to predict individual responses as a function of reinforcement history; however, no class of Markov chain can incorporate changing response units in their predictions.

Key words: response acquisition, behavioral unit, response sequences, behavioral variability, response stereotypy, chunking, rats

How do functional response units develop? Consider the process of shaping the lever-press response in rats. The procedure of shaping by successive approximations entails reinforcing patterns of behavior that approximate the desired behavior. The response–reinforcement contingency is successively modified so that closer approximations to the desired behavior are required for reinforcement. Previously reinforced behavior patterns no longer produce reinforcement. The individual movements comprising the lever-press response become integrated into an adaptive functional unit that can be used in a wide variety of reinforcement schedules.

Identification of the functional response

This research was supported by grants from Wofford College to the first author and from the South Carolina Independent Colleges and Universities to the second author. This experiment was conducted in partial fulfillment of the requirements for the BS degree in psychology at Wofford College awarded to the second author. We thank Susan Schneider for her helpful comments on the manuscript. Portions of this paper were presented at the annual meeting of the Association for Behavior Analysis, Orlando, Florida, May 1998. We thank Ellen Burriss, Stephen Gray, and Elizabeth Hubbard for their help in conducting the experiments.

Please address correspondence to Alliston Reid, Department of Psychology, Wofford College, 429 N. Church St., Spartanburg, South Carolina 29303-3663 (E-mail: reidak@wofford.edu).

unit has long been a problem in the experimental analysis of behavior (cf. Catania, 1973). The processes responsible for the transition from simpler to more complex integrated responses have received even less attention. Two complementary theoretical views have been proposed to account for the development of these functional units. Reinforcement theory assumes that reinforcement selects or strengthens the constituent components of the fully integrated response. Once the movement pattern has become integrated into a functional unit, the new unit itself is strengthened by each successive reinforcement. Reinforcement models have rarely addressed the problem of identifying which response is actually being strengthened during this dynamic process (for exceptions, see Shimp, 1979, 1984; Staddon & Zhang, 1991). Instead, “reinforcement is assumed to selectively strengthen the ‘reinforced response’ without any explicit discussion of how the organism knows what that response is” (Staddon & Zhang, p. 280).

A related theoretical approach, originating mostly from computational modeling in artificial intelligence, is to view the task as a problem to be solved by the organism. The assignment-of-credit problem asks how the organism assigns credit to the components of

its behavior stream in such a way that an adaptive pattern of behavior develops. The credit-assignment problem is a necessary part of reinforcement theory. The main mechanism by which reinforcement is said to solve the credit-assignment problem is that of temporal contiguity, which produces greater strengthening of those responses that are more closely followed by reinforcement. Temporal contiguity, however, is clearly not a necessary and sufficient condition for the assignment of credit. Operant learning can occur after substantial delays between a response and reinforcement (e.g., Lattal & Gleeson, 1990), and activities that are contiguous with reinforcement may not be those that are actually strengthened (e.g., Breland & Breland, 1961).

With extended training, adaptive patterns of behavior sometimes become integrated into functional response units (Fetterman & Stubbs, 1982; Schneider & Morris, 1992; Schwartz, 1981, 1982, 1986; Shimp, 1982; Stubbs, Fetterman, & Dreyfus, 1987; Thompson & Zeiler, 1986). In an intriguing examination of functional response units, Fetterman and Stubbs reinforced sequences of two responses in a reinforcement schedule in which matching of response sequences was pitted against matching of individual key pecks. With extended training, matching was obtained with response sequences rather than with individual key pecks. This provides an example of how behavior may become organized into complex functional response units that preclude simpler organization as smaller functional units such as individual key pecks.

How could a model of learning handle this finding? A learning model that assumes that the individual key peck is the response unit that gains strength through reinforcement will have considerable problems unless it includes some mechanism for changing the nature of the response unit. Some recent models of sequence learning (e.g., Machado, 1997) have avoided this problem by simply assuming the existence of functional units consisting of multiple responses, each unit having response strength, without specifying precisely how these units develop. Several studies have demonstrated that complex response units sometimes develop on various reinforcement schedules (Fetterman &

Stubbs, 1982; Schneider & Morris, 1992; Schwartz, 1981, 1982, 1986). However, relatively few studies have investigated how these response units change within individual subjects as a function of their reinforcement history. Therefore, the processes responsible for this shift to more complex units are not well understood. Our emphasis differs from these other studies of behavioral units in that we are interested in the processes responsible for the development of the units during acquisition rather than in their steady-state performance.

Shimp and his colleagues (Shimp, 1978; Shimp, Childers, & Hightower, 1990) have argued cogently that functional response units represent a local behavioral organization reflecting statistical regularities in the organism's dynamic (internal and external) environment. Organisms may organize behavior around changes in their environment. Most experiments demonstrating functional response units have utilized discrete-trials procedures, which provide cues to demarcate the beginning and end of trials (cf. Reed, Schachtman, & Hall, 1991; Reid, 1994; Schwartz, 1982; Silberberg & Williams, 1974). These demarcating stimuli are thought to aid in the temporal organization of behavior just as demarcating stimuli help people to remember telephone numbers (Shimp, 1978).

Nevertheless, Fetterman and Stubbs (1982) and others (e.g., Machado, 1993, 1997) have demonstrated that complex functional response units may develop on free-operant schedules even in the absence of demarcating stimuli. Interestingly, the particular behavioral units obtained in these studies may differ from those produced by a discrete-trials procedure. For example, the behavioral units observed by Fetterman and Stubbs consisted of overlapping pairs of key pecks. That is, a series of left (L) and right (R) pecks such as LLRLRR was interpreted as an overlapping series of behavioral units (LL) (LR) (RL) (LR) (RR). The demarcating stimuli present in most discrete-trials procedures do not produce overlapping behavioral units; rather, each lever press or key peck contributes to only one behavioral unit. One goal of the present research was to compare the behavioral units obtained in a free-operant schedule in which the presence and absence of demarcating stimuli are manipulated to

determine if both conditions generate the same behavioral organization (i.e., the same behavioral units).

The studies above were concerned with steady-state behavior, so they did not provide data concerning the acquisition of the functional response units. We are particularly interested in how behavioral organization develops over time, that is, how these complex units are learned. A central question for the current research is: Do explicit demarcating stimuli influence the speed of development of complex functional response units? Precisely how would the addition of demarcating stimuli influence the development of such units? Finally, once complex response units develop in the presence of demarcating stimuli, would the units maintain their integrity if the demarcating stimuli were subsequently removed?

A final question concerns the role of changing reinforcement probability in the development of functional response units. Consider once again the example of shaping the lever-press response. Successful shaping usually entails modifying the reinforcement contingency in such a way that the behavior stream becomes controlled by differential reinforcement probabilities for its constituent movements. Without changes in the contingency, the lever-press response is unlikely to be learned (but see Lattal & Gleeson, 1990; Sutphin, Byrne, & Poling, 1998). Does the development of functional response units require changes in reinforcement probability for the current behavior pattern, or can they develop even with unchanging reinforcement contingencies?

This experiment was designed to answer the questions concerning the roles of demarcating stimuli and differential reinforcement probability on the development of functional response units. We examined the development of the units in a free-operant schedule in which the presence of demarcating stimuli was manipulated. We exposed two groups of rats to a free-operant schedule that differentially reinforced some two-response sequences with higher probability than others. That is, a computer flipped an imaginary coin following every even response (never after odd responses) and reinforced changeovers from one lever to the other (LR, RL) with greater probability than perseverating on the same le-

ver (LL, RR). We compared the development of functional response units in Group ABA with that of Group BAB, which differed in the order of exposure to demarcating stimuli. Condition A did not provide demarcating stimuli to indicate the beginning or end of the contingent sequences (i.e., when the coin was flipped), whereas Condition B did provide the demarcating stimulus. Our goal was to determine if the subjects would identify the undemarcated sequences (in Condition A) that differentially produced reinforcement with higher probability. Would early exposure to the demarcating stimuli facilitate development of functional response units? Would the behavioral units be the same in both conditions?

Notice that this procedure does not require subjects to detect when the coin is tossed: Strict alternation between levers would produce maximal reinforcement rate and would not require the organization of any complex functional units. We wished to determine if behavior would become organized into functional units consisting of exactly two responses, and if so, would the behavioral organization occur in the presence of an unchanging reinforcement contingency?

METHOD

Subjects

Eight rats (*Rattus norvegicus*) of mixed strain, approximately 4 months old at the beginning of the experiment, were divided into two groups: Group ABA and Group BAB. All subjects were maintained at 80% of their free-feeding body weights by supplemental feeding after experimental sessions with Marlan Tekland Rodent Diet. Subjects were housed individually in home cages, with free access to water, in a room with natural lighting and approximately constant temperature and humidity.

Apparatus

Two identical BRS two-lever rat chambers (30 cm by 23 cm by 24 cm) were used. Each was located inside an isolation chamber with a ventilation fan that masked external noises. A Gerbrands Model D-2 feeder dispensed 45-mg Noyes pellets (Formula A/I). The two response levers were located 8.5 cm above the

floor and were separated by 10.5 cm. Centered between the 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 houselights (GE-1819) were located at the top of the rear wall. A tone generator (Mallory MCP32082) was located directly above the feeder tray, 19 cm above the floor. Individual microcomputers controlled the experiment and recorded every bar press and stimulus change along with their times of occurrence.

Procedure

Pretraining. After their body weights had been reduced to 80% of free-feeding weights, subjects in both groups were trained to press levers by a successive approximations procedure. They were then exposed for one session to a schedule of continuous reinforcement (each lever press produced a food pellet), followed by two sessions of a modified fixed-ratio (FR) 4 schedule. This modified FR schedule delivered reinforcement following the completion of four lever presses distributed on two levers, provided that both levers were pressed at least once. Sessions lasted until 100 pellets were delivered.

Experimental procedure. Subjects in Group ABA were exposed to two conditions, presented in an ABA design. Subjects in Group BAB were exposed to the same two conditions but in the noted order. Subjects remained in each condition until visual observation of their interresponse times (IRTs) showed no increasing or decreasing trend for a minimum of seven sessions. Sessions in all conditions lasted until 100 pellets were delivered. Experimental sessions occurred at approximately the same time each day, 7 days per week.

Condition A. Subjects were exposed to a probabilistic reinforcement schedule that delivered a food pellet with a probability of .6 following every pair of responses containing a changeover from one lever to the other (left-right or right-left, called "heterogeneous" sequences) and with a probability of .2 following two responses on the same lever ("homogeneous" sequences). Therefore, a minimum of two responses was always required for reinforcement. The computer tossed an imaginary biased coin (based on its

pseudorandom number generator) with every even (but not odd) response, with the coin's bias determined by the pair of responses produced. Sequences were nonoverlapping. Thus, each response could contribute to only one two-response sequence (i.e., a moving window of two responses, not one response, was used to define the sequences). Each two-response sequence was undemarcated; that is, no stimulus changes in the subject's environment accompanied the coin tosses other than occasional reinforcement. All lever presses produced a 0.3-s period in which the three panel lamps over the levers blinked off and responses were ineffective. When food was delivered, it occurred at the end of the 0.3-s period.

Condition B. Condition B was exactly like Condition A except in one respect. Every even (but not odd) response produced a 0.3-s tone, independent of reinforcement delivery. The tones coincided with the flip of the imaginary coin. The tones served to demarcate the beginning and end of each two-response sequence. When food was delivered, it occurred at the end of the 0.3-s tone.

RESULTS

Our main interest in this study was the development of functional response units. We assume that lever-press training and early exposure to the ratio schedule produced behavioral units consisting of individual left or right lever presses. The experimental procedure, however, differentially reinforced pairs of lever presses such that heterogeneous sequences were reinforced with a higher probability than were homogeneous sequences. Therefore, a natural comparison is to examine whether the behavior stream became organized at the level of the individual lever press or around pairs of lever presses in the first experimental condition for each group. We want to know whether the tone facilitated this behavioral organization, and whether the same behavioral units were formed in the presence and absence of the explicit demarcating stimulus. We begin by examining the number of trials per session of each of the four possible two-response sequences.

The differential reinforcement probability for heterogeneous versus homogeneous sequences did influence the frequency of each

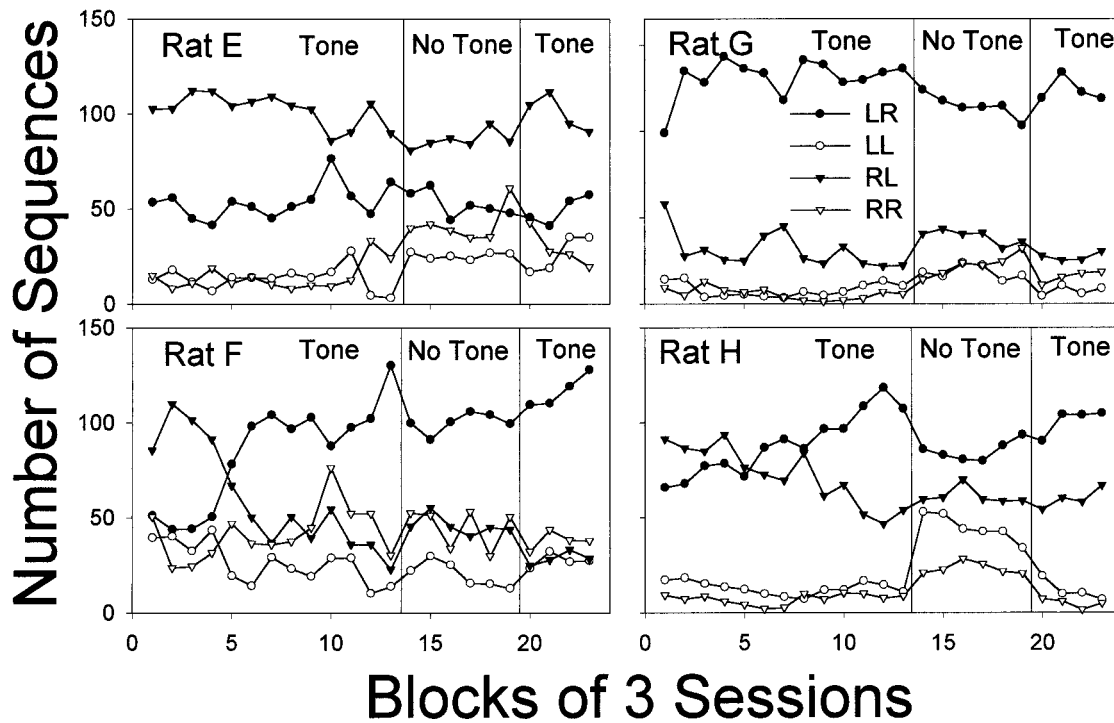


Fig. 1. The number of each response sequence produced in each session in each experimental condition for each subject in Group BAB, averaged across blocks of three sessions. Filled symbols represent heterogeneous sequences. Open symbols represent homogeneous sequences.

response sequence for all subjects in both groups. Figures 1 and 2 depict the number of each of the four types of response sequence generated in each session in each experimental condition for each subject in Group BAB and Group ABA, respectively. All subjects in Group BAB produced heterogeneous sequences more often than homogeneous sequences beginning with the first session. For each subject, one particular heterogeneous sequence became dominant, even though both heterogeneous sequences were reinforced with equal probability. Homogeneous sequences occurred infrequently for all subjects. Group ABA, which began training without a demarcating stimulus, showed the same basic effect (Figure 2), although some subjects required more sessions for stable preference to develop for the heterogeneous sequences. By the end of the first condition, 3 of the 4 subjects (Rat 2 was the exception) produced heterogeneous sequences more often than homogeneous sequences. These subjects also generated one heterogeneous sequence (LR for Rat 1 and

RL for Rats 3 and 4) substantially more often than the other heterogeneous sequence. Homogeneous sequences occurred infrequently for most rats. Rat 2, however, produced RR approximately as often as the two heterogeneous sequences, even though it rarely produced LL.

Following exposure to the first condition, subjects were shifted to the second condition and subsequently returned to the original condition. For Group BAB, the demarcating tone was eliminated during the second condition, whereas for Group ABA the demarcating tone was added. These manipulations of the presence or absence of the demarcating stimulus did not systematically affect the number of each sequence generated in either group.

The most common strategy for determining whether the behavior stream was organized around functional units composed of one or two lever presses is to compare the degree to which each matched its relative reinforcement rate. For example, Fetterman and Stubbs (1982) and Schneider and Morris

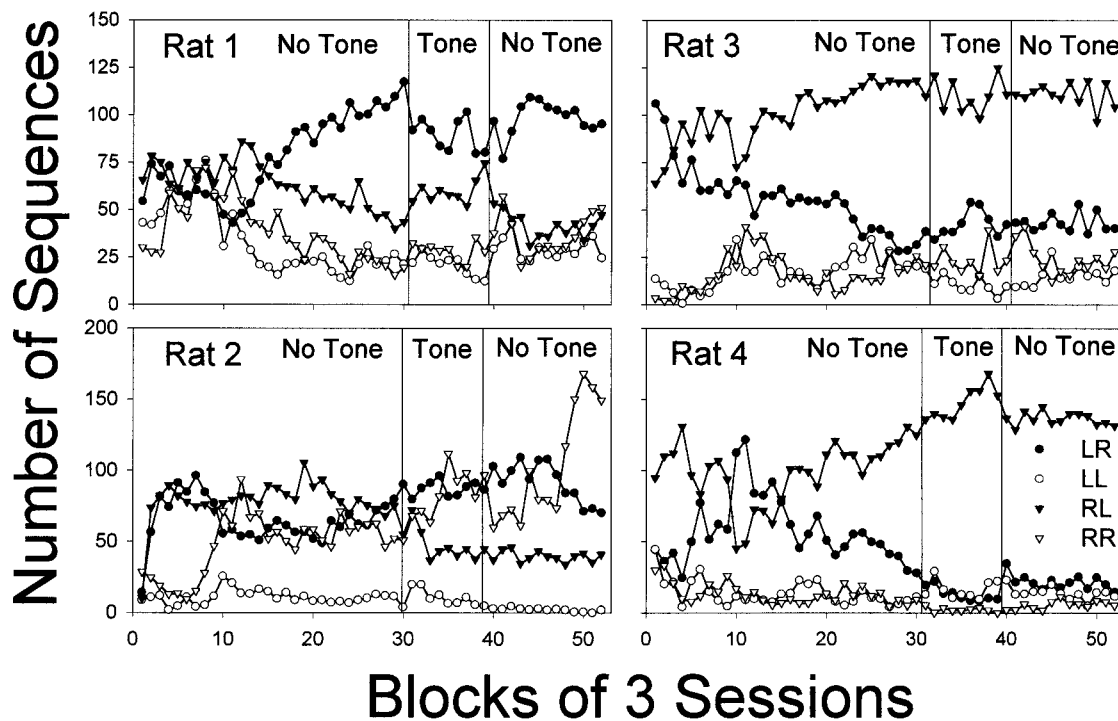


Fig. 2. The number of each response sequence produced in each session in each experimental condition for each subject in Group ABA, averaged across blocks of three sessions. Filled symbols represent heterogeneous sequences. Open symbols represent homogeneous sequences.

(1992) concluded that their experiments demonstrated the existence of behavioral units consisting of two responses, rather than one, because matching of response sequences better accounted for their data than did matching of individual responses. In the current study we can determine the extent to which matching of individual responses occurred, but the observation of matching of response sequences would provide little or no useful information. This is because probabilistic schedules of reinforcement ensure some conformity to matching. Each sequence was reinforced with a fixed probability, so the schedule is similar to a variable-ratio schedule for response sequences. Within the constraints imposed by the computer's random number generator, a fixed frequency of any particular sequence would generate a relatively fixed reinforcement rate. Therefore, matching of response sequences should be generated automatically by the probabilistic schedule, even though matching of individual lever presses is not forced by the schedule.

In evaluating the extent to which matching of *individual* lever presses occurred, consider the case of the individual left lever press. In heterogeneous sequences, the programmed probability of reinforcement for pressing the left and right levers was always equal. In homogeneous sequences, the probability of reinforcement for pairs of left lever presses was always equal to that for pairs of right presses. Therefore, the relative rate of reinforcement *programmed* for left presses was always equal to that for right presses. The *obtained* relative rate of reinforcement will depend upon the frequency of each sequence produced, just as it does in concurrent schedules. For example, if a subject showed a strong preference for RL, then the obtained relative rate of reinforcement for individual left lever presses would be greater than that obtained for right presses because food delivery can occur only at the end of the response sequence. If behavior were organized around the individual lever press, one would expect the proportion of left presses to match the proportion of reinforcement for those presses.

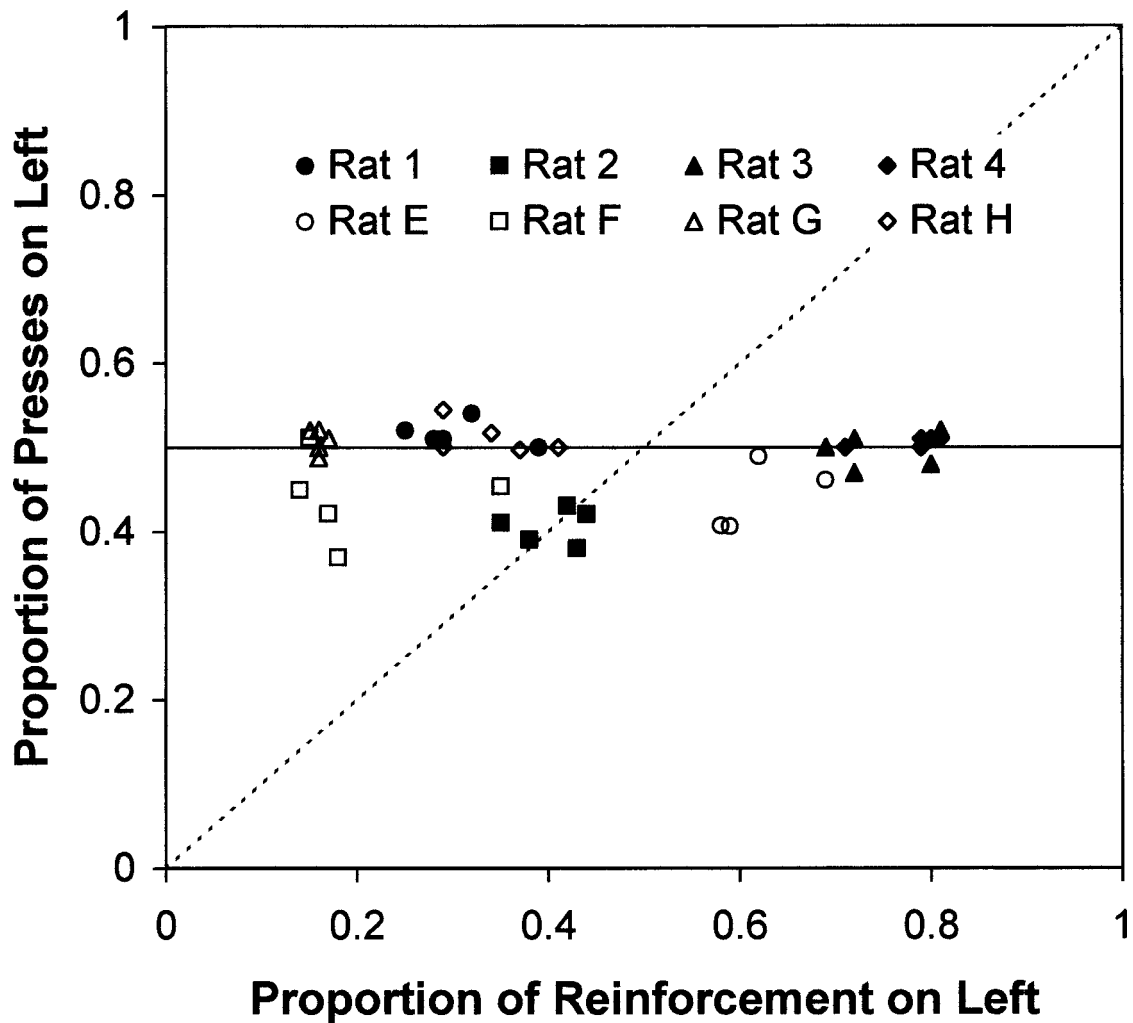


Fig. 3. Comparison of the proportion of left lever presses to the proportion of reinforcement produced by left presses for the last five sessions of the first condition for all subjects in Experiment 1. Each point represents data for 1 subject in one session. The no-tone condition for Group ABA is represented by filled symbols, and the tone condition for Group BAB is represented by open symbols.

This analysis is presented in Figure 3 for the last five sessions of the first condition for all subjects in both groups. The proportion of reinforcement for left presses was calculated by counting the number of reinforcers following left lever presses and dividing by the total number of reinforcers during the session. The dashed line along the diagonal shows the predictions of matching. The horizontal line represents a constant 50% proportion of left presses, regardless of obtained reinforcement for those presses.

For 7 of the 8 subjects, the proportion of left presses did not match the proportion of

reinforcement produced by those left presses. Instead, the proportion of left presses remained at approximately 50%, regardless of the obtained proportion of reinforcement. The exception was for Rat 2, for which the proportion of left presses approximately matched the obtained proportion of reinforcement for left presses. Recall that Rat 2 was also the only subject that generated a large number of homogeneous (RR) sequences. The unique behavioral organization produced by Rat 2 is discussed in more detail below. Because matching of individual responses did not occur in either group, the

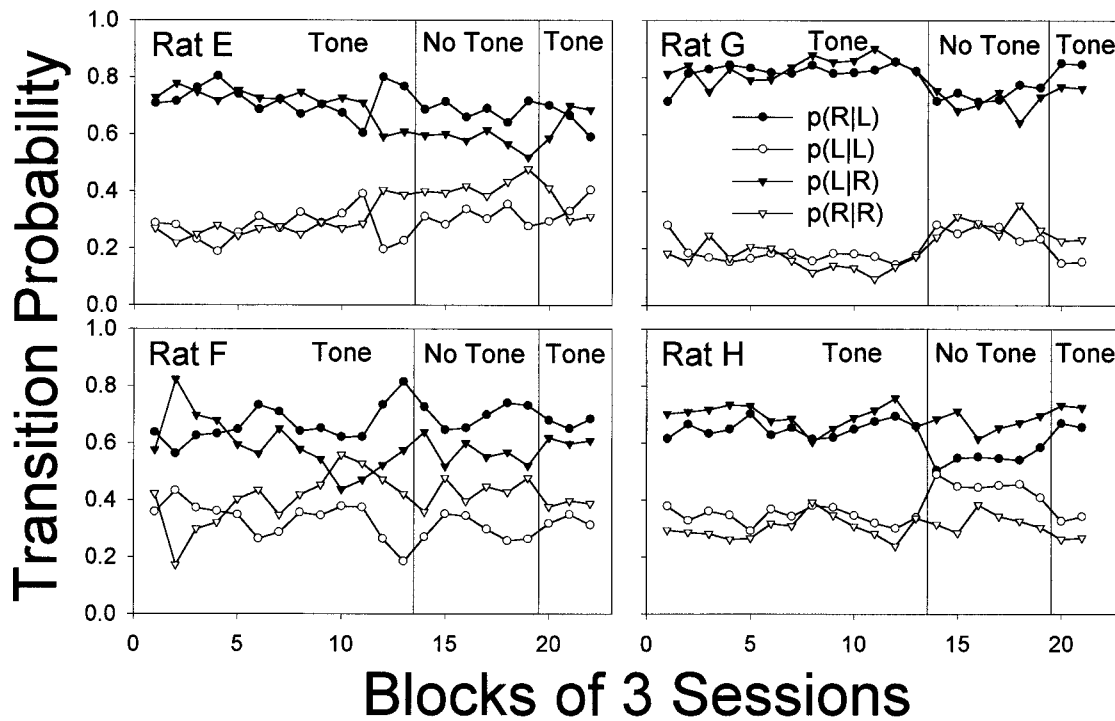


Fig. 4. Transition probabilities for each two-response sequence across sessions for each subject in Group BAB, averaged across blocks of three sessions. Filled symbols represent the conditional probabilities involved in heterogeneous sequences, and open symbols represent those for homogeneous sequences.

individual lever press did not appear to function as the behavioral unit.

Another way of identifying the organization in the behavior stream is to examine the stability of the transition probabilities between responses. That is, one would expect the conditional probability of a left press, given that a left press has just occurred, $p(L|L)$, to be lower than the probability of a right press, because homogeneous sequences have a lower reinforcement probability than heterogeneous sequences. This analysis is different from a simple count of each sequence, because transition probabilities could be high even if the actual frequency of a particular sequence is low. Transition probabilities are useful indicators of the degree to which a response serves as a predictor of the next response in the behavior stream. Responding on a particular lever could have discriminative stimulus (S^D) properties that influence the probability of the next response. Because the presence or absence of the tone as a demarcating stimulus may serve as an additional S^D , the transition probabilities may indicate

how these discriminative stimuli function in their various combinations.

Figures 4 and 5 depict the transition probabilities for each two-response sequence across sessions for each subject in Groups BAB and ABA, respectively. These conditional probabilities were calculated without regard to whether a pattern of responding was reinforced or not. Because we are primarily concerned with the development of functional response units, we first examine the influence of the presence or absence of the demarcating stimulus in the first condition for the two groups.

All rats in Group BAB showed highly stable transition probabilities, beginning with the first session. The values of the four transition probabilities remained nearly constant across all sessions of the tone condition for all 4 subjects. All subjects in this group showed a clear separation between homogeneous and heterogeneous sequences, with no increasing or decreasing trend over the 40 sessions of the tone condition. All subjects in Group ABA produced transition probabilities that be-

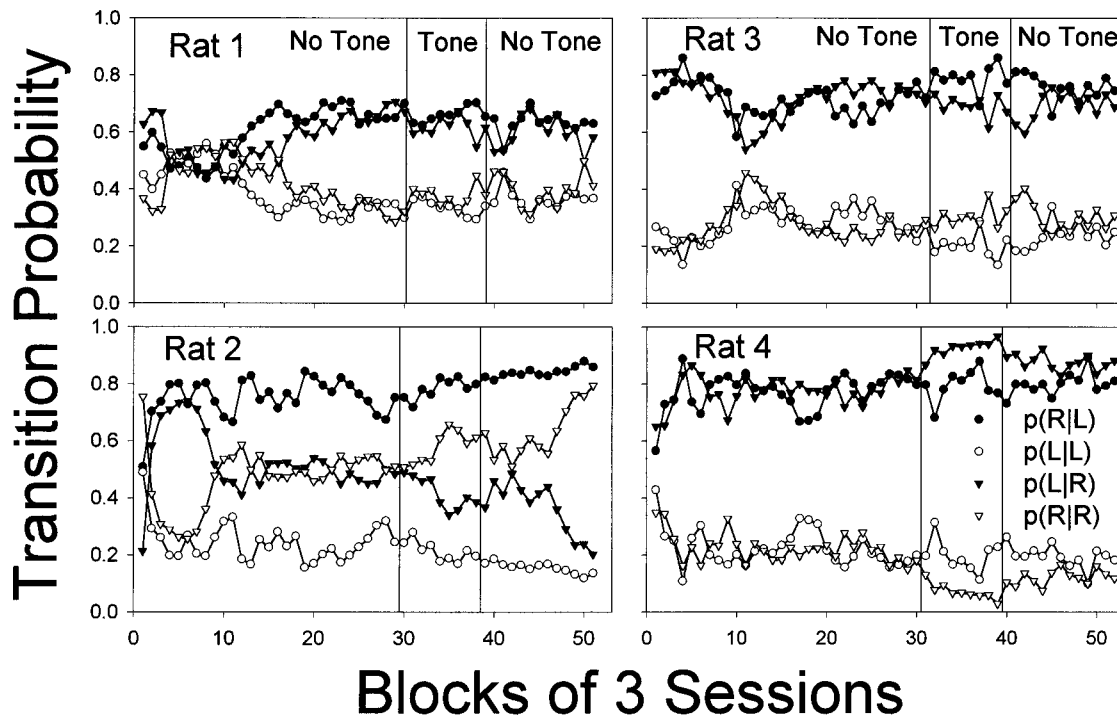


Fig. 5. Transition probabilities for each two-response sequence across sessions for each subject in Group ABA, averaged across blocks of three sessions. Filled symbols represent the conditional probabilities involved in heterogeneous sequences, and open symbols represent those for homogeneous sequences.

came highly stable after 30 to 40 days of exposure to the reinforcement schedule, with those of Rat 4 stabilizing much earlier. Three of the 4 subjects (Rat 2 was the exception again) showed a clear separation between homogeneous and heterogeneous sequences. Once stable, the values of the four transition probabilities remained nearly constant across the sessions of the no-tone condition for these 3 subjects. Rat 2 also had stable transition probabilities, but showed differential control of behavior by earlier presses. A right lever press was not predictive of the next lever selected, $p(L|R) = p(R|R)$, even though a left press was highly predictive of the next press, $p(R|L) > p(L|L)$.

In general, the transition probabilities in the first condition of both groups became stable across sessions, even though only one group had a demarcating stimulus. The tone may have facilitated the development of this stable pattern because the stable pattern usually occurred earlier in subjects with the tone (Group BAB) than in subjects with no tone

(Group ABA). Rat 3, however, may have been an exception to this observation.

The second condition for both groups manipulated the presence or absence of the demarcating stimulus long after the transition probabilities were stable. The effect of adding or removing the tone was small, and it appeared to affect the behavior of only some subjects (Rats G, H, 3, 4, and possibly 2). When an effect was observed in each case in both groups, the presence of the tone aided the predictability of the prior response. That is, the differences in conditional probabilities were slightly greater with the demarcating stimulus than without it.

An examination of the temporal properties of the behavior stream can indicate local behavioral organization and may help to identify functional response units. If functional response units consisting of exactly two nonoverlapping responses developed in this procedure, then it would be reasonable to expect to find longer IRTs associated with responses that produced the flip of the imagi-

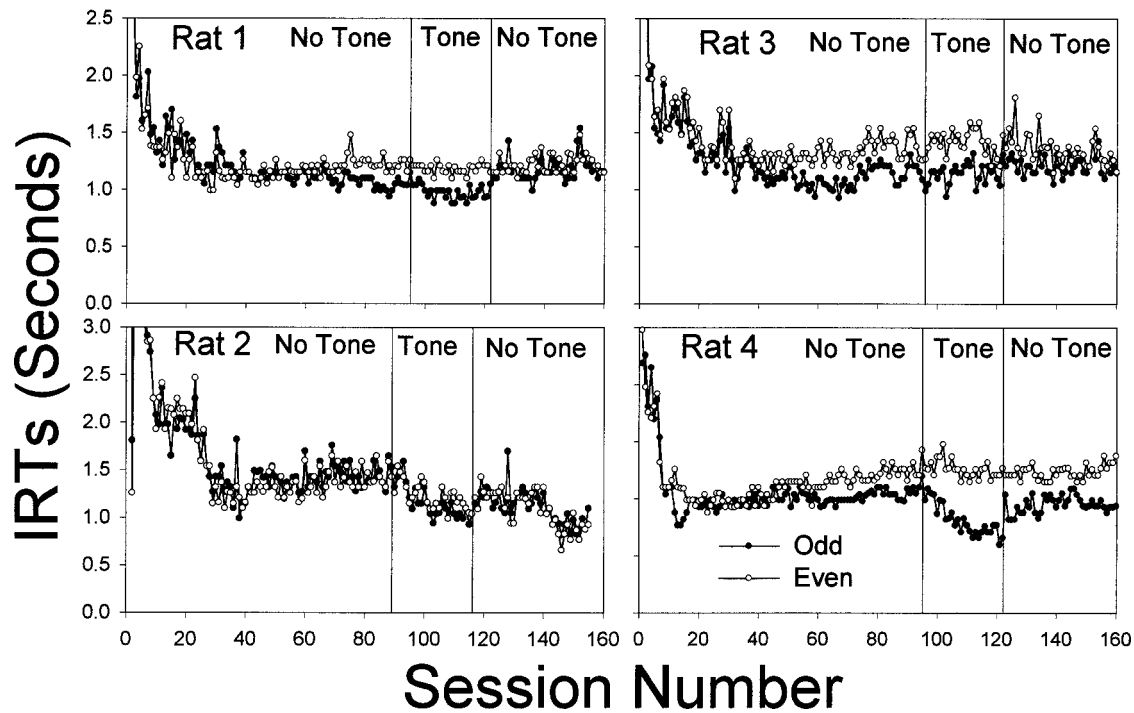


Fig. 6. Median interresponse times (IRTs) following odd (filled circles) and even (open circles) number responses for each subject in Group ABA. Responses producing reinforcement were excluded.

nary coin, that is, following even instead of odd responses. The first lever press following food delivery was never reinforced, but the second press sometimes did produce food. Therefore, one might expect the odd IRTs (time between the first and second responses following an imaginary coin flip) to be shorter in duration than the even IRTs (time between the second and third responses in unreinforced sequences). This finding would be expected, and somewhat trivial, if a demarcating stimulus always occurred following even responses, as in Group BAB. However, a reliable difference between even and odd IRTs in the absence of a demarcating stimulus in Group ABA would be strong evidence for the development of complex behavioral units. We examine this difference in Figure 6, which depicts the median IRTs following odd- and even-numbered responses for each subject in Group ABA. We are primarily concerned with the IRTs that occurred in the first undemarcated condition (no tone). Following 40 to 65 daily sessions, 3 of the 4 subjects (Rat 2 was the exception) produced even IRTs that were significantly longer in du-

ration than were odd IRTs. A paired *t* test on the last 20 sessions of the no-tone condition yielded significant differences for 3 of the 4 subjects: Rat 1, $t(19) = 12.64$, $p < .001$; Rat 2, $t(19) = -1.38$, $p = .91$; Rat 3, $t(19) = 13.77$, $p < .001$; Rat 4, $t(19) = 12.26$, $p < .001$. With extended exposure to the schedule, therefore, a clear difference in the timing of the responses developed, even before the demarcating stimulus was provided in the subsequent tone condition. This difference in IRTs demonstrates that the temporal pattern of responding on this free-operant schedule was organized around the flip of the imaginary coin that occasionally produced food. Note that the differences in these IRTs were not artifacts of food delivery or the additional time required to eat the food pellets, because sequences producing reinforcement were excluded from this analysis.

Rat 2 did not produce different IRTs following even versus odd responses. It is not clear why this subject's behavior was not organized around the coin flip. Nevertheless, its data confirm that the divergence of IRTs

shown by the other subjects was not an artifact of the procedure used to calculate IRTs.

In the second condition for this group, the brief (0.3-s) tone was presented as a demarcating stimulus during every coin flip (i.e., the tone condition). Differences in even and odd IRTs were maintained or increased. Rat 2 continued to show no difference in IRTs, even with the demarcating stimulus. The differences disappeared for 2 of 3 rats when the tone was discontinued.

To know precisely how the tone affected IRTs, a finer grained analysis is necessary. The analysis of even versus odd IRTs (Figure 6) included many responses that were temporally distant from reinforcement. Because reinforcement was probabilistic, rats often produced series of eight or more responses before food was delivered. It is likely that discrimination of differential reinforcement probability (coin flips) becomes poorer with increases in the length of the series of responses. Therefore, we calculated the median IRTs following the first three unreinforced responses following food deliveries. These IRTs are depicted in the three curves of Figures 7 and 8 for Groups ABA and BAB, respectively.

Figure 7 shows that by the end of the no-tone condition, 3 of the 4 subjects in Group ABA (Rat 2 was the exception) showed a clear difference in the IRTs following the first and second responses following reinforcement. This difference is similar to the difference observed in Figure 6 between even and odd IRTs, except that the current analysis was limited to only those responses immediately following reinforcement. Subjects showed little or no difference between the second and third IRTs following reinforcement in this condition, even though reinforcement was possible only after the second response. When the tone was introduced as a demarcating stimulus in the second condition, the IRTs following the third response decreased to or below the durations obtained for the first IRTs (Rats 1, 2, and 3) or reached values between those obtained following the first and second responses (Rat 4). When the demarcating stimulus was removed in the third condition, the IRTs following the third response regained their previous levels for all subjects in the group. The tone served as an explicit discriminative stimulus that helped to distinguish the differences in local reinforce-

ment probability between the second and third responses following food delivery. When the tone was present, all subjects (even Rat 2) responded quickly following the third response, but when the tone was absent, subjects treated the second and third responses similarly and produced longer IRTs.

Figure 8 shows that early exposure to the demarcating stimulus for Group BAB resulted in patterns of IRTs in the first condition that were different in two main respects from those observed for Group ABA (Figure 7). First, with the tone, subjects in Group BAB began pausing after the second response very early in training, whereas subjects in Group ABA (no tone) began pausing only after 30 to 65 sessions. (Note that the first 40 sessions are not depicted in Figure 7 to show more clearly the differences in IRTs at the end of the condition.) The tone facilitated the development of complex behavioral units.

Second, Figure 8 shows that IRTs following the second response were longer than those following *either* the first or third response. The IRTs for the first and third responses were nearly identical in the tone conditions. The no-tone condition for Group ABA (Figure 7) did not produce differences in first and third IRTs. However, once Group ABA was exposed to the demarcating stimulus after considerable training without one, the tone condition produced decreases in the IRTs following the third response. Informal observations indicated that the rats stopped looking for food after lever presses that did not produce a tone. In each of the tone conditions for both groups, the tone became an external S^D to examine the food tray, and the absence of the tone became an S^A .

The observation of reliably different IRTs in Group ABA (Figure 7), even in the absence of the tone, demonstrates some degree of control of behavior by number of responses produced since food delivery. The effects of this S^D appeared to be supplemented by the tone, which was a more explicit S^D (during tone conditions), to increase the organization of behavior around the flips of the imaginary biased coin.

DISCUSSION

The behavior of most subjects in this experiment became organized around the flips

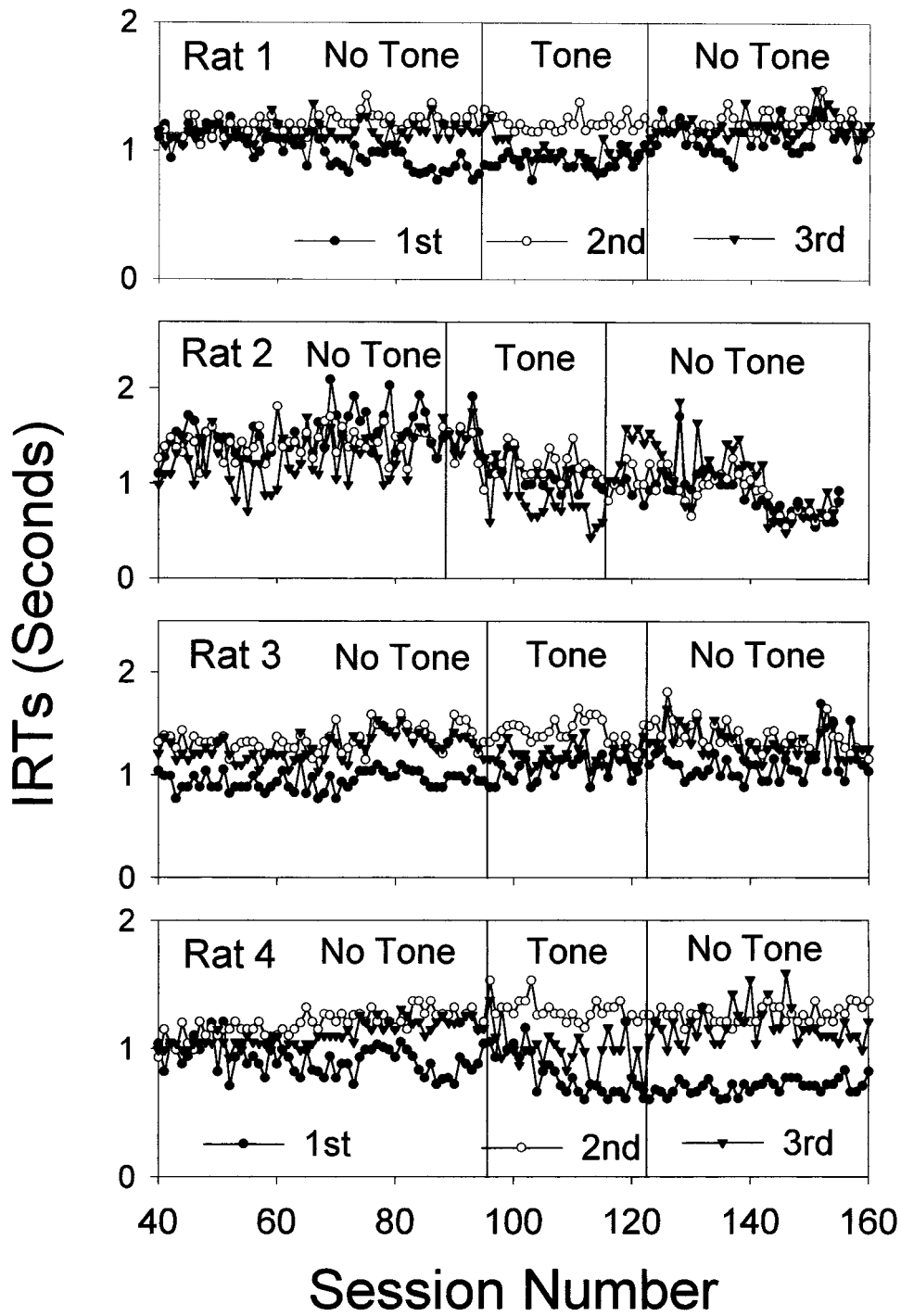


Fig. 7. Median IRTs following the first, second, and third unreinforced responses following each food delivery for subjects in Group ABA. Reinforced responses were excluded. The first 40 sessions are not depicted to show more clearly the differences in IRTs at the end of the no-tone condition.

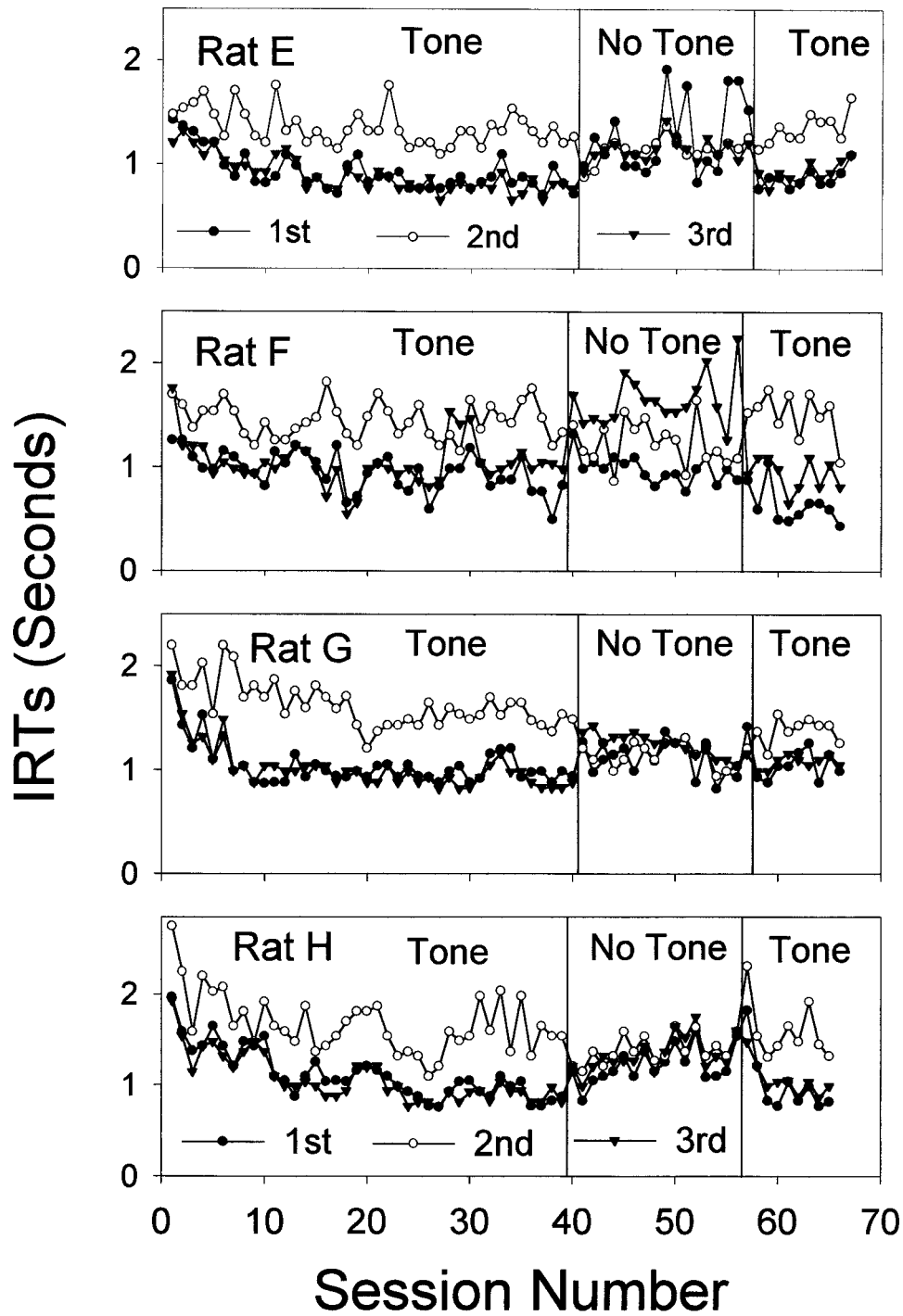


Fig. 8. Median IRTs following the first, second, and third unreinforced responses following each food delivery for subjects in Group BAB. Reinforced responses were excluded.

of the imaginary biased coin. Imaginary coin flips represented changes in local reinforcement probability because reinforcement was possible after even lever presses but never after odd presses. Recall that when behavior became organized in the first condition (no tone) for Group ABA, coin flips were not associated with any changes in the subjects' environment other than occasional reinforcement. This finding demonstrates that extended exposure to the differential reinforcement probability was sufficient to organize behavior into sequences of two responses and also that changes in reinforcement contingency, such as those that occur during shaping procedures, were not necessary for behavioral organization to occur.

The evidence for behavioral organization as functional response units came from five converging measures: (a) The reinforcement contingency controlled the frequency of each two-response sequence (see Figures 1 and 2). (b) The individual lever press did not appear to act as a functional response unit: Subjects that showed the behavioral organization did not produce a matching relation between the relative proportion of individual lever presses and the obtained reinforcement generated by these presses (see Figure 3). Rather, the proportion of left presses was independent of the frequency of reinforcement they produced. (c) Transition probabilities between all two-response sequences were highly stable (see Figures 4 and 5). (d) All subjects that showed the behavioral organization also produced shorter IRTs after odd lever presses than after even presses (see Figure 6). (e) The finer grained analysis of the temporal pattern of behavior (comparing IRTs after the first and second lever presses following every food delivery) showed shorter IRTs after the first press than after the second press (see Figures 7 and 8). This difference generally increased when a demarcating stimulus (the brief tone) was added to each coin flip.

Complex functional response units developed in both conditions. The behavioral units obtained with the demarcating tone were the same as those obtained without explicit demarcating stimuli. Behavioral units consisted of two separate lever presses, framed by flips of the imaginary coin, that were associated with different reinforcement probabilities. Behavioral units did not consist

of overlapping pairs of responses, such as those observed in the free-operant studies of Fetterman and Stubbs (1982) and Machado (1997). In these earlier studies, reinforcement probability did not change in a cyclic pattern around sequences of behavior of fixed size, such as pairs of responses. Thus, differential reinforcement probability may not have served as effectively as a discriminative stimulus that could demarcate non-overlapping response patterns. Although this argument may help to explain why nonoverlapping behavioral units were observed in this experiment, it does not explain why behavioral units would overlap in other studies.

The fact that behavior became organized around changes in the environment resembles chunking (Miller, 1956). Although generally applied to humans, chunking frequently has been proposed as a perceptual grouping process contributing to behavioral organization in nonhumans (Fountain & Annau, 1984; Fountain, Henne, & Hulse, 1984; Terrace, 1987, 1991; Terrace & Chen, 1991a, 1991b). Terrace (2001) points out that the relation of chunking to serial behavior patterns may not be as straightforward as it might seem. For example, chunking is generally acknowledged to facilitate short-term memory, whereas experiments with rats and pigeons measure behavior over months, thus relying heavily on long-term memory processes. When the temporal organization of behavior sequences is used to define chunks, as in this study, Terrace (2001) argues that one should distinguish between these "output chunks" and the "input chunks" of the type demonstrated by Miller (1956). Input chunking refers to the organization of newly encoded information, whereas output chunking refers to the organization of familiar information retrieved from long-term memory.

Whatever memory processes are involved, it is clear that extended exposure to an unchanging reinforcement contingency produced changes in the behavioral unit from individual lever presses to pairs of lever presses. Behavior became organized around two types of environmental changes: (a) the demarcating tone and (b) the changing reinforcement probabilities following even and odd lever presses. The demarcating tone was an explicit signal associated with increased reinforcement probability. The tone facilitated

the development of complex behavioral units, consistent with the framework of Shimp and his colleagues (Shimp, 1978; Shimp et al., 1990). They occurred earlier in training, and IRTs were longer in tone conditions than in no-tone conditions.

Modeling the Development of Functional Response Units

How could a model of learning handle these findings? It is informative to compare two classes of models. The first class derives from a theoretical approach proposed in a series of articles by Shimp and his colleagues (Shimp 1978, 1979, 1984, 1992; Shimp et al., 1990; Shimp & Friedrich, 1993).

Shimp's (1978, 1992) associative learner is a dynamic stochastic model designed specifically to show how reinforcement contingencies can create local organization in behavior. Several versions of this model have been applied to temporal psychophysics, IRT distributions on simple reinforcement schedules, molar versus molecular performance on concurrent schedules, and both discrete-trials and free-operant schedules. Recent versions of this cognitive processing model have been applied to a molecular analysis of human operant behavior (Shimp et al., 1990) and to spatial attention (Shimp & Friedrich, 1993). To our knowledge, it has not been applied directly to the learning of complex response sequences distributed on two or more levers or keys like the sequences learned in this study. Nevertheless, the theoretical framework shows much potential to explain how complex functional response units develop, but existing versions may require modification for this particular application. We will not attempt to extend this model at this time.

The other approach is to examine sequences of responses as Markov chains. Markov chains have been frequently used to describe response sequences and as stochastic models of operant learning. Markov chains may be used as descriptive statistics, describing properties of transitions between activities. They are also widely used as a parsimonious foundation for stochastic models of serial behavior (e.g., Machado, 1997). The present data have implications beyond those of any particular Markov model of serial behavior (e.g., Machado, 1997); therefore, we discuss the overall approach of using Markov chains to model

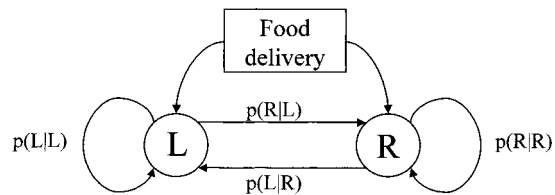


Fig. 9. State transition diagram for a two-state Markov chain. Food delivery may serve as a discriminative stimulus affecting the unconditional probability of a left or right lever press.

serial learning and discuss particular models only as examples.

Markov chains may take many forms. For example, they may include states with one response or several responses; they may be first or higher order; discrete or continuous; transition probabilities may be stationary or non-stationary; and time may (semi-Markov chain) or may not (Markov chain) be explicitly included. Because of the considerable explanatory power of Markov chains and their popularity as models of serial behavior, an examination of their feasibility with the current data should be useful.

Consider the case of the simplest first-order Markov chain with two independent states, representing left and right lever presses. The state transition diagram for this model is depicted in Figure 9. The primary goal is to reduce the uncertainty in predicting the occurrence of the next response. This "next response" is defined as a simple lever press in Figure 9. (Note, however, that as more complex functional response units are formed, this predicted response must become more complex.) Markov chains attempt to reduce uncertainty by including knowledge of prior events. Therefore, all Markov chains are based on a comparison of two probabilities: Assuming the sequence RRL was observed, the unconditional probability of the left lever press $p(L)$ is compared to the conditional probability of a left press given that a right press had just occurred $p(L|R)$. The analysis calculates the improvement in prediction gained by including the immediately prior response as a predictor of the next response. If one includes responses that occur before this prior response, one can generate Markov chains of higher order. These chains also are based on comparisons of probabilities. For example, a second-order

Markov chain would compare $p(L|R)$ with $p(L|RR)$ to determine if prediction of the left press is improved by including a history of two responses rather than only one.

Markov chains are useful tools for analyzing sequences of behavior because they provide an explicit way of separating two mechanisms important for sequence learning: (a) the contribution of the history of responding to the prediction of the next response and (b) the characteristics of the next response predicted to occur. History effects may be of two types. The effects of recent history of responding may be treated as discriminative stimuli that influence the probability of the next response. Markov chains may also include other possible discriminative stimulus effects. For example, Figure 9 includes food delivery as a discriminative stimulus. Each food delivery may alter the probability of left or right responses. Markov chains of first, second, or higher order differ in the amount of history included in the prediction of the next response. However, the characteristics of the individual predicted response do *not* change. If the independent states of the chain are defined as individual left and right lever presses (as in Figure 9), then Markov chains of any order can predict *only* these individual presses. Using the Chapman-Kolmogorov equations (cf. Kemeny & Snell, 1976), one can calculate the probabilities of particular sequences composed of these individual responses. However, the complexity of the individual response (the functional response unit) will not change. This point is most important for our purposes. Our demonstration that extended exposure to the reinforcement schedule produced changes in the functional response unit implies that the response to be predicted by the model must change.

At first glance, a first-order Markov chain with stationary transition probabilities appears to fit our data well; however, it fits if and only if we limit our predictions to the next left or right lever press, ignoring the development of more complex units and their temporal patterning. The impressive stability of the transition probabilities depicted in Figure 4 demonstrates that the transition probabilities were stationary. "Knowledge" of the last response was an accurate predictor of the next response in all conditions of both experiments. However, Figure 2 showed that

most subjects generated one two-response sequence substantially more often than the other three sequences. For example, Rat 3 produced RL several times more often than the other sequences. One RL sequence usually followed another RL sequence. Although this chain could be extended to one of second order, the increase in predictive power would not be significant because $p(R|L) \approx p(R|RL)$. If we only wanted to predict which lever would be pressed next on this reinforcement schedule, a first-order Markov chain would suffice.

A first-order Markov chain cannot account for the behavioral organization observed, however. Recall that most subjects produced longer IRTs after each even (but not odd) lever press (see Figures 6 through 8). How could a Markov chain account for such pauses after each second response? There are three options (Kemeny & Snell, 1976). One option is to replace the Markov chains with semi-Markov chains. Markov chains do not include time as an explicit variable, but semi-Markov chains do. However, Gottman and Roy (1990) point out that one should use semi-Markov models only if one believes that transitions to another state are a function of the time spent in an antecedent state. The present data indicate that behavior became organized around flips of the imaginary coin; that is, events were the organizing features, not elapsed time. Therefore, we will not consider the option of using semi-Markov chains further.

A second option is to realize that the subjects had to be doing something during the longer IRTs. Informal observations indicated that they were examining the food tray. Therefore, one could incorporate one more behavioral state into the chain, forming a Markov chain with three independent states: left press (L), right press (R), and check for food (C). Can a Markov chain with these three states account for the longer IRTs that occurred after even presses? Consider the frequencies of the sequences produced by Rat 1 depicted in Figure 1. This subject generated heterogeneous sequences (LR, RL) more often than homogeneous sequences. Longer IRTs usually followed both types of heterogeneous sequences. That is, this subject generated the chains LRC and RLC. These data do not conform to a first-order Markov chain

with stationary transition probabilities. In the first chain L should predict R, but L should predict C in the second chain (occurring in the same session). Similarly, R should predict C in one instance and L in the other. One could, in principle, make post hoc assumptions to allow these data to be characterized by a second-order Markov chain with three states. However, it would be equivalent to assuming the existence of an undetected, hypothetical modulator that controlled predictions made by each response. The current data provided no evidence for such a modulator.

A final option to account for the IRT data with a Markov chain is to assume that the transition probabilities are not stationary. That is, the transition probability from one particular response to another is assumed to be a function of N , the number of responses produced; that is, $p(R|L) = f(N)$. For example Machado (1997) proposed a first-order Markov chain with nonstationary transition probabilities to account for the variability of sequences of eight responses obtained in pigeons when changeovers between keys were reinforced probabilistically. He used a discrete-trials procedure. In his application of the Markov chain, the counter for the number of responses (N) was reset at the beginning of each trial. In free-operant procedures, such as in the present experiment, it is more difficult to identify what the value of N might be at any point in the session. Nevertheless, one could assume that food delivery always resets N , which would then increment with each lever press until the next food delivery resets it. To account for the present IRT data, $f(N)$ would have to represent a threshold change in probability occurring precisely between the first and second responses after food delivery. This is because $p(R|L)$ must be very high following the first response after food delivery in the sequence LRC [$p(R|L) \gg p(C|L)$, so that L predicts R, rather than C], but very low after the second response in the sequence RLC occurring in the same session [$p(R|L) \ll p(C|L)$, to reverse the prediction]. This post hoc hypothesis is unlikely for two reasons: (a) The parameter values that provided best fits for Machado's pigeons produced much slower changes in $f(N)$ than those required for the current data; and (b) the impressive stability of the transi-

tion probabilities depicted in Figure 4 do not indicate that transition probabilities between the first and second responses of each sequence were changing so dramatically.

Because these three options are not feasible, we are left to conclude that a first-order Markov chain cannot account for the behavioral organization observed in the current study, particularly the IRT data. Extending to second-order chains would not provide much additional predictive value. All of these Markov chains have the same weakness: They predict the next individual response rather than the next sequence of responses. The current results showed that the functional response unit changed with extended training. Markov chains include no principled way of changing the components of each state from a single lever press to a particular two-response sequence. To account for the IRT data, one could easily propose Markov chains with four states, each composed of two lever presses. This chain, however, would no longer be suitable for the early training sessions. It would be equivalent to assuming the existence of the complex functional response units whose development we wish to explain. Changes in functional response units produce critical challenges to theories of sequence learning, particularly those based on Markov chains. Theories of learning must include some principled way of predicting changes in the functional response units.

REFERENCES

- Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, *16*, 681-684.
- Catania, A. C. (1973). The concept of the operant in the analysis of behavior. *Behaviorism*, *1*(2), 103-116.
- Fetterman, J. G., & Stubbs, D. A. (1982). Matching, maximizing, and the behavioral unit: Concurrent reinforcement of response sequences. *Journal of the Experimental Analysis of Behavior*, *37*, 97-114.
- Fountain, S. B., & Annau, Z. (1984). Chunking, sorting, and rule-learning from serial patterns of brain-stimulation reward by rats. *Animal Learning & Behavior*, *12*, 265-274.
- Fountain, S. B., Henne, D. R., & Hulse, S. H. (1984). Phrasing cues and hierarchical organization in serial learning in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 30-39.
- Gottman, J. M., & Roy, A. K. (1990). *Sequential analysis: A guide for behavioral researchers*. New York: Cambridge University Press.
- Kemeny, J. G., & Snell, J. L. (1976). *Finite Markov chains*. New York: Springer-Verlag.
- Lattal, K. A., & Gleeson, S. (1990). Response acquisition

- with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, 16, 27–39.
- Machado, A. (1993). Learning variable and stereotypical sequences of responses: Some data and a new model. *Behavioural Processes*, 30, 103–130.
- Machado, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of switching between two keys. *Journal of the Experimental Analysis of Behavior*, 68, 1–25.
- Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–96.
- Reed, P., Schachtman, T. R., & Hall, G. (1991). Effect of signaled reinforcement on the formation of behavioral units. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 475–485.
- Reid, A. K. (1994). Learning new response sequences. *Behavioural Processes*, 32, 147–162.
- Schneider, S. M., & Morris, E. K. (1992). Sequences of spaced responses: Behavioral units and the role of contiguity. *Journal of the Experimental Analysis of Behavior*, 58, 537–555.
- Schwartz, B. (1981). Reinforcement creates behavioral units. *Behaviour Analysis Letters*, 1, 33–41.
- Schwartz, B. (1982). Interval and ratio reinforcement of a complex, sequential operant in pigeons. *Journal of the Experimental Analysis of Behavior*, 37, 349–357.
- Schwartz, B. (1986). Allocation of complex, sequential operants on multiple and concurrent schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 45, 283–295.
- Shimp, C. P. (1978). Memory, temporal discrimination, and learned structure in behavior. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12, pp. 39–76). New York: Academic Press.
- Shimp, C. P. (1979). The local organization of behaviour: Method and theory. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behavior: Vol. 1. Reinforcement and the organization of behavior* (pp. 261–298). New York: Wiley.
- Shimp, C. P. (1982). Choice and behavioral patterning. *Journal of the Experimental Analysis of Behavior*, 37, 157–169.
- Shimp, C. P. (1984). Timing, learning, and forgetting. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Vol. 423, pp. 346–360). New York: New York Academy of Sciences.
- Shimp, C. P. (1992). Computational behavior dynamics: An alternative description of Nevin (1969). *Journal of the Experimental Analysis of Behavior*, 57, 289–299.
- Shimp, C. P., Childers, L. J., & Hightower, F. A. (1990). Local patterns in human operant behavior and a behaving model to interrelate animal and human performances. *Journal of Experimental Psychology: Animal Behavior Processes*, 16(2), 200–212.
- Shimp, C. P., & Friedrich, F. J. (1993). Behavioral and computational models of spatial attention. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 26–37.
- Silberberg, A., & Williams, D. R. (1974). Choice behavior on discrete trials: A demonstration of the occurrence of a response strategy. *Journal of the Experimental Analysis of Behavior*, 21, 315–322.
- Staddon, J. E. R., & Zhang, Y. (1991). On the assignment-of-credit problem in operant learning. In M. L. Commons, S. Grossberg, & J. E. R. Staddon (Eds.), *Quantitative analyses of behavior: Neural network models of conditioning and action* (pp. 279–293). Hillsdale, NJ: Erlbaum.
- Stubbs, D. A., Fetterman, J. G., & Dreyfus, L. R. (1987). Concurrent reinforcement of response sequences. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 205–224). Hillsdale, NJ: Erlbaum.
- Sutphin, G., Byrne, T., & Poling, A. (1998). Response acquisition with delayed reinforcement: A comparison of two-lever procedures. *Journal of the Experimental Analysis of Behavior*, 69, 17–28.
- Terrace, H. S. (1987). Chunking by a pigeon in a serial learning task. *Nature*, 325, 149–151.
- Terrace, H. S. (1991). Chunking during serial learning by a pigeon: I. Basic evidence. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 81–93.
- Terrace, H. (2001). Chunking and serially organized behavior in pigeons, monkeys, and humans. In R. G. Cook (Ed.), *Avian visual cognition*. Retrieved from www.pigeon.psy.tufts.edu/avc/terrace/
- Terrace, H. S., & Chen, S. (1991a). Chunking during serial learning by a pigeon: II. Integrity of a chunk on a new list. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 94–106.
- Terrace, H. S., & Chen, S. (1991b). Chunking during serial learning by a pigeon: III. What are the necessary conditions for establishing a chunk? *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 107–118.
- Thompson, T., & Zeiler, M. D. (1986). *Analysis and integration of behavioral units*. Hillsdale, NJ: Erlbaum.

Received August 27, 1999
Final acceptance July 13, 2001