Serial Learning by Rhesus Monkeys: I. Acquisition and Retention of Multiple Four-Item Lists

Karyl B. Swartz
Lehman College, City University of New York

Shaofu Chen and H. S. Terrace
Columbia University

Two rhesus monkeys were trained to learn eight 4-item lists, each composed of 4 different photographs. Lists were trained in successive phases: A, A→B, A→B→C, and A→B→C→D. After List 4, retention, as measured by the method of savings, was, on average, 66% (range: 44–84%). Indeed, all 4 lists could be recalled reliably during a single session with neither a decrement in accuracy nor an increase in the latency of responding to each item. Response latencies on a subset test employing all possible 2- and 3-item subsets of each 4-item list support the hypothesis that monkeys form linear representations of a list. Latencies to Item 1 of a subset varied directly with the position of that item in the original list. On List 1, latencies to Item 2 varied directly with the number of intervening items between Item 1 and Item 2 in the original list. During the acquisition of Lists 5–8, both Ss mastered the A→B and A→B→C phases of training in the minimum number of trials possible.

Recent studies of serial learning by monkeys and pigeons have provided ample evidence that these nonverbal organisms can learn to produce arbitrary lists composed of arbitrary stimuli (D'Amato & Colombo, 1988; Straub & Terrace, 1981; Terrace, 1987). While learning to produce a list, pigeons and monkeys also acquire knowledge of nonadjacent items (D'Amato & Colombo, 1988; Terrace, 1987) and of the ordinal position of individual items (D'Amato & Colombo, 1989; Terrace 1986a). Chunking of an arbitrary list was demonstrated in studies in which pigeons were trained to produce lists consisting of clustered subsets of similar items such as colors and achromatic geometric forms (Terrace, 1987, 1991; Terrace & Chen, 1991a, 1991b).

Taken together, these findings pose some intriguing questions about serial learning. How, for example, does a nonverbal organism represent an arbitrary list of arbitrary items? The answer to that question appears, to some extent, to depend on the organism. The available evidence suggests that monkeys form linear representations of their lists and that pigeons do not (D'Amato & Colombo, 1988; Terrace, 1991). One should also keep in mind one fundamental aspect about

an animal's representation of an arbitrary list. Neither the list items, nor their positions, are encoded linguistically. List learning by nonverbal organisms perforce draws upon cognitive processes for serially organizing arbitrary items that antedate language phylogenetically.

When attempting to answer questions concerning the ability of nonverbal organisms to perform serial tasks, it is all too easy to overlook the fact that linguistic competence is by no means the only factor that distinguishes animal and human subjects in experiments on list learning. Each study of list learning in animals has used subjects who, at best, had learned but a single list prior to the experiment in question (Terrace, 1986a) or who were experimentally naive with respect to list learning (e.g., Terrace, 1986b, 1987, 1991; Terrace & Chen, 1991a, 1991b). By contrast, ever since Ebbinghaus (1885/1964), investigators of human memory have used subjects who are not only able to encode linguistically the lists they memorize but who have also had the benefit of list-learning strategies that they have developed in the course of mastering innumerable lists before serving as a subject in an actual list-learning experiment. In that sense, subjects of typical list-learning experiments are naive only with respect to the demands of the serial task required by a particular experiment. Were it not for the thousands of lists such subjects had learned previously, they may have lacked essential skills for performing the task(s) asked of them. Although the nature of those skills has been the focus of well-known discussions of the serial organization of behavior stimulated by Lashley's (1951) classic report, we are not aware of any theory of the ontogeny of those skills, either in animals or in human beings.

The purpose of the present experiment was to teach successive lists to a nonverbal organism with the ultimate goal of revealing changes in strategies used to master these lists. The absence of a comparable literature in studies of human memory creates difficulties in attempting to anticipate how an animal would cope with a multiple-list problem. It is, of course, well established that human adults rely on efficient strategies of encoding, storing, and retrieving lists with increas-

ing list-learning experience (Crowder, 1976). It is by no means clear, however, how those strategies develop. Striking developmental changes in the use of rehearsal strategies by children have been reported by Flavell, Beach, and Chinsky (1966), yet no data are available regarding the list-learning strategies of preverbal infants (see Chalmers & McGonigle, 1984, for an interesting analysis of precursors of list learning in young children).

We are well aware that it will take a substantial number of experiments to develop a model that characterizes adequately the acquisition of serial competence by a nonverbal organism and to define the best methods for training nonverbal organisms to learn to produce lists "on demand," as a young child might. Only then can one focus on the contribution of verbal skills in comparative studies of list learning by animals and human subjects. It should be recognized, however, that preverbal human infants face problems similar to our monkeys when they attempt to master their initial lists (e.g., arranging blocks or other toys in particular sequences) and that, at present, no information is available about the acquisition of lists by preverbal infants. We expect, therefore, that the systematic investigation of the development of list-learning strategies by a monkey will ultimately place in a comparative perspective analogous strategies that a preverbal infant might use before language becomes a relevant factor, and that experiments of the type described in this study will draw attention to the need to address the much neglected problem of the ontogeny of the serial organization of behavior.

The paradigm we used, a simultaneous chain, was developed in previous research on serial learning in nonverbal organisms (Terrace, 1984). Certain features of the simultaneous chaining paradigm distinguish it from traditional chaining paradigms. A simultaneous chain presents n list items simultaneously (e.g., A, B, C, D, E) in arrays in which the spatial arrangement of the items is varied across trials. The subject's task is to respond to each item in the sequence defined by the experimenter (e.g., A→B→C→D→E). The correct execution of that sequence results in food reward.

Because the configuration of the items to which the subject responds is changed from trial to trial, there is no unique chain of conditioned responses that will produce reinforcement. The subject is not given any differential feedback concerning the correctness of the sequence it produces until the sequence is completed. The list items to which the subject is required to respond are presented until an error occurs or until the sequence is correctly executed. It follows that the stimulus array of list items provides no information regarding the serial order of a correct sequence.

These features of a simultaneous chain have an important bearing on interpretations of the subject's performance. Consider, for example, a specific transition during the execution of the sequence. Having responded to Item B, the subject is not provided with any discriminative stimulus that would indicate that its next response should be directed to C (as opposed to A, D, or E), and so on, for all other transitions. Thus, like human subjects producing a rote list of arbitrary stimuli, a nonverbal organism must fall back on its representation of the sequence in working memory to obtain a cue as to which item it must next respond to, which items to avoid at each position of the sequence, or both.

Experiment 1: Acquisition of Multiple Lists

The purpose of Experiment 1 was to train rhesus monkeys to learn to produce a series of lists, each consisting of different photographs. Photographs (in each instance, of natural objects) were selected as list items to minimize the likelihood of interlist interference that would result from lists composed of artificial stimuli (e.g., colors or geometric shapes). The basic questions addressed in Experiment 1 were whether there would be any reduction in the training time needed to master successive lists and whether systemic patterns of error elimination would emerge as each list was acquired.

Method

Subjects. The subjects were two male rhesus monkeys (*Macaca mulatta*), Franklin and Rutherford. They were born in the primate facility of the New York State Psychiatric Institute and housed there in accordance with current National Institutes of Health (NIH) guidelines. At the beginning of training, Franklin and Rutherford were 8 and 9 years old, respectively, and experimentally naive. Prior to the current study, Rutherford served as a pilot subject in a study similar to the present one, but he was not trained to criterion. The subjects were fed a diet of monkey chow, high-protein bread, and fruit to supplement the food obtained during experimental sessions. They were maintained at 80% to 95% of their ad libitum body weights throughout the experiment.

Apparatus. The animals were tested individually in a chamber whose dimensions were 48 (w) × 53 (d) × 53 (h) cm. One wall of the chamber, which was made of aluminum, contained a cutout to accommodate a 30.48-cm diagonal Sony color monitor. The other walls of the chamber were made of lexan 1.9 cm thick. A Gerbrands model 5110 pellet dispenser delivered 190-mg Noyes banana-flavored food pellets to a food cup that was located at floor level on the right side of the metal wall. The experimental chamber was located in a BRS/LVE sound-insulated box 0.97 (w) × 0.97 (d) × 1.27 (h) cm. Outside noises were masked by pink noise (modulated white noise) presented inside the experimental chamber. A color video camera located inside the chamber allowed monitoring and videotaping of each experimental session.

Responses to the monitor were detected by a touch frame (BFA Technologies) attached to the front of the Sony monitor. The touch frame contained 32 horizontal and 24 vertical infrared light-emitting diodes and detectors. Each time the infrared beams were interrupted, their x, y coordinates were transmitted to the computer and mapped onto a 512 × 400 pixel grid to determine the stimulus to which the monkey responded. Stimulus presentation, response recording, and all other aspects of the experimental session were controlled by an AT type computer.

The stimuli were stored on the hard drive of the computer, as digitized images, and presented on the video monitor in any of nine positions by means of a Targa 16 image digitizer. Their positions on the monitor were defined by the openings of a thin transparent lexan template containing nine 4 × 3.4-cm holes arranged in a symmetrical 3 × 3 matrix. The function of the template, which was positioned directly in front of the monitor, was to eliminate swiping movements from one stimulus to another. That topography of responding caused reading errors when the computer attempted to define the stimulus to which the subject responded.

The lists that our subjects were required to produce were composed of colored video images 150 (w) × 120 (h) pixels, or 7.1 (w) × 5.6 (h) cm. The 16 images we used were obtained primarily from high-quality photographs (from books, magazines, and calendars). Four lists, each containing four different photographs, were composed as follows: List 1, bird → flower → frog → seashells; List 2, tree →

weasel → dragonfly → water; List 3, elk → rocks → leaves → person; and List 4, mountain → fish → monkey → tomatoes.

Procedure. The training procedure was similar to that used in previous studies of sequence learning in pigeons (Straub & Terrace, 1981; Terrace, 1987, 1991; Terrace & Chen, 1991a, 1991b) and cebus monkeys (D'Amato & Colombo, 1988). Each list was trained, progressively, in four phases, with one item added at each phase. During Phase 1, only Item A was presented. The subject was rewarded for touching A in any of the nine positions in which it could appear on the monitor. During Phase 2, A and B were presented simultaneously, and the subject was rewarded for touching those two items in the sequence A→B. Thus, the four phases used to train each list were A, A→B, A→B→C, and A→B→C→D. Subjects were moved from one phase to the next following two consecutive sessions in which they completed correctly at least 75% of the trials. When the subjects satisfied the accuracy criterion on the four-item phase of the first list, they began training, during the next session, on the single-item phase of the succeeding list, and so on, until they completed training on the fourth list.

The monkeys were tested daily in 60-trial sessions. Two additional initial "warm-up" trials were not included in data analyses. Each trial was signaled by a 2-s warning signal (a flashing houselight). A new trial began after a variable intertrial interval (ITI) whose mean value was 10 s. On each trial, the items composing the list were presented simultaneously in any of the nine possible positions. A subroutine of the program that controlled each session used an algorithm for generating random configurations that determined which positions would be occupied by which items on each trial. List items were available for 20 s or until a correct sequence was produced, whichever occurred first. The first response to an item (that occurred in the appropriate order) produced a brief (330 ms) presentation of a red border around that item (0.4 cm wide) and a brief (330 ms) high-frequency tone. Multiple touches to an individual item were permitted so long as they did not violate the sequence, but they were not accompanied by feedback. For example, A→A→A→B was permitted, but A→A→B→A was not. Correct sequences were reinforced with a food pellet. Incorrect responses (touching an item out of order, in either a backward or forward direction) resulted in the immediate termination of a trial and a 16-s delay of the onset of the next trial. During the time-out delay (TO), the houselight was extinguished. Failure to respond on any trial, within 20 s of its onset, had the same consequences as an error of commission.

Results and Discussion

Acquisition. The two- and three-item phases of training were mastered with increasing efficiency over the course of successive lists. Indeed, Franklin satisfied the accuracy criterion in the minimum possible number of sessions during the two-item phase of training on Lists 2–4 and during the three-item phase on Lists 3 and 4. Rutherford also completed the two-item phase in the minimum possible number of sessions following List 1 and, with the exception of List 3, needed progressively fewer trials to satisfy the accuracy criterion at the three-item phase on each succeeding list. Acquisition of the four-item phase was more variable.

The relevant data are shown in Figure 1. A two-way (List × Phase of Training) analysis of variance (ANOVA) of trials to criterion for both animals showed a significant effect of phase of training, $F(2, 2) = 253.14, p < .01$, with no significant effect of list and no significant List × Training interaction.

Post hoc analyses (Tukey's honestly significant difference [HSD] tests) of the phase-of-training effect showed that Phase 4 took significantly longer than either Phase 2 or Phase 3 and that Phases 2 and 3 did not differ from one another. In this test, and all subsequent post hoc tests, the alpha level was set at .05.

The rapid acquisition of the three-item sequence during Phase 4 can be considered as the first step toward developing a list-learning strategy. None is needed during the two-item phase because to produce the sequence A→B, the subject simply has to search for the first item (A) and then respond to the other item (B) by default. Once a third item is added, however, the subject must choose between B and C to execute the sequence correctly. The rapid acquisition of A→B→C suggests that the subject has, by virtue of his training on the A and A→B phases, unitized A→B as a chunk and that he then responds to C by default.

Errors. An analysis of the relative frequencies of errors was calculated for each subject at each position. That analysis, which is summarized in Table 1, showed that, during the four-item phase of training, there were more errors at Position C than at any of the other positions. Because Position C was the only position at which a forward (A→B→D) and a backward (A→B→A) error could occur, a larger number of errors would be expected at that position by chance. However, at the beginning of training on each list, the distribution of errors at C was asymmetrical, with more forward errors than backward errors. As can be seen in Figure 2, which shows the relative frequency of forward and backward errors during the first two and the last two sessions on each list, the forward errors at C and backward errors at C and D decreased significantly during training. The error distributions for each list are shown in Figure 2.

Given that only two monkeys were available for this study and that there was no opportunity to scale the difficulty of each of our lists, we have no way of knowing to what extent particular error patterns reflect individual differences, list difficulty, or some interaction of those variables. For reasons that we could not discern, List 4 seemed more difficult than the preceding lists. As the following analyses suggest, it also appears that our two subjects reacted differently to some of the earlier lists, either because of the nature of the lists or because of variations in mood (a not uncommon problem with our monkeys). Further research with large groups of subjects and appropriate controls is needed to unconfound all of these variables.

Three-way ANOVAs (List × Position × Training) of the arcsine transformed relative frequency of forward errors at Positions A, B, and C were calculated for each animal. For Rutherford, there was a significant decrease in forward errors across training, $F(1, 1) = 220.89, p < .05$, and a significant List × Position interaction, $F(6, 6) = 9.86, p < .01$. A simple main effects post hoc analysis of the List × Position interaction showed that List 1 had more errors at B than Lists 2 and 4, with List 3 at an intermediate level. More errors occurred on List 4 at Position B than at A or C.

An analysis of Franklin's forward errors showed a significant position effect, $F(3, 3) = 24.47, p < .05$, and significant
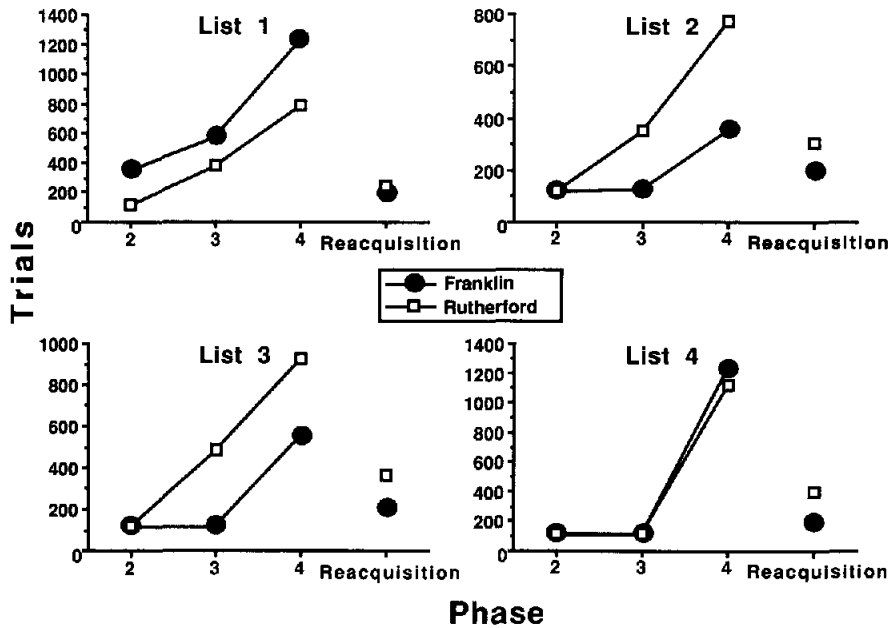
Figure 1. The mean number of trials to criterion for each list by the rhesus monkeys used in Experiment 1. (The right-hand points show the number of trials needed to satisfy the accuracy criterion during the reacquisition of each list [Experiment 2].)

interactions of List × Position, $F(6, 6) = 6.19$, $p < .05$, and Position × Training, $F(2, 2) = 62.08$, $p < .05$. Post hoc analyses of the List × Position interaction showed that the errors on Lists 1 and 4 did not differ across positions and that there were fewer errors at Position A than at B or C on Lists 2 and 3. Post hoc analyses of the Position × Training interaction showed that there were more errors at Position C at the beginning of training and more errors remaining at Position B at the end of training. A significant decrease in errors from the beginning to the end of training was observed only at Position C.

Three-way ANOVAs (List × Position × Training) were also conducted for the arcsine transformed relative frequencies of

Table 1
Percentage of Errors at Each Position During Four-Item Phase of Training (All Sessions) on Lists 1–4

Subject	List	Position			
		A	B	C	D
Franklin	1	7.4	11.2	51.6	26.4
	2	5.6	11.2	16.0	14.6
	3	2.2	9.8	28.9	22.8
	4	6.6	22.2	32.6	11.8
Rutherford	1	10.0	32.6	18.2	10.8
	2	8.2	11.6	26.6	2.2
	3	8.4	20.4	22.0	11.6
	4	4.8	12.2	32.0	19.4

backward errors at Positions C and D for each subject. A significant List × Position interaction was obtained from Rutherford's errors, $F(3, 3) = 15.07$, $p < .05$. Simple main effects analysis of this interaction showed a significant list effect at Position C, with more errors at C on Lists 2 and 3 than on Lists 1 and 4, although post hoc analyses of this effect using Tukey's HSD test showed no significant difference across means. There was also a marginally significant decrease in Rutherford's backward errors from beginning to end of training, $F(1, 1) = 111.86$, $p < .06$.

The analysis of Franklin's backward errors showed significant main effects of position, $F(1, 1) = 7226.2$, $p < .01$, and of training, $F(1, 1) = 2776.69$, $p < .05$, and a significant List × Training interaction, $F(3, 3) = 11.88$, $p < .05$. Simple main effects post hoc analyses of the interaction showed a significant decrease in backward errors across training during List 1 and marginally significant decreases during Lists 2 ($p < .07$) and 4 ($p < .06$). There was no decrease in backward errors across training on List 3 for Franklin.

Latencies. Median response latencies on correctly completed trials were similar across all lists. The longest latencies occurred to the first item. Responses to the two interior items were shorter and did not differ from one another, and latencies to the last item were longer for one animal (Rutherford) and did not differ from the interior items for the other animal (Franklin). Latencies on incorrect trials showed the same pattern but were consistently longer than latencies on correct trials. Figure 3 shows the means of the median latency of responding on correct trials by each animal at each of the four positions on each list.
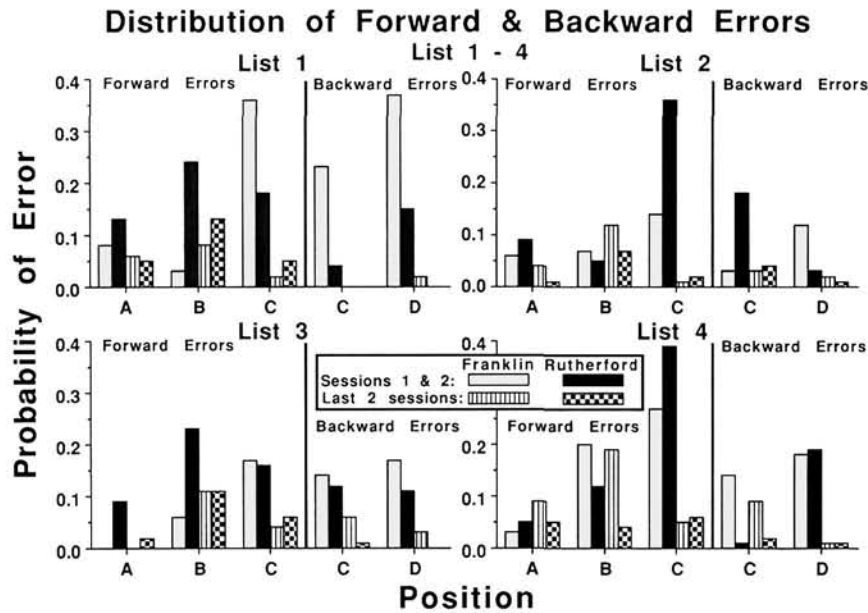
Figure 2. The relative frequency of forward and backward errors at each position in the list during the first two and the last two sessions, averaged across lists, for each subject on each list in Experiment 1.

Two-way ANOVAs (List × Position) were performed on median latencies for the last five sessions of the four-item phase of acquisition for each subject. Rutherford showed significant effects of list, $F(3, 12) = 23.46$, $p < .01$, and of position, $F(3, 12) = 322.99$, $p < .01$, and a significant List × Position interaction, $F(9, 36) = 6.99$, $p < .01$. Post hoc analyses of the interaction showed that latencies at Position A differed significantly from one another on each list. Latencies to B were lower on Lists 2 and 4 than on Lists 1 and 3. Latencies to C were longer on Lists 1 and 3 and latencies to D were longer on List 1 than on the other three lists. Across all lists, the latency to A was significantly longer than the latency at any other position, and, with the exception of List 3, the latency to D was significantly longer than those to B and C, which did not differ.

Analyses of Franklin's latencies yielded only a significant main effect of position, $F(3, 12) = 221.53$, $p < .01$. Post hoc analyses using Tukey's HSD test showed that the latency to Item A was significantly longer than that it was to any of the other three items and that the latencies to Items B, C, and D did not differ from one another.

The two main features of pattern of latencies to each list item that we observed, both on correct and on incorrect trials, were similar to those reported by D'Amato and Colombo (1988). Latencies of correct responses to A were significantly longer than latencies to any other item. This could result from a variety of mechanical or psychological factors. Of least interest is the mechanical explanation that the subject was surprised by the trial onset and had to orient himself to the stimulus display. We regard this as unlikely because direct daily observations of each subject by means of a video monitor and reviews of videotapes showed that the 2-s warning prior to the onset of a trial (a momentary change in the intensity of the houselight) was effective on virtually every trial in that, following the warning stimulus, the subject positioned himself

facing the monitor, with one or both hands poised to touch the monitor. In D'Amato and Colombo's study, the subjects initiated trials, thereby maximizing their readiness to respond to the discriminanda used in that study.

Given the different configuration of items on each trial, the long latency to A may reflect the search time needed to locate A. While searching for A, the subject may notice the locations of B, C, and D, hence the relatively short latency to those items. If this were the case, it would be expected that a negative correlation would exist between the latency to A and the latency to B. That is, the longer the subject searches for A, the more likely he is to locate B during that search, leading to a quick response to B. Conversely, the faster he locates A, the less likely it is that he will have located B, leading to a longer latency to B. To test this hypothesis, correlations between the latency to A and the latency to B were calculated for each animal for each list during the first two and the last two sessions of the four-item phase of training. None of the negative correlations were statistically significant. The analysis of Rutherford's data provided correlations that ranged between $-.25$ and $.29$ ($M = -.06$), with no significant values. The correlations from Franklin's data fell between $-.31$ and $.65$, with a mean of $.06$. Three of the correlations were statistically significant (List 1, Session 1, $r = .65$, $p < .05$; List 1, last session, $r = .38$, $p < .05$; and List 3, last session, $r = .32$, $p < .05$).

Despite the analysis just described, it remains to be determined whether the subject formulates a plan of executing the sequence on the basis of the relative positions of A, B, C, and D while searching for A. It would, for example, be of interest to record eye fixations during the execution of a trial. Having located A, one might anticipate two types of visual fixation sequences: one in which the subject moves from item to item in synchrony with touching each item (evidence of a plan formulated at the beginning of the sequence) and one in
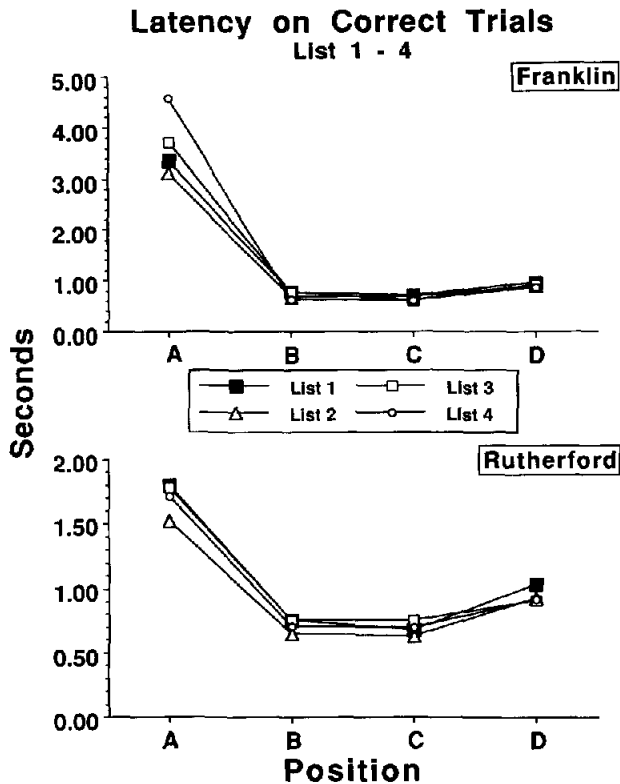
Figure 3. Mean of the median response latencies of Franklin and Rutherford to each item for Lists 1–4 in Experiment 1. (Note the differently scaled ordinates for each subject.)

which the subject searches through the items before making successive responses (evidence of having to calculate his position at each step of the sequence).

Experiment 2: Reacquisition of Lists and Subset Tests

Although there was no overlap in the photographs that composed each of the lists used in Experiment 1, one might anticipate a high degree of forgetting of each list as more lists are learned because all of the list items were photographs. To determine the extent to which individual lists were forgotten after training on all lists was completed, each list was retrained in Experiment 2. The method of savings (Ebbinghaus, 1885/1964) was used to evaluate the retention of each list.

Experiment 2 also evaluated the subjects' knowledge of the ordinal relationship between nonadjacent events. The subset test we used, which was modeled after one developed for pigeons (Straub & Terrace, 1981) and later used by D'Amato and Colombo (1988), presented all possible pairs of two items from one of the four-item lists after the acquisition criterion for that list was satisfied.

Method

Subjects. The subjects of Experiment 2 were the same as those who served in Experiment 1 (Franklin and Rutherford).

Apparatus and stimuli. The same apparatus was used to present the same four 4-item lists used in Experiment 1.

Procedure. Experiment 2 followed Experiment 1, with no break in training. Following the acquisition of the fourth list of Experiment 1, both subjects were retrained on each list, in the same order followed in Experiment 1. In each instance, the same acquisition criterion was used (two consecutive 60-trial sessions in which the subject completed correctly at least 75% of all trials, i.e., at least 45 correctly completed trials per session).

Because it was expected that the subjects would make few errors on early phases of familiar lists, the first day of training on each list consisted of 20-trial blocks of each phase of training (1, 2, 3, and 4 items). During the first session only, subjects were advanced from one phase to the next independently of their performance on each phase. In each instance, however, performance was above chance (and generally quite high) on each block of each list. The percentage of correctly completed sequences for all 20-trial blocks of Phase 1 was 100% on all lists for both subjects, between 90% and 100% for Phase 2, and, with one exception (65% by Rutherford on List 3), between 80% and 95% for Phase 3. For the four-item phase, the percentage correct was somewhat lower (Rutherford, 74%, 53%, 45%, and 27%; and Franklin, 90%, 74%, 80%, and 80%, respectively, for Lists 1–4). Following the first 20-trial session, all sessions consisted of 60 four-item trials. Training on a particular list was continued until the accuracy criterion was satisfied.

To determine the extent of the subjects' knowledge of adjacent and nonadjacent items, each animal was tested on all possible two-item subsets that could be derived from that list (AB, AC, AD, BC, BD, and CD) following criterion performance on each four-item list. The subset test consisted of three blocks of 24 trials. Each subset was presented four times within each block. The program that controlled the experiment used an algorithm to generate random orders of subsets within each block with the constraint that there were four presentations of each subset in each block. A three-item subset test was administered on the following day. Each subset (ABC, ABD, ACD, and BCD) was presented in a random order in two 24-trial blocks. All aspects of the experimental procedure, including reinforcement contingencies, remained in effect during each subset test.

Results and Discussion

Reacquisition. With the exception of the first session, all reacquisition sessions consisted of training on four-item lists. Accordingly, we will focus on that phase of training in evaluating the retention of each list. As shown in the right-hand section of Figure 1, reacquisition of each list was much more rapid than original acquisition. The mean number of trials to criterion for the four lists was 875 for original acquisition and 264 for reacquisition. On Lists 2, 3, and 4, Franklin satisfied the accuracy criterion in two sessions, the minimum number of trials in which it was possible to do so. Rutherford needed 3, 6, and 7 sessions, respectively, to satisfy the accuracy criterion for those lists. A two-way ANOVA (Lists × Acquisition–Reacquisition) of trials to criterion showed a significant decrease in the number of trials required for reacquisition, $F(1, 1) = 285.48, p < .05$.

Comparisons of individual performance for the four-item phase on each list during Experiment 1 (acquisition) and Experiment 2 (reacquisition) are shown in Table 2. Across lists, the average savings shown by Franklin was 68.9%; by Rutherford, 63.6%. Franklin's savings score on List 2 (44.6%) is low not because he had difficulty relearning that list but because he acquired the list in relatively few trials (359). The 199 trials he needed to relearn the list is very close to the

Table 2

Trials to Criterion for Acquisition and Reacquisition of Four-Item Phase, Lists 1–4

Subject	List	Acquisition	Reacquisition	% savings
Franklin	1	1231	199	83.8
	2	359	199	44.6
	3	557	204	63.3
	4	1240	199	84.0
Rutherford	1	793	249	68.6
	2	775	304	60.8
	3	927	361	61.1
	4	1118	402	64.0

minimum number of trials required to reach criterion (140) during reacquisition of the lists in Experiment 2.

Errors. The rapid reacquisition of the lists is evident in the low relative frequencies of errors at all positions for both subjects. These are shown in Table 3. The pattern of Rutherford's errors was similar to that observed during acquisition, albeit on a much smaller scale (cf. Figure 2). However, Franklin, who reacquired each of the four lists almost immediately, produced very few forward errors at C and very few backward errors at D. His overall pattern of errors during reacquisition was similar to that observed at the end of acquisition.

The similarity between the error distributions observed in Experiments 1 and 2 and the high degree of savings we obtained indicate that the photographs that composed each list differed sufficiently so as to minimize interference between them. If, for example, Item 1 of List 2 resembled Item 3 of List 4, one would expect a high frequency of initial responses to Item 3 on List 4. No such deviations from the distribution shown in Table 3 were observed.

Latencies. As shown in Table 4, the pattern of latencies during the four-item phase was similar to that observed during acquisition. Three-way ANOVAs (List × Position × Acquisition–Reacquisition) of median latencies for each animal during the final two sessions replicated the significant position effects obtained in Experiment 1: Franklin, $F(3, 3) = 541.32$, $p < .01$, and Rutherford, $F(3, 3) = 632.86$, $p < .01$. Both animals responded more slowly to A. Rutherford responded more slowly to D than to B or C. However, analyses of the latencies of both animals yielded a significant List × Position interaction: Franklin, $F(9, 9) = 8.89$, $p < .01$, and Rutherford, $F(9, 9) = 3.20$, $p < .05$. For Rutherford, this interaction was a consequence of variability in the latency of responding to the last item of each list. Latency to responding to D was significantly longer than to B or C on Lists 1 and 2, but it did not differ significantly from B and C on Lists 3 and 4.

Analyses of Franklin's latencies showed a significant List × Acquisition–Reacquisition interaction, $F(3, 3) = 15.19$, $p < .05$, and a significant List × Position × Acquisition–Reacquisition interaction, $F(9, 9) = 14.48$, $p < .01$. These interactions resulted from a significant increase in Franklin's latency to A during reacquisition of List 4.

Performance on subset tests. With the exception of the subset CD, accuracy of responding on all two-item subsets was consistently high. The relevant data are shown in Figure 4. Accuracy on all three-item subsets was uniformly high (>67% in each instance). A two-way ANOVA (List × Subset) was conducted on percentage correct for each animal on two-item subsets. A significant main effect of subsets was obtained from both subjects' data: Franklin, $F(5, 10) = 17.25$, $p < .01$, and Rutherford, $F(5, 10) = 11.46$, $p < .01$. Also, there was a significant List × Subset interaction: Franklin, $F(15, 30) = 4.22$, $p < .01$, and Rutherford, $F(15, 30) = 2.10$, $p < .05$. Post hoc analyses of the interaction for both animals showed that although the CD subset had the lowest percentage for all lists, it was significantly lower than the other subsets only on Lists 2 and 4 for Franklin and List 1 for Rutherford. A similar analysis of percentage correct on three-item subsets yielded no significant effects for Franklin's data. The analysis of Rutherford's data showed a significant List × Subsets interaction, $F(9, 9) = 3.66$, $p < .05$; however, a simple main effects post hoc analysis yielded no significant effects.

The data obtained on the two- and three-item subset tests of the present experiment are similar to those reported by D'Amato and Colombo (1988), who trained cebus monkeys to master five-item lists. Unlike pigeons, who, in Terrace's (1987) study, performed at chance on subsets constructed from interior items (BC, BD, and CD), cebus monkeys performed accurately on those and all other subsets. As in the present experiment, however, a decrement in accuracy of responding was observed to the subset containing the last two items of the list (DE).

Another difference between the performance of pigeons and monkeys on two-item subsets of lists they have learned to produce was revealed by analyses of the latencies of responding to the first and second items of each subset. The data obtained from Franklin and Rutherford, which are shown in Figure 5, are similar to those reported by D'Amato and Colombo (1988). The farther removed Item 1 was from the beginning of the list, the longer was the latency to Item 1 of each two-item subset. Two-way ANOVAs (List × Position) of the latencies to Item 1 of each subset for each animal yielded significant position effects for both subjects: Franklin, $F(2, 4) = 80.62$, $p < .01$, and Rutherford, $F(2, 4) = 41.39$, $p < .01$. Post hoc analyses (Tukey HSD tests) showed that, for

Table 3

Relative Frequencies of Errors Reacquisition of Lists 1–4 in Experiment 2

Early vs. late training by position	Franklin					Rutherford				
	Forward			Backward		Forward			Backward	
Position	A	B	C	C	D	A	B	C	C	D
Early (first 2 sessions)	0.11	0.08	0.04	0.06	0.03	0.05	0.13	0.13	0.03	0.01
Late (last 2 sessions)	0.03	0.04	0.03	0.03	0.01	0.03	0.05	0.03	0.01	0.02

Table 4

Means of Median Latencies of Responding to Each Item, Across Lists, During Criterion Sessions of Experiment 1 (Acquisition) and Experiment 2 (Reacquisition)

Subject by criterion session	Position			
	A	B	C	D
Franklin				
Acquisition	3.34	0.93	0.91	0.94
Reacquisition	3.74	1.19	0.94	0.95
Rutherford				
Acquisition	1.85	0.94	0.86	1.02
Reacquisition	1.97	0.79	0.8	0.99

both animals, the latency to respond to the first item was significantly longer when that item was C than when it was either A or B. The latencies to A and B did not differ. Franklin's latencies also yielded a significant List × Item interaction, $F(6, 12) = 3.32$, $p < .05$. Simple main effects analyses of this interaction (which can be seen in the left-hand panel of Figure 5) showed that on all lists, the latency to respond to Item 1 was significantly longer when it was C than when it was A or B, and that that effect was stronger during the first two than the last two lists.

The latency of responding to Item 2 of each subset varied with the number of items from the original list that intervened between Item 1 and Item 2 of the subset. The relevant data are shown in the right-hand panel of Figure 5. Two-way ANOVAs (List × Number of Intervening Items) were conducted for each animal. Rutherford's latencies yielded a significant effect of intervening item, $F(2, 4) = 15.68$, $p < .05$. Post hoc analyses (Tukey's HSD test) showed that Rutherford's latency to the second item was significantly longer with two intervening items than with zero or one. Franklin's latencies yielded a significant interaction of List × Intervening Items, $F(6, 12) = 4.10$, $p < .05$. Simple main effects analyses of this interaction showed that, on Lists 1–3, the latency to
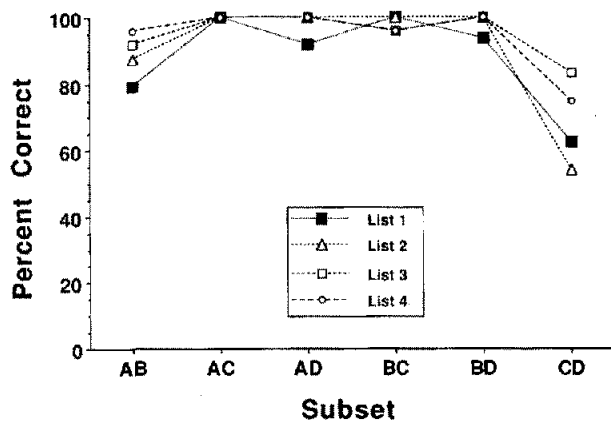

Figure 4. The mean percentage of correctly completed sequences in Experiment 2 for each two-item subset that could be derived from the four-item lists mastered by both subjects. (See text for additional details.)

the second item was longer with two intervening items than with zero or one, but that the difference disappeared on the last list. The same effect of extended training on the diminution of differences in latency as a function of number of intervening items was observed in cebus monkeys by D'Amato (M. R. D'Amato, personal communication, January 25, 1990).

Pigeons, given similar subset tests, showed no evidence of monotonic latency functions (Straub & Terrace, 1981; Terrace, 1991). This, and the differences in accuracy on subset tests described earlier, suggests that monkeys formed linear representations of the sequences they learned and that pigeons did not (see D'Amato & Colombo, 1988, and Terrace, 1991, for discussions of these differences).

Experiment 3: Production of Multiple Lists Within a Single Session

A familiar feature of human list memory that is all too easily taken for granted is the ability to produce any of a large number of alternative lists on demand (e.g., a particular telephone number, the succession of stations on a particular railroad line, and so on). The cognitive precursors of that ability are by no means clear. It is not known, for example, just when and how children develop strategies to produce particular lists on demand and, in particular, the extent to which that ability requires verbal mediation. Because multiple-list learning in animals has not been investigated previously, there is no empirical basis for predicting the degree to which an animal could produce a particular list under circumstances in which it cannot predict which of a set of alternative lists will be required. The purpose of Experiment 3 was to determine whether monkeys could produce, on demand, any of the four lists they learned in Experiment 1 within the same session.

Method

Subjects. Franklin and Rutherford served as the subjects in Experiment 3.

Apparatus and stimuli. The apparatus and stimuli were identical to those in Experiments 1 and 2.

Procedure. Following completion of the subset tests of Experiment 2, each subject was trained to execute two of the four lists he was trained to produce in Experiments 1 and 2 within a *single* session. All aspects of list training followed in Experiment 3 were identical to that followed in Experiment 2 with the exception of the values of the ITI and TO parameters and the number of lists on which the subject was tested in a particular session.

During Experiment 3, the ITI was reduced from 10 to 3 s, and the TO that followed incorrect responses was reduced from 16 to 6 s. On any trial, the four items of one of a pair of lists would appear. Which of the two lists chosen for that session would be presented on a given trial was determined by an algorithm for randomizing sequences, with the restriction that p(List 1) $= p$(List 2) $= .5$ and that no more than three presentations of either list could occur on successive trials. Each combination of two lists was presented until the accuracy criterion used in Experiment 1 and 2 was satisfied. As in Experiment 2, the first day of training on each pair of lists consisted of one 20-trial block of each of the four phases of training (1, 2, 3, and 4 items). Independent of the performance on any of these blocks (which, in each instance, was greater than 75% correct), the subject's next session
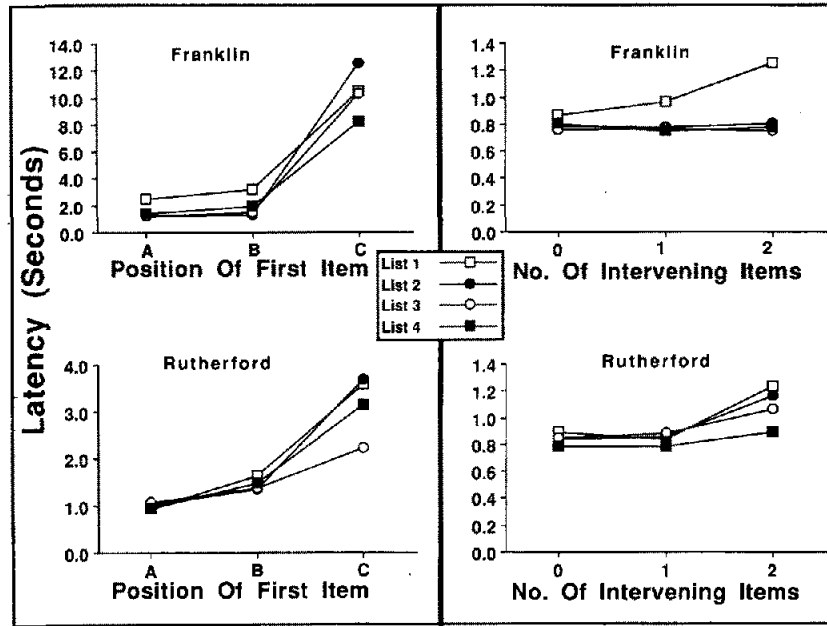
Figure 5. Mean of the median latency of responding to the first item during two-item subset tests as a function of the position of the first item (left panel) and as a function of the number of intervening items in the original list (right panel).

consisted of 60 trials of two 4-item lists (30 trials of each list, randomly distributed within the session as described earlier). Training on a particular pair of lists continued until the subject responded correctly on 75% of all trials during two successive sessions. Once the criterion was met, training began on a second pair of lists. This procedure was repeated until all of the possible six pairs of lists that could be derived from the original four lists were tested during the same session. The sequence of pairs used during multiple list training was 1 and 2, 3 and 4, 1 and 3, 2 and 4, 1 and 4, and 2 and 3.

Training on all four lists began during the session that followed training on the final pair of lists. The first session in which the subject was required to produce all four lists consisted of three 20-trial blocks. During the first block, the subject was required to respond to the final two lists on which he had been tested the previous day (last session of two-list training). During the second block, a third list was added. The fourth list was added in the final block. In each instance, the subject was required to produce a four-item list. Independent of the level of performance achieved during the first session (which in each instance was at least 65% correct), the subject was required to respond to each of the four lists during all subsequent sessions, each of which contained 80 trials. The probability of items of a particular list appearing on a particular trial was .25. The sequence of lists was generated by an algorithm that ensured that the range of the number of presentations of each list was 17–23. As previously, 2 consecutive days were required on which the overall performance level was 75% correctly completed trials per session. In addition, each subject was required to respond correctly, on average, to at least 75% of the trials on *each* list during the two criterion sessions.

Results and Discussion

Acquisition. As shown in Table 5, the acquisition of pairs of four-item lists was comparable to the reacquisition of the individual lists. During the four-item phase of training, Frank-

lin required a mean of 238 trials per pair to satisfy the accuracy criterion; Rutherford required a mean of 237 trials per pair. Thus, acquisition of pairs of four-item lists was much more rapid than the acquisition of the individual lists. A one-way ANOVA was performed on trials to criterion for the four-item phase of acquisition (mean of all four lists), reacquisition (mean of all four lists), and each of the pairs (1 and 2, 3 and 4, 1 and 3, 2 and 4, 1 and 4, and 2 and 3). A significant difference was obtained between acquisition and all other conditions, with no differences between reacquisition and any of the pairs, $F(7, 7) = 12.24, p < .01$.

During training in which they were required to respond to all four lists during a single session, both animals required only four full sessions to satisfy the accuracy criterion (following the initial 20-trial block). This came to 340 trials, a number that is deceptively high because of the requirement that the criterion be met on all four lists. During the first 20-trial block in which they were tested on all four lists, Franklin and Rutherford responded correctly on 65% and 85% of the trials, respectively. The overall level of accuracy for both animals was at least 75% during all subsequent sessions. Were

Table 5

Mean Trials to Criterion for Acquisition of Individual Lists and Multiple Lists, Lists 1–4

Subject	Acquisition	Reacquisition	Acquisition of pairs	Acquisition of multiple lists
Franklin	847	200	238	340
Rutherford	903	329	237	340

it not for the requirement that, on average, at least 75% of the trials for each list be completed correctly, multiple-list training would have ended after two sessions. A one-way ANOVA was conducted on mean trials to criterion across lists for acquisition, reacquisition, acquisition of pairs, and acquisition of multiple lists. There was no difference between trials to criterion for reacquisition, acquisition of pairs, and acquisition of multiple lists, but, as previously, significantly more trials were required during the initial acquisition of each list, $F(3, 3) = 42.69, p < .01$.

Latencies. Our latency data suggest that the subject does not have to be set for a particular list to perform accurately when different lists are presented during the same session. Latencies were not affected by the inclusion of additional lists within a session, as shown in Table 6. Two-way ANOVAs (Training Condition [acquisition, pairs, or multiple lists] × List) of the latency to respond to the first item were conducted for each subject. No significant effects were found, suggesting that the subjects were able to execute, with no decrease in accuracy, any list without prior knowledge of which list would be presented.

Experiment 4: Acquisition of Lists 5–8

The results of Experiments 1–3 show that monkeys can retain four arbitrary lists, each composed of four arbitrary photographs, and that they can produce a given list under circumstances in which knowledge of all four lists is required. It was also shown that, during initial training on each list, subjects learned the first three items of successive lists with increasing efficiency. In many instances, the three-item phase of training was mastered in the minimum possible number of trials needed to satisfy the accuracy criterion. Taken together, these results suggest that experimentally naive monkeys acquire rudimentary list-learning skills during the course of mastering a mere four lists but that they are far from approaching the limits of their list-learning capacity. Despite rapid decreases in trials to criterion during the two- and three-item phases following acquisition of the first list, no such decrease was observed during the four-item phase, even while learning the fourth list. To determine whether the time needed

Table 6

Means of Median Latencies of Responding to Each Item, Across Lists, During Acquisition of Single Lists, Pairs of Lists, and Multiple Lists, Lists 1–4

Subject by list type	Position			
	A	B	C	D
Franklin				
Single lists	3.69	0.71	0.64	0.90
Pairs	2.54	0.56	0.61	0.76
Multiple lists	2.62	0.59	0.56	0.78
Rutherford				
Single lists	1.70	0.72	0.69	0.95
Pairs	1.70	0.60	0.61	0.87
Multiple lists	1.74	0.57	0.60	0.90

to master Phase 4 could be reduced further, four additional four-item lists were trained.

Method

Subjects. Franklin and Rutherford served as the subjects in Experiment 4.

Apparatus and stimuli. The apparatus was identical to that in Experiments 1–3. The stimuli were digitized images similar to those used in Experiments 1–3. The four lists used in this experiment consisted of the following photographs: List 5, red leaf → clouds → snail → vegetables; List 6, cat → seashell → ant → flowers; List 7, water scenery → human face → butterfly → cherries; and List 8, garlic braid → eagle → autumn leaves → jackal.

Procedure. The training procedure was the same as that used in Experiment 1. Each list was trained in phases, with one item added after the acquisition criterion was satisfied for that phase. The criterion for moving from one phase to the next was two consecutive sessions with at least 75% of the trials correct. All session parameters (number of trials, stimulus duration, ITI, TO, and reinforcement contingencies) were identical to those used in Experiment 3.

Results and Discussion

Acquisition. Superficially, the results of Experiment 4 resemble those of Experiment 1 with respect to the number of trials needed to master successive phases of training. However, an analysis of errors during Phase 4 suggests that, by List 8, Franklin and Rutherford were adopting a new strategy for learning new lists.

The two-item phase was acquired in the minimum number of trials possible by both animals on all lists. Franklin acquired the three-item phase in the minimum number of trials (120) for two lists (Lists 6 and 8) and in only 231 trials and 212 trials for Lists 5 and 7, respectively. Rutherford reached criterion on the three-item phase in the minimum number of trials for three lists and required 388 trials on List 6. Acquisition of the four-item phase was, as in the previous four lists, variable and showed no downward trend. These data are shown in Figure 6. A two-way ANOVA (List × Phase) of trials to criterion for Lists 1–8 showed a significant effect of phase, $F(2, 2) = 124.39, p < .01$, but no list effect or interaction. Phase 4 took significantly more trials than either Phases 2 or 3, which did not differ from one another.

Errors. As shown in Table 7, the percentage of errors at each position was somewhat lower than that obtained during the acquisition of Lists 1–4 (cf. Table 1). As on those lists, this aspect of our analysis of errors showed that on Lists 5–8 the pattern of errors at each position was similar, with better performance at Positions A and D than in the middle of the list.

Although the pattern of forward errors during the four-item phase of Lists 5–8 was similar to that for Lists 1–4, the frequency of backward errors on Lists 5–8 was lower than it was during the acquisition of Lists 1–4 at the beginning as well as the end of training. The relative frequencies of both types of errors for each animal at each position for the first two and last two sessions of Lists 5–8 are shown in Figure 7. Three-way ANOVAs (List × Position × Training) of the arcsine transformed forward errors during the first two and
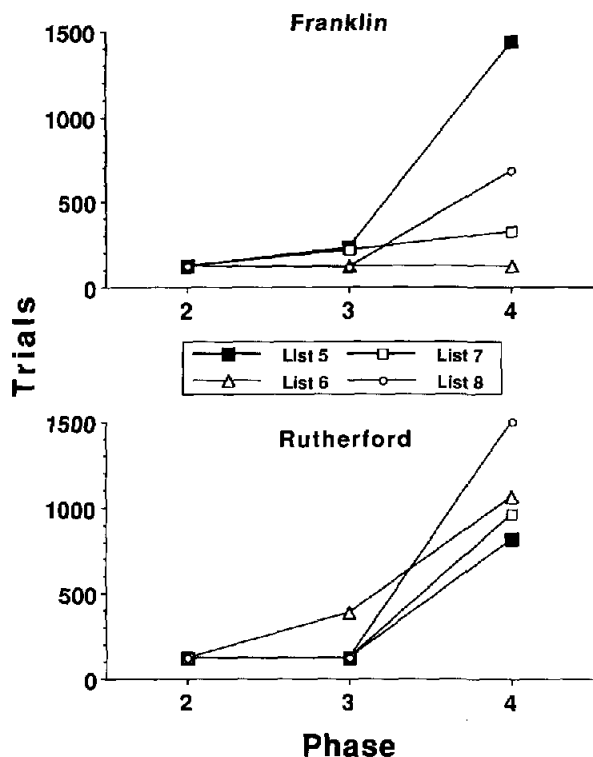
Figure 6. The mean number of trials to criterion for each of the second set of four lists learned by the rhesus monkeys in Experiment 4.

last two sessions of each list were conducted for each subject. For Franklin, errors during List 6 were eliminated from the analysis because he acquired the four-item phase in only two sessions. A significant training effect was obtained in the case of Rutherford's errors, $F(1, 1) = 505.00, p < .05$, as well as significant List × Training, $F(3, 3) = 9.63, p < .05$, and Position × Training, $F(2, 2) = 26.15, p < .05$, interactions. Post hoc analyses of the List × Training interaction showed a significant decrease in errors on all lists but List 7. The Position × Training interaction was a function of a significant decrease in errors at Position C across training, but not at A or B. No significant effects were obtained from Franklin's data, although the acquisition effect was marginal, $F(1, 1) = 129.92, p < .06$.

Three-way ANOVAs (List × Position × Training) of the arcsine transformed backward errors were conducted for each subject. No significant effects were obtained from analyses of Rutherford's data. However, a significant List × Position × Training interaction, $F(2, 2) = 33.32, p < .05$, was observed in an analysis of Franklin's errors. Post hoc analyses of this interaction produced no significant differences; the proportion of backward errors at any position for any of the lists for Franklin was so small that this interaction was spurious.

The decrease in backward errors during acquisition of Lists 5-8 compared with those shown during acquisition of Lists 1-4, even though there were no overall decreases in the number of trials to criterion during the four-item phase of the last four lists, suggests that some strategy related to efficient

list learning was developing during the acquisition of Lists 5-8. Both animals showed a consistent pattern of error change. The large number of forward errors at C at the beginning of training sharply decreased during the acquisition of the four-item phase (Franklin's data for List 6 are not included in the figure because he acquired this list in only two sessions). Although backward errors were virtually eliminated after extensive training on Lists 1-4, they occurred only at minimal levels at the beginning of training during Lists 5-8.

Latencies. The latency patterns observed during the four-item phases of Lists 5-8 were similar to those observed during the four-item phases of the first four lists. The median latencies at each position of Lists 5-8 are shown in Figure 8. The latency of responding to the first item was longest, with no difference among the latencies to the other three items. Two-way ANOVAs (List × Position) of the median latency to each position during the last five sessions of each list for Rutherford and during the last two sessions for each list for Franklin were conducted. Significant effects of list were obtained in both instances: Franklin, $F(3, 3) = 42.10, p < .01$, and Rutherford, $F(3, 12) = 4.21, p < .05$. Also, significant effects of position were found in both instances: Franklin, $F(3, 3) = 841.45, p < .01$, and Rutherford, $F(3, 12) = 53.41, p < .01$. Finally, there were significant List × Position interactions: Franklin, $F(9, 9) = 19.04, p < .01$, and Rutherford, $F(9, 36) = 4.47, p < .01$. Post hoc analyses of Rutherford's interaction showed that, for List 5, the latency of responding to A was significantly longer than the latency of responding to any other position, and that the latency of responding to D was longer than it was to B or C, which did not differ. On the subsequent three lists, only the latency of responding to A was significantly longer than any of the other latencies, which did not differ. Franklin's interaction was a function of a significant increase in the latency of responding to B during Lists 7 and 8, but the latency of responding to A was significantly longer than all of the other latencies on all lists.

General Discussion

Our results show that a monkey can become a proficient list learner during the course of learning a succession of new lists. As noted earlier, the literature on human list learning

Table 7

Percentage of Errors at Each Position During Four-Item Phase of Training (All Sessions) on Lists 5-8

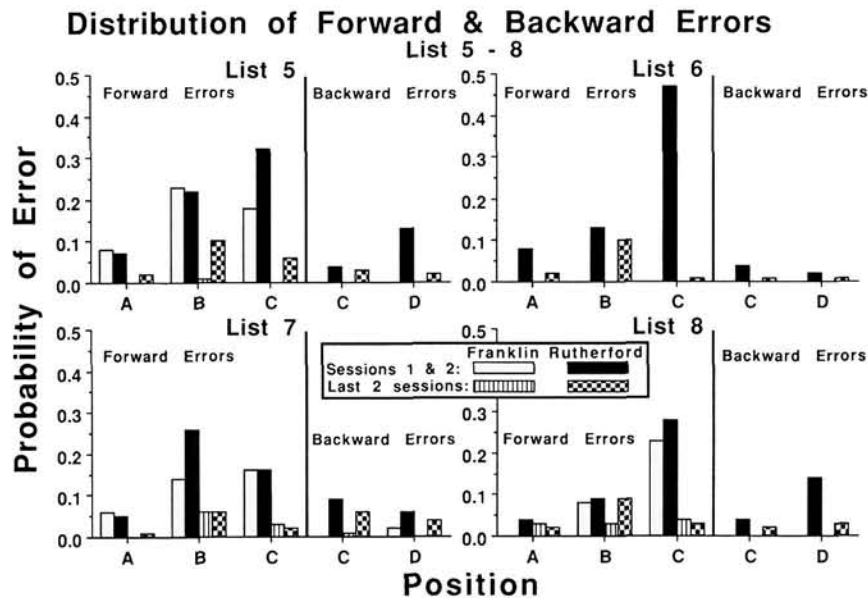| Subject	List	Position			
		A	B	C	D
Franklin	5	2.1	8.3	5.6	0.2
	6	3.5	9.5	4.0	2.0
	7	1.8	12.2	9.3	0.7
	8	2.8	7.4	9.7	1.9
Rutherford	5	8.3	13.6	21.4	3.6
	6	6.1	18.8	20.0	2.2
	7	5.0	18.1	11.4	3.5
	8	5.4	17.0	22.0	7.6

Distribution of Forward & Backward Errors

Figure 7. The relative frequency of forward and backward errors at each position in the list during the first two and last two sessions of the four-item phase of acquisition for the second set of four lists learned (Experiment 4). (Note that Franklin's data for List 6 are not presented because he acquired the four-item phase of that list in only two sessions.)

provides no theories or models of list acquisition that allow comparisons of sequence learning by monkey and human subjects. The few relevant studies on serial learning in children have used subjects who have extensive language skills (e.g., Frankel & Rollins, 1982, Perlmutter, Sophian, Mitchell, & Cavanaugh, 1981; and Brown, Brown, & Caci, 1981) and who have had many opportunities to develop strategies for performing serial learning tasks by engaging in the large variety of games in which children arrange objects in particular orders.

The task of describing what cognitive processes our monkeys brought to bear while improving their efficiency in mastering successive lists and maintaining each list in reference memory is made all the more difficult by the absence of any evidence that either monkey had approached the limits of his list-learning ability or his ability to retain the lists he learned. It seems likely, for example, that additional training on four-item lists would produce further reductions in the acquisition process than those observed in this study. How many such lists a monkey can retain is also an open question. Yet another factor (among many) that has to be explored is the repetition of items across lists. Initially, one would expect that the repetition of an item, especially when the repeated item occupies a new ordinal position, would retard list learning. One cannot, however, rule out the possibility that extensive training on such lists might result in flexible strategies for reassigning the ordinal position of a particular item. With these caveats in mind, we will try to characterize the processes our subjects used in learning successive lists by focusing first on two aspects of the acquisition process: the types of errors the subjects made while mastering each list and their latencies of responding to each item.

Elimination of errors. As in the acquisition of other complex tasks, the systematic elimination of particular types of errors provides important clues as to the nature of the learning process. In the present instance, most of the errors occurred during the four-item phase of training on each list (cf. Figure 1). This is not surprising given the procedure we used to train each list. Recall that, during the first phase of training, the subject's task is simply to respond to Item A. During Phase 2, Items A and B are presented simultaneously on each trial. Because the subject is disposed to respond to A by virtue of the training received during the first phase of training, it is a simple matter to execute an A→B sequence by first locating A and then responding to B by default. That strategy would account for our finding that, after acquiring List 1, both subjects satisfied the accuracy criterion during the two-item phase of training in the minimum possible number of sessions (cf. Figure 1).

It is not until the three-item phase of training that a default rule proves inadequate as a means of completing the sequence. During three-item training on List 1, both subjects showed a high tendency to make forward errors to Item C after responding to Item A. By the beginning of training on List 2, Franklin's tendency to make A→C errors was reduced to the point at which he was able to complete the three-item phase of training in the minimum number of sessions needed to satisfy the accuracy criterion. The same was true of Rutherford at the beginning of training on List 4. It seems reasonable to assume that the transition from two- to three-item lists (one that can be characterized as a transition that requires the subject to go beyond a default rule) was facilitated by the salience of Items A and C: Item A by virtue of its introduction at the start of training and Item C by virtue of its novelty and
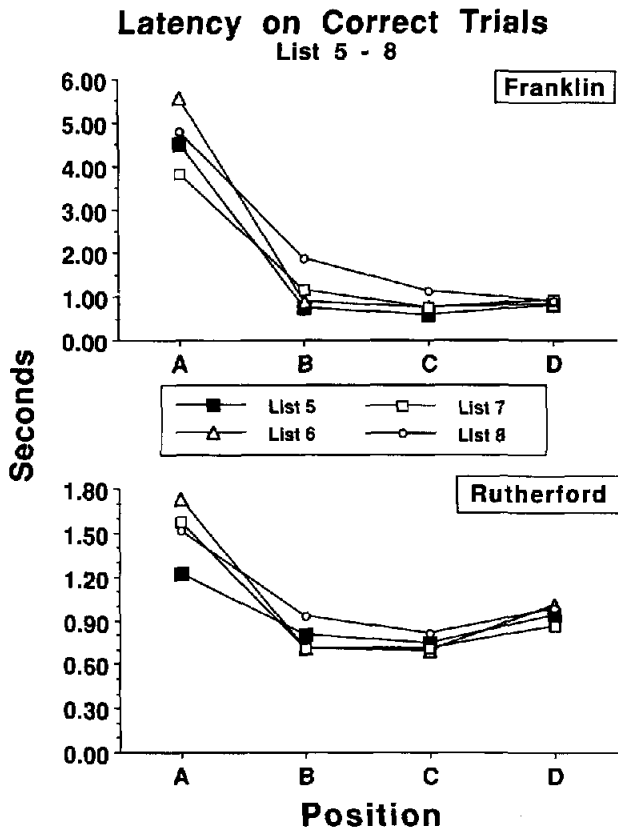
Latency on Correct Trials
List 5 - 8

Figure 8. Mean of the median response latencies of Franklin and Rutherford to each item for Lists 5–8. (Note the differently scaled ordinates for each subject.)

its association with reinforcement. Four-item lists pose yet another complication by requiring the subject to choose between two "middle" items, B and C.

As can be seen in Figure 2, both monkeys acquired four-item lists by learning to avoid two kinds of error: one-skip forward errors at Position C (B→D) and backward errors at Position D (A→B→C→B or A→B→C→A). The relative frequency of both types of error dropped precipitously by the time each subject satisfied the accuracy criterion. Furthermore, the tendency to make backward errors decreased between the first four and last four lists (cf. Figure 7), suggesting that the subjects were adopting a more efficient list-learning strategy with successive lists. The remaining errors, forward errors at C, did not decrease during the mastery of successive lists. If they had done so, Franklin and Rutherford would have been able to learn new four-item lists in the minimum possible number of trials. We have no reason to believe that training on new four-item lists would not result in the adoption of other strategies (e.g., a means of quickly eliminating errors at C) that would provide further increments in list-learning efficiency.

Latencies. The pattern of latencies to successive items on each list remained highly stable. In each instance, the latency to Item A was longest, typically twice as long as the latency

to any of the other items. Furthermore, Rutherford increased the latency of his response to the last item. For both animals, latencies to the two interior items were similar and consistently low.

The three most obvious explanations of the long latencies to A are the time needed to orient to the display of items at the onset of each trial, the time needed to search for A (and possibly other items) at the beginning of each trial, and the time needed to plan the execution of the list. Video observation of each subject's behavior revealed the reliable occurrence of search patterns of item arrays at the onset of each trial. Given the design of the present study, we have no means of assessing the relative contributions of each factor to the latency of responding to A. It is of interest, however, to note that the number of lists our subjects had to produce within a single session (1, 2, or 4) did not appear to influence the search time for Item A. This suggests that Item A of each list is recognized on an absolute basis and that the subject's representations of each list are independent of one another. As noted earlier, analyses of visual fixation patterns to the items of a list may clarify how the subject plans the execution of each list.

Retention of lists. Once a list was learned it was retained without any difficulty. Indeed, the data summarized in Table 5 show that the retention of a list was not diminished by the additional demands placed on reference memory by requiring the subject to produce as many as four different lists within the same session. The uniformly high level of savings obtained during reacquisition (Experiment 2) was not influenced by the difficulty of the list (as measured by the number of trials needed to satisfy the accuracy criterion; cf. Figure 1). Similarly, the amount of savings did not appear to be influenced by the sequence in which the lists were trained. The data shown in Table 2 indicate that learning list *n* had no adverse consequences on the retention of list *n* − 1.

This savings analysis of list retention is confounded with experience in list learning. That is, a decrease in trials to reacquire List 1 might simply be a function of list sophistication and might not reflect memory for that list. That explanation is unlikely, however, because trials to criterion did not systematically decrease with succeeding lists during reacquisition (see Table 2). Furthermore, fewer trials were required for reacquisition of Lists 1–4 than for acquisition of Lists 5–8 (cf. Table 2 and Figure 6).

These preliminary findings regarding the retention of multiple lists by monkeys suggest that one has yet to approach, by a wide margin, the limits of their capacity to retain multiple lists. If the number of photographs that a monkey can recognize is the limiting factor of its capacity to retain multiple lists, one can assume that one has barely begun to approach that limit. Although a monkey's capacity to recognize different photographs has not been put to explicit test (cf. experiments on human memory for photographs, e.g., Shepard, 1967; Standing, 1973), data obtained by Wright and his associates on a monkey's performance on a serial probe recognition task (Sands & Wright, 1980) and by Herrnstein and his associates on a pigeon's ability to categorize photographs of different natural objects (Herrnstein, 1984; Herrnstein, Loveland, & Cable, 1976; Vaughan & Greene, 1984)

suggest that monkeys are capable of recognizing hundreds of photographs.

Much stronger candidates for a constraint on the number of unique lists a monkey can retain are the manner in which they are represented and the process(es) a monkey uses to execute a list. We will assume that lists are represented linearly on the basis of latency data obtained from two-item subsets in D'Amato and Colombo's (1988) study of the acquisition of a single five-item list by cebus monkeys and on similar latency data obtained in the present study. On this view, the correct execution of a sequence requires the subject to locate Item A, to use the internal association that Item A occasions as a basis for responding to B, and so on, until it responds to the last item. Although one would expect that the number of such lists that a monkey can retain is substantially smaller than the number of Item As it can recognize, it is not clear why that is so. Given that the subject can access but a single list on the basis of its recognition of Item A, the capacity of short-term memory would not appear to be a limiting factor. One issue that we will address in future research is the difference between recovering a single item and using a single item as the first step in processing the other three list items that, by virtue of list training, function as a chunk.

All told, our results suggest that the successive list paradigm used in the present study may provide a basis for studying list learning in monkeys that is analogous to the method introduced by Ebbinghaus to study list learning by human subjects. Photographs appear to be at least as discriminable as nonsense syllables. Of greater relevance is the availability of a nonverbal organism that can readily acquire and retain many lists of photographs. That combination of an intelligent nonverbal primate and a paradigm that allows one to train numerous arbitrary lists provides a basis for establishing an animal model of serial learning, one that should help place in perspective the role of language in list learning by infant and adult human beings.

We must caution, however, as the massive literature on human list learning bears testament, that teasing apart the criterial variables relevant to list learning by monkeys will be a lengthy process. All we can conclude from our present study is that somehow two monkeys became more efficient at learning lists during the course of mastering successive four-item lists. Among the many questions we cannot answer are how representative our results are of monkeys who have been reared in more socially enriched environments, to what extent there were significant individual differences in the list-learning abilities of our two subjects, and to what extent our lists were equally difficult. We can take comfort, however, in the fact that we have a reliable procedure for studying list learning in a nonverbal organism, one that should ultimately provide a basis for delineating similarities and differences in the sequential competencies of monkeys and human beings.

References

Brown, R. M., Brown, N. L., & Caci, M. (1981). Serial position effects in young children: Temporal or spatial? *Child Development, 52,* 1191–1201.

Chalmers, M., & McGonigle, B. (1984). Are children any more logical than monkeys on the five-term series problem? *Journal of Experimental Child Psychology, 37,* 355–377.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

D'Amato, M. R., & Colombo, M. (1988). Representation of serial order in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 131–139.

D'Amato, M. R., & Colombo, M. (1989). Serial learning with wild card items by monkeys (*Cebus apella*): Implications for knowledge of ordinal position. *Journal of Comparative Psychology, 103,* 252–261.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Original work published 1885)

Flavell, J. H., Beach, D. H., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37,* 283–299.

Frankel, M. T., & Rollins, H. A., Jr. (1982). Age-related differences in clustering: A new approach. *Journal of Experimental Child Psychology, 34,* 113–122.

Herrnstein, R. J. (1984). Objects, categories, and discriminative stimuli. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition: Proceedings of the Harry Frank Guggenheim Conference, June 2–4, 1982* (pp. 233–261). Hillsdale, NJ: Erlbaum.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 285–302.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffries (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

Perlmutter, M., Sophian, C., Mitchell, D. B., & Cavanaugh, J. C. (1981). Semantic and contextual cuing of preschool children's recall. *Child Development, 52,* 873–881.

Sands, S. F., & Wright, A. A. (1980). Serial probe recognition performance by a rhesus monkey and a human with 10- and 20-item lists. *Journal of Experimental Psychology: Animal Behavior Processes, 6,* 386–396.

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior, 6,* 156–163.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25,* 207–222.

Straub, R. O., & Terrace, H. S. (1981). Generalization of serial learning in the pigeon. *Animal Learning and Behavior, 9,* 454–468.

Terrace, H. S. (1984). Simultaneous chaining: The problem it poses for traditional chaining theory. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes.* (pp. 115–137). Cambridge, MA: Ballinger.

Terrace, H. S. (1986a). A nonverbal organism's knowledge of ordinal position in a serial learning task. *Journal of Experimental Psychology: Animal Behavior Processes, 12,* 203–214.

Terrace, H. S. (1986b). Positive transfer from sequence production to sequence discrimination in a nonverbal organism. *Journal of Experimental Psychology: Animal Behavior Processes, 12,* 215–234.

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. 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.

Vaughan, W., & Greene, S. L. (1984). Acquisition of absolute discrimination in pigeons. In M. L. Commons, R. J. Herrnstein, & A.

R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (pp. 231–238). Cambridge, MA: Ballinger.

AMERICAN PSYCHOLOGICAL ASSOCIATION
SUBSCRIPTION CLAIMS INFORMATION

Today's Date:_____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do NOT duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION

MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL)

ADDRESS

DATE YOUR ORDER WAS MAILED (OR PHONED):_____

P.O. NUMBER:_____

CITY   STATE/COUNTRY   ZIP

____PREPAID ____CHECK ____CHARGE
            CHECK/CARD CLEARED DATE:_____

YOUR NAME AND PHONE NUMBER

(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.)

ISSUES: ____MISSING ____DAMAGED

TITLE   VOLUME OR YEAR   NUMBER OR MONTH

_____

_____

_____

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4-6 weeks.

———————— (TO BE FILLED OUT BY APA STAFF) ————————

DATE RECEIVED: _____   DATE OF ACTION:_____
ACTION TAKEN: _____   INV. NO. & DATE: _____
STAFF NAME: _____   LABEL NO. & DATE:_____

SEND THIS FORM TO: APA Subscription Claims, 1400 N. Uhle Street, Arlington, VA 22201-2969

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.