

Serial Learning by Rhesus Monkeys: II. Learning Four-Item Lists by Trial and Error

Karyl B. Swartz

City College of the City University of New York

Shaofu Chen and H. S. Terrace

Columbia University

Three rhesus macaque monkeys were trained to produce novel 4-item lists ($A \rightarrow B \rightarrow C \rightarrow D$) on which all items were displayed from the start of training. Subjects were previously trained to produce 4-item lists by adding one item at a time (A , $A \rightarrow B$, $A \rightarrow B \rightarrow C$, and $A \rightarrow B \rightarrow C \rightarrow D$; lists K. B. Swartz et al., 1991). Those lists could be mastered by responding to each new item last. To learn lists on which all items were displayed from the start of training, subjects had to recall the consequences of errors and correct responses to each item. Errors ended the trial; correct responses to A, B, or C allowed the trial to continue. A correct response to D produced food reward. Although the probability of executing a 4-item list correctly by chance was .04, each subject mastered 4 novel 4-item lists by trial and error. The ability of monkeys to use a trial-and-error strategy to learn novel lists provides a basis for studying the development of serial expertise in animals.

The experiment described in this article concerns the development of serial expertise in monkeys, specifically, a monkey's ability to memorize arbitrary sequences of arbitrary stimuli. Recent research has shown that cebus apella (D'Amato & Colombo, 1988; D'Amato & Colombo, 1989, 1990), Japanese macaques (Ohshiba, 1997), and rhesus macaques (Chen, Swartz, & Terrace, 1997; Swartz, Chen, & Terrace, 1991) can learn sequences composed of as many as five arbitrary stimuli. In the Swartz et al. (1991) study, monkeys mastered eight successive lists, each composed of four novel items. The ability of monkeys to learn that many lists make them good candidates for experiments on the development of serial expertise.

The lists trained in these studies were *simultaneous chains* (Terrace, 1984). A simultaneous chain differs in many respects from a *successive chain*, the traditional task used to train animals to produce specific response sequences (Skinner, 1938). The basic

difference between a simultaneous and a successive chain is the manner in which stimuli are presented. All of the items of a simultaneous chain are displayed continuously throughout each trial. In addition, the location of each item is varied from trial to trial to prevent subjects from learning the correct sequence as a series of rote movements. The subject's task is to respond to each item in a particular order, regardless of its spatial position. In a successive chain, the subject encounters each cue successively, and the position of the cues remains constant from trial to trial (e.g., the choice points in a maze).

During the execution of a successive chain, each correct response produces a distinctive change in the subject's external environment. Consider, for example, the following four-item successive and simultaneous chains:

Successive chain: $S_A: R_A \rightarrow S_B: R_B \rightarrow S_C: R_C \rightarrow S_D: R_D \rightarrow S^R$;

Simultaneous chain: $S_A S_B S_C S_D: R_A \rightarrow R_B \rightarrow R_C \rightarrow R_D \rightarrow S^R$.

In a successive chain, R_A results in the disappearance of S_A and the appearance of S_B . For example, if R_A was a left turn at the first choice point of a maze (S_A), the consequence of R_A is the appearance of the next choice point (S_B), and the disappearance of S_A . S_B also functions as a secondary reinforcer for R_A and as a discriminative stimulus for R_B . Because only one discriminative stimulus is present at each position of a successive chain, subjects could produce such chains by learning a specific response to each stimulus, rather than the ordinal relationships between successive stimuli.

The external environment provides no information during the execution of a simultaneous chain that could serve as a discriminative stimulus for a particular response. Thus, following R_A , no information is provided that the next response should be R_B (as opposed to R_C or R_D). The same is true at any of the other transitions. In each instance, the subject must rely on a representation of its position in the sequence when moving from one item to the next.

Karyl B. Swartz, Department of Psychology, Lehman College, City College of the City University of New York; Shaofu Chen and H. S. Terrace, Department of Psychology, Columbia University.

This research was supported in part by Grants MH-40462 and MH-49747 from the National Institute of Mental Health, Grant GM08225 from the National Institute of General Medical Sciences, Grant 6-60403 from the North Atlantic Treaty Organization, and the Whitehall Foundation. The data reported in this article are from Experiment 1 of a doctoral dissertation (Shaofu Chen, Identification and Retention of Ordinal Position of List Items by Rhesus Monkeys, 1993, Columbia University). We thank Brendan McGonigle for helpful comments regarding an earlier draft of this article.

Correspondence concerning this article should be addressed to Karyl B. Swartz, Department of Psychology, Lehman College, City College of the City University of New York, 250 Bedford Park Boulevard West, Bronx, New York 10468, or to H. S. Terrace, Department of Psychology, Columbia University, 1190 Amsterdam Avenue, New York, New York 10027. Electronic mail may be sent to kbslc@cunyvm.cuny.edu or to terrace@columbia.edu.

The execution of a simultaneous chain can be illustrated by the following hypothetical example. Imagine trying to enter your personal identification number (PIN) number at a cash machine on which the positions of the numbers were changed each time you went to the bank. Suppose that your PIN number was 9-2-1-5. You could not enter that (or any) PIN number by executing a sequence of distinctive motor movements, that is, first pressing the button in the lower right corner of the number pad to enter 9, then pressing the button in the upper middle position to enter 2, and so on. Instead, you would have to search for each number and keep track of your position in the sequence as you pressed different buttons. As difficult as it may seem to enter a PIN number on a cash machine on which the numbers change positions, it would be inordinately more difficult to discover the correct PIN number by trial and error. With the exception that photographs were used instead of numbers, that is precisely the task on which the subjects of this study were trained.

On a simultaneous chain, any error terminates the trial. To learn a particular list, subjects have to attend to the consequences of different types of error at each position. For example, on a four-item simultaneous chain ($A \rightarrow B \rightarrow C \rightarrow D$), there are 6 possible forward errors, $B, C, D, A \rightarrow C, A \rightarrow D, A \rightarrow B \rightarrow D$; and three possible backward errors, $A \rightarrow B \rightarrow A, A \rightarrow B \rightarrow C \rightarrow A$, and $A \rightarrow B \rightarrow C \rightarrow B$. Subjects also have to attend to the consequences of correct responses. Each correct response to A, B, or C was followed by brief auditory and visual feedback to indicate that the response was detected. That feedback was nondifferential because it provided no information as to the identity of the next item of the sequence. For example, the feedback for responding to Item B correctly provided no information that the next item was C (as opposed to A or D). A correct response to D ended the trial with a reward.

In all previous experiments on simultaneous chaining, lists were trained incrementally because of the sheer difficulty of learning the correct sequence by trial and error if all of the items were present at the start of training. (Chen, Swartz, & Terrace, 1997; D'Amato & Colombo, 1988, 1989, 1990; Ohshiba, 1997; Swartz et al., 1991). Only Item A was presented at the start of training. The remaining items were added one at a time whenever the subject satisfied an accuracy criterion. Thus, a four-item list, $A \rightarrow B \rightarrow C \rightarrow D$, would be trained in four phases, A, $A \rightarrow B$, $A \rightarrow B \rightarrow C$, and $A \rightarrow B \rightarrow C \rightarrow D$.

For a naive subject, the chance probability of selecting the first item of a four-item list correctly is .25. Because repeat responses to the same item are not counted as errors (as they would when entering a PIN number), $p(B | A) = .33$. Similarly, $p(C | A \rightarrow B)$ and $p(D | A \rightarrow B \rightarrow C) = .33$. Thus, the chance probability of completing a four-item sequence correctly is $.25 \times .33 \times .33 \times .33 = .009$. The chance probability of executing a four-item list correctly on two successive trials is less than .0001. In the face of such daunting odds, one would expect responding to extinguish, even if a subject occasionally completed the list correctly by chance. Subjects who made such lucky guesses would have much in common with the proverbial monkey who produced a line from Shakespeare when given access to a typewriter. Neither monkey could be said to know the sequence he produced.

In the Swartz et al. study, the relative frequency of backward errors decreased after subjects mastered their first four-item list. This suggests that subjects had learned to execute lists in a forward

direction. Once subjects have learned not to respond to a previously selected item, the probability of responding correctly to each item of a four-item sequence by chance is $.25 \times .33 \times .5 \times 1.00 = .04$. Accordingly, the probability of guessing the correct sequence on a new list increases by more than one order of magnitude, from .009 to .04. To the extent that experienced monkeys could also learn to reduce the relative frequency of forward errors, it seemed possible to train the subjects of the Swartz et al. (1991) study to learn new lists without the incremental method, that is, to rely exclusively on a trial-and-error strategy to determine the correct order of responding to items from a new list.

In this experiment, the subjects who served in the Swartz et al. (1991) study were trained on new four-item lists on which the incremental procedure was gradually eliminated. Training on the first four-item list began with two items (A and B, rather than just Item A). Three items were presented simultaneously from the start of training on subsequent lists (A, B, and C). Finally, new four-item lists were trained on which A, B, C, and D were presented simultaneously at the start of training.

Method

Subjects

The subjects were 3 male rhesus monkeys (*Macaca mulatta*), Rutherford, Franklin, and Dudley. They were, respectively, 10, 9, and 6 years old at the start of the experiment. Each monkey was born and housed in the primate facility of the New York State Psychiatric Institute. In addition to the food they obtained during experimental sessions, subjects were fed a mixed diet of primate chow (Prolab 26; Purina-Mills, Inc., St. Louis, MO), high-protein bread, and various fruit. Subjects' weights were monitored daily.

Rutherford and Franklin were trained in an earlier study to produce 8 four-item lists, each composed of novel photographs of natural objects (Swartz et al., 1991). Dudley, who was not a subject of that experiment, was trained to produce the same four-item lists on which Rutherford and Franklin were trained. In each instance, lists were trained by the incremental method. Prior to the start of this experiment, Rutherford and Franklin also served in an experiment on knowledge of an item's ordinal position (Chen et al., 1997).

Apparatus

Subjects were tested in an experimental chamber, 48 cm wide \times 53 cm deep \times 53 cm high, that was housed in a BRS/LVE (Silver Spring, MD) sound-isolated booth, 97 cm wide \times 97 cm deep \times 127 cm high. The front wall of the experimental chamber accommodated the picture tube of a color video monitor (30.48 cm diagonal). Banana-flavored pellets (190 mg, P. J. Noyes improved formula L; Noyes, Lancaster, NH) were dispensed by a Gerbrands pellet dispenser.

Responses to each item were detected by a light-sensitive touch frame (BFA Technology Inc.). The touch frame, which contained 32 horizontal and 24 vertical infrared light-emitting diodes and detectors, sent a signal to a computer whenever any of the infrared beams were interrupted. The computer determined the location of each touch by calculating its coordinates on a 512×400 pixel map.

List items were digitized images of color photographs of natural objects, each measuring 150×120 pixels, 7.1 cm wide \times 5.6 cm high, that were obtained from magazines, books, calendars, and travel slides. Photographs were used as list items (rather than colors or geometric forms) to provide a large set of highly discriminable stimuli for generating lists of novel items. No assumptions were made as to what a subject perceived in the photographs (i.e., natural objects or discriminable collages of colored

pixels; Herrnstein, 1985; Wasserman & Bhatt, 1992). Sample lists are shown in Figure 1.

A transparent lexan template was placed between the front wall of the experimental chamber and the touch frame to minimize the likelihood that subjects would swipe the screen as they moved from one stimulus to another. Swipes cause errors to the extent that one or more incorrect items lay in the path between two correct items. The template contained nine 4-cm \times 3.4-cm cutouts (in a 3 \times 3 matrix). These corresponded to the nine positions on the video monitor at which list items could appear.

Procedure

The procedure for training subjects to produce lists was the same as that used by Swartz et al. (1991). Subjects were trained daily in 60-trial sessions. Each trial was signaled by a 2-s ready signal (a flashing house-light). Trial duration was 20 s. Trials were separated by an intertrial interval, whose duration ranged from 5–15 s. Each session began with two warm-up trials. They were not included in the data analysis.

Brief visual and auditory stimuli were presented, following the first correct response to each photograph (a 0.5-cm wide red border around the photographs and a 1,200-Hz tone, each presented simultaneously for 0.3 s). These stimuli provided feedback for the subject that monitor detected its response. Two banana pellets were delivered following the correct completion of each sequence. Any error terminated the trial and was followed by a 10–20-s time-out during which the house light of the experimental chamber was extinguished. Neither the red border nor the auditory stimulus was presented after an error. A failure to complete a trial within 20 s also ended the trial with a time-out. Repeat responses (which occurred on fewer than 1% of the trials) did not produce any feedback and were considered correct if they occurred in the required order, for example, A \rightarrow B \rightarrow B \rightarrow C \rightarrow C \rightarrow D.

The configuration of list items was varied from trial to trial to ensure that the list could not be executed as a rote motor sequence. Configurations were selected randomly from the sets of two-, three-, and four-item configurations that could be generated by using each of the 9 stimulus positions on the video monitor: 72 [9!(9-2)] two-item; 504 [9!(9-3)] three-item, and 3,024 [9!(9-4)] four-item configurations.

The elimination of the incremental method began with each subject's 9th four-item list. Training on List 9 began with two items (A & B) rather than a single item (A). As in the Swartz et al. (1991) study, the accuracy criterion was the correct completion of 75% of the trials during two successive sessions. That criterion determined when a subject was advanced from one phase of training to the next (e.g., on List 1, from A₁ \rightarrow B₁ to A₁ \rightarrow B₁ \rightarrow C₁ and from A₁ \rightarrow B₁ \rightarrow C₁ to A₁ \rightarrow B₁ \rightarrow C₁ \rightarrow D₁), or to the first phase of training a new list (e.g., the A₂ \rightarrow B₂ phase of List 2). Consider, for example, training on Lists 9 (*water* \rightarrow *grasshopper* \rightarrow *rodent* \rightarrow *shells*) and 10 (*baby* \rightarrow *wolf* \rightarrow *grass* \rightarrow *apples*). Training on List 9 began with the first two items (*water* and *grasshopper*). If the subject satisfied the accuracy criterion for the two-item phase of training on List 9, training on List 10 began with three items (*baby*, *wolf*, and *grass*). If not, training on List 10 began with the items *baby* and *wolf*. After a subject satisfied the criterion on a list on which three items were displayed from the start of training, four-item lists were trained on which all items were displayed from the start of training.

Rutherford and Franklin were trained on 6 four-item lists on which all items were present from the start; Dudley was trained on two such lists.¹ Table 1 summarizes the list-learning histories of each subject and the training procedure used in this experiment. The right-hand column shows the probability of a subject executing a new list correctly by chance at the start of training. These probabilities are conservative in that they reflect subjects' prior experience in learning simultaneous chains. Specifically, they assume that subjects have learned not to make backward errors. Without that assumption, the probability of executing a new list correctly by chance would be considerably lower, for example, on a four-item list it would be .025 \times .33 \times .33 \times .33 = .009.

Results

Each subject learned to produce at least 2 four-item lists, each composed of novel and arbitrarily selected photographs, on which all items were presented at the start of training. Prior to those lists, subjects learned to produce 12–20 four-item lists. Those lists were also composed of novel arbitrarily selected photographs. The first 8 lists were trained by an incremental method whereby each list item was introduced individually (A, A \rightarrow B, A \rightarrow B \rightarrow C, and A \rightarrow B \rightarrow C \rightarrow D). On subsequent lists, the incremental method was gradually eliminated. New lists were introduced on which two or three items were present from the start of training (A \rightarrow B and A \rightarrow B \rightarrow C). On those lists, subjects had to use a trial and error strategy during the initial phase of training to determine the ordinal position of each item. Specifically, they had to remember if a response to a particular item at a particular position of the list allowed the trial to continue or if it ended the trial. Subsequent phases of training could be mastered by the strategy of responding to a newly added item last. However, on the remaining lists, on which all of the items were present from the start of training, subjects had to determine the ordinal position of all items by trial and error.

The number of trials each subject needed to satisfy the acquisition criterion on each four-item list is shown in Figure 2.² The left-hand panel shows the number of trials needed on lists trained with all phases of the incremental method. The middle panel shows the number of trials needed on lists trained with two and three phases, and the right-hand panel shows the number of trials needed on lists on which all four items were present from the start of training.

The decrease in the number of trials needed to master lists on which the incremental method was eliminated could reflect one or both of the following factors: the experience of learning Lists 1–8 and a shift to a trial-and-error strategy as the incremental method was eliminated. In an effort to assess the role of experience, a control group would be needed that learned as many lists by the incremental method as the subjects in this experiment learned without the incremental method. Because this experiment was not designed to evaluate the relative contribution of each factor, we focus on the development of a trial-and-error strategy to determine the ordinal position of new list items.

Lists on Which Two Items Were Displayed From the Start of Training

Only two–four sessions were needed to complete the A \rightarrow B phase of training on new lists. This was not surprising in view

¹ After mastering their second four-item list on which all items were present from the start of training, Rutherford and Franklin served in an experiment that evaluated their knowledge of the ordinal position of list items (Chen et al., 1997). On completion of that experiment, Rutherford and Franklin were trained on four new four-item lists on which all items were present from the start of training.

² Trials to criterion are complete trials, that is, trials that were terminated by a reward because the subject produced the entire sequence correctly, or trials that were terminated by a time out because the subject made an error. An incomplete trial is one that was terminated by the failure to respond to any of the items or to complete the required sequence in the allotted time.



Table 1
Summary of Training Phases

Subjects	Number of lists	Phases of training	Chance level of accuracy during first phase of training ^a
Dudley	8	A ^c	
Franklin	8	A → B	$P(A) = 1.0$
Rutherford	8	A → B → C	
Dudley	1 ^b	A → B → C → D	
Franklin	3	A → B ^d	$P(A) \times p(B A) = .5 \times 1.0 = 0.5$
Rutherford	1	A → B → C	
Dudley	5	A → B → C → D	$P(A) \times p(B A) \times p(C B) = 0.33 \times 0.5 \times 1.0 = .17$
Franklin	7	A → B → C ^e	
Rutherford	4	A → B → C → D	
Dudley	2		$P(A) \times p(B A) \times p(C B) \times p(D C) = 0.25 \times 0.33 \times 0.25 \times 1.0 = 0.04$
Franklin	6	A → B → C → D ^f	
Rutherford	6		

^a Chance accuracy was calculated under the conservative assumption that subjects made no backward errors and that repeat responses to a particular item were not errors. ^b The number of lists each subject learned during three-phase and two-phase training was determined by their performance during the first two sessions of training on each list. If the accuracy criterion was satisfied, then the length of the next list was increased by one item during the first session of training (see text for additional details). ^c Four phases. ^d Three phases. ^e Two phases. ^f One phase.

of that fact that subjects had a 50–50 chance of responding correctly to A. After choosing A, they could obtain reinforcement by responding to B by default. Rutherford and Dudley mastered the two-item phase on lists on which training began

with two items in two sessions, the minimum number of sessions needed to satisfy the acquisition criterion. Franklin missed satisfying that criterion on his first such list by one trial.

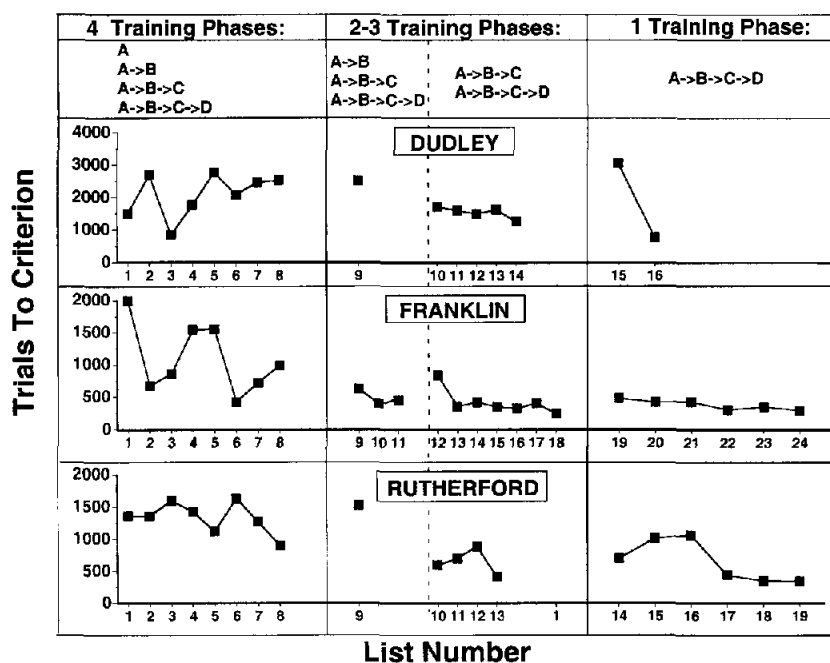


Figure 2. The number of trials needed to satisfy the accuracy criterion on lists trained by all phases of the incremental method (left panel), on lists on which the incremental method was gradually eliminated (middle panel) and on lists on which four items were displayed from the start of training (right panel). The data shown for Franklin and Rutherford for Lists 1–8 are from Swartz et al., 1991.

Lists on Which Three Items Were Displayed From the Start of Training

As on lists on which two items are presented from the start of training, subjects had to identify A by trial and error on lists on which training starts with three items. After identifying A, however, subjects could no longer use a default strategy as a basis for responding to B. Instead they had to determine the identity of B by trial and error. On lists on which three items are displayed from the start of training, subjects also had to learn not to make a backward error ($A \rightarrow B \rightarrow A$).

With each new list, subjects became more proficient at determining the ordinal position of each item. On the first list on which three items were displayed from the start of training, none of the subjects exceeded the chance level of accuracy during the first 40 trials (17%). On the last such list, each subject surpassed the 80% level of accuracy within the first 40 trials. The relevant data are shown in Figure 3. A repeated measure ANOVA of accuracy levels (List \times Block) during the first 120 trials of the first and last lists on which training began with three items yielded a significant effect of list, $F(1, 2) = 20.7$, $p < .05$, of block, $F(5, 10) = 4.18$, $p < .05$, but not of the interaction of those factors, $F(5, 10) = 1.92$, $p > .1$. Subjects' proficiency at mastering lists on which three items were displayed from the start of training was also evident in the steady decrease in the total number of trials needed to satisfy the accuracy criterion for each new list. This is shown in Figure 4. Linear regression coefficients fitted to each of the functions shown in Figure 4 were significant in each instance ($t = 14.2$, 7.5 , and 6.9 , and $p < .01$; for Rutherford, Franklin, and Dudley, respectively).

Subjects should be equally disposed to select A, B, or C as the first item on lists on which three items were displayed from the start of training. By contrast, a subject who was trained to learn the sequence $A \rightarrow B$ should select A reliably at the beginning of $A \rightarrow B \rightarrow C$ training. Both predictions were confirmed. The relevant

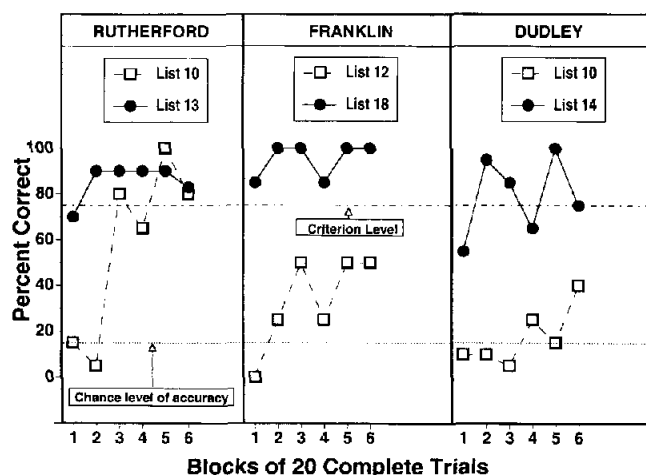


Figure 3. Accuracy of performance on lists on which three items were displayed from the start of training. The dashed functions show accuracy during the first 120 complete trials of the first list trained in this manner; the filled functions show accuracy on the last 120 complete trials of the last list trained in this manner.

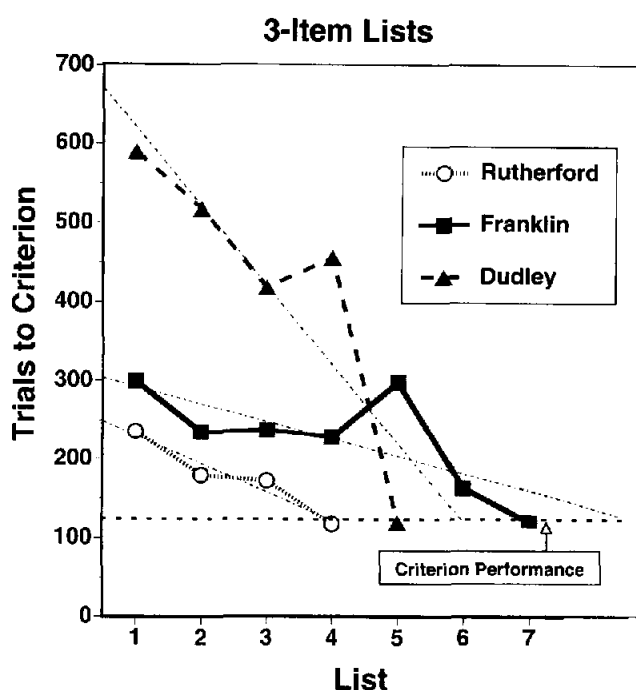


Figure 4. The number of trials needed to satisfy the accuracy criterion on the three-item phase of training on lists on which three items were displayed from the start of training.

data are shown in Table 2. On List 9 (on which training began with two items), each subject responded correctly to A during at least 87% of the first 30 trials of $A \rightarrow B \rightarrow C$ training (Table 2, upper panel). Only 1 subject (Rutherford) selected A at a greater than chance level (87%) during the first 30 trials on which three items were presented from the start of training (Table 2, bottom panel). The range of the other 2 subjects was 10%–17%. The selection of items B and C was also highly idiosyncratic. For example, Rutherford never chose B as the first item. Franklin and Dudley chose B and C, respectively, on 77% and 67% of the first 30 trials.

Forward errors were much more frequent than backward errors on lists on which training began with three items. This was true even when allowance was made for the fact that there were more

Table 2
Choice of Initial Item During First 30 Trials of Phase 3

Item	Rutherford	Franklin	Dudley
List 9 (Trained by successive phase method, starting with Phase 2)			
A	.97	.93	.87
B	.03	.07	.10
C	—	—	.03
First list on which three items are displayed from the start of training ^a			
A	.87	.10	.17
B	—	.77	.17
C	.13	.13	.67

^a Rutherford's list was List 10, Franklin's list was List 12, and Dudley's list was List 10.

opportunities to make a forward error (B, C, $A \rightarrow C$) than a backward error ($A \rightarrow B \rightarrow A$). The avoidance of backward errors is of interest in that it provides evidence that subjects were learning that lists had to be executed in a forward manner. Subjects lacked that knowledge while attempting to master their initial lists. On those lists, there was no difference in the relative frequencies of forward and backward errors (Swartz et al., 1991).

The relative frequencies of forward and backward errors are shown in Figure 5. A two-way repeated measure ANOVA (Error Type \times Stage of Training) of the relative frequency of forward and backward errors showed that there were more forward errors than backward errors, $F(1, 2) = 17.34$, $p < .05$, and that both types of errors decreased significantly at the end of training on each list, $F(1, 2) = 23.4$, $p < .05$. The interaction between these factors, which was also significant, $F(1, 2) = 15.23$, $p < .05$, appears to reflect a floor effect.

The low frequency of backward errors is also evident in the conditional probability of responding to C following responses to A and B. Figure 6 shows the conditional probabilities of a correct response at each step of the sequence, in blocks of three trials, during the first and the last lists on which training began with three items (left- and right-hand panels, respectively). The conditional probabilities of correct responses to individual items should not be confused with the probability of correctly completed trials (the product of the conditional probabilities). For example, even though the conditional probabilities of a correct response to A, B, and C by Rutherford during the first block of three sessions of training on List 1 were, respectively, .33, .80, and .50, he completed only 0.13 of the trials correctly.

The most striking feature of the functions in the left-hand panel of Figure 6 is the initially high value of $p(R_C | R_A \rightarrow R_B)$. The

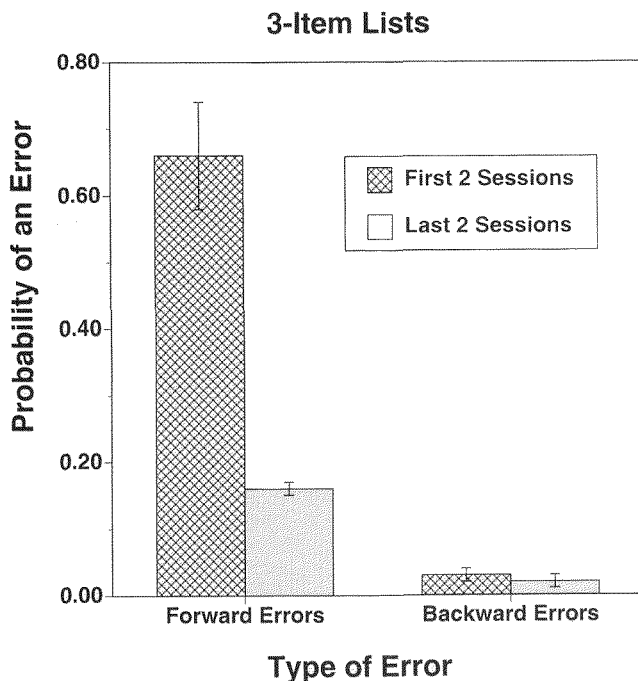


Figure 5. The probability of forward and backward errors during the three-item phase of training on lists on which training began with three items. See text for description of different types of errors.

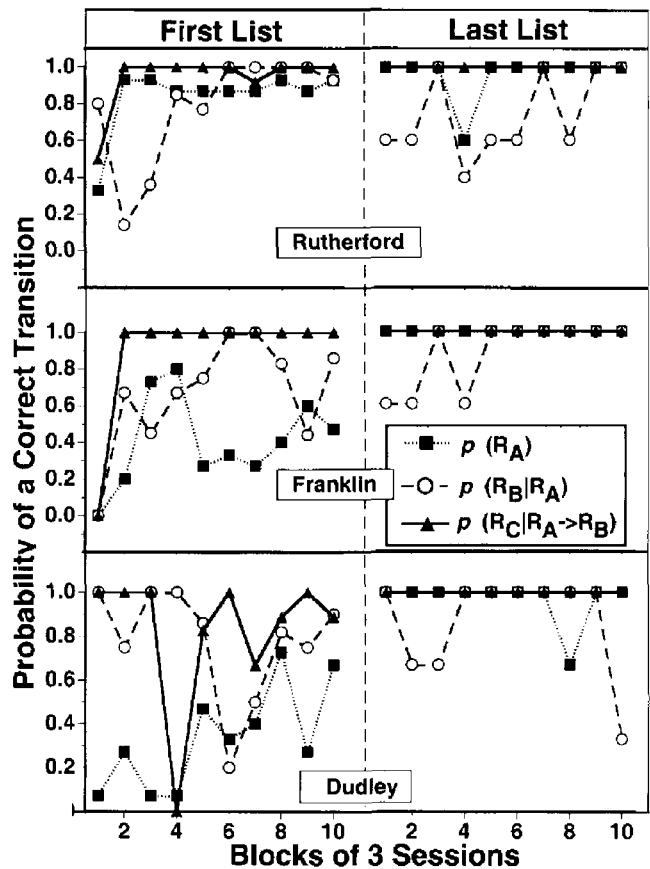


Figure 6. The probability of a correct response at each transition of an A, B, C sequence on the first (left-hand panel) and the last (right-hand panel) list on which training began with three items. On the first list, block size is 15 completed trials. On the second list, block size is three completed trials.

$p(R_C | R_A \rightarrow R_B)$ functions suggest that subjects responded to C by default, having responded correctly to A and B. Figure 6 also shows that Rutherford and Franklin never made a backward error after the seventh block of trials. Dudley's performance at the third step of the sequence was more variable but $p(R_C | R_A \rightarrow R_B)$ was higher than the other conditional probabilities. Two-way repeated measure ANOVAs (Transition \times Training) for each subject confirmed that there was a significant effect of transition [Rutherford: $F(2, 18) = 6.37$, $p < .01$; Franklin: $F(2, 18) = 13.94$, $p < .01$; Dudley: $F(2, 18) = 6.45$, $p < .01$], and that there was significant effect of training for Franklin and Dudley [Franklin: $F(1, 9) = 20.39$, $p < .001$; Dudley: $F(1, 9) = 1.49$, $p < .001$], but not for Rutherford [Rutherford: $F(1, 9) = 1.81$, $p > .2$]. The Transition \times Training interaction was significant for Franklin and Dudley [Franklin: $F(2, 18) = 14.33$, $p < .001$; Dudley: $F(2, 18) = 7.75$, $p < .01$] but not for Rutherford [$F(2, 27) = .73$, $p > .5$]. The interaction between those factors reflects a sharp increase in correct transitions during training on the first list.

The functions depicting the values of $p(R_A)$ in the right-hand panel of Figure 6 are of interest for two reasons. Of minor interest is that each subject guessed the identity of A on the first trial. The probability of that confluence of lucky guesses is 1/27. These

lucky guesses cannot be explained by the salience of Item A. The items used to compose the last list on which three items were presented from the start of training differed for each subject. Of greater interest is that, having guessed the identity of A, each subject understood the consequence of his lucky guess. Rutherford and Dudley made no errors at the first position of the sequence during the first 30 trials. Franklin made only one. On earlier lists on which subjects began training with three items, subjects perseverated in making errors at the first position, even after their first correct response to A. Retaining the consequences of a correct response to A on a three-item list is a small, but important step, in the development of serial expertise and a step that appeared for the first time after training on more than a dozen other lists.

Acquisition of Novel Lists by Trial and Error

During the final phase of the experiment, each subject was trained on four-item lists on which all of the items were present from the start of training. To master these lists, subjects had to determine the ordinal position of all of the items by trial and error. Each subject satisfied the accuracy criterion on the lists trained in this manner (Franklin and Rutherford, six lists; Dudley, two lists). The relevant data are shown in the right-hand panel of Figure 2.

A three-item list can be represented as two distinct end items and a single interior item. The first item is distinct because it is the item most closely associated with trial onset; the last item is distinct because it is most closely associated with reinforcement. On three-item lists, the middle item can therefore be identified by default. On four-item lists, the two interior items have to be differentiated from both the start item, the last item, and from each other. The difficulty of that task was immediately evident on lists on which training began with four items. Dudley made more errors in satisfying accuracy criterion on List 15 (his first four-item list on which all items were present from the start of training), than on any of his previous lists (cf. Figure 2). Two subjects (Rutherford and Dudley) failed to complete any trials correctly during their first session. Indeed, both subjects failed to exceed a zero level of accuracy during two of their next three sessions. Remarkably, both subjects continued to respond during each session. It is doubtful that a naive subject would show such resistance to extinction.

As on new three-item lists, there was no bias toward any item on lists on which four items were presented from the start of training. By contrast, the relative frequency of a correct response to A was high at the start of $A \rightarrow B \rightarrow C \rightarrow D$ training on lists trained with one or more phases of the incremental method. The relevant data are shown in Table 3. Each entry is the probability of choosing A, B, C, and D as the first item during the first 30 trials of $A \rightarrow B \rightarrow C \rightarrow D$ training. Lists trained by the incremental method include Lists 7 and 8 (the last two lists trained with all phases of the incremental method) and all lists on which training began with two- or three-items. On these lists, the probability of selecting A was .91. The entries in the bottom row of Table 3 are from the first 30 trials of lists on which four items were present from the start of training. The probability of selecting A, B, C, or D was close to .25, the value expected if the items were selected by chance.

Table 3

Probability of Selecting A, B, C, and D During First 30 Trials on Four-Item Lists Trained With and Without the Incremental Method

Incremental method	Item			
	A	B	C	D
With	.91	.06	.02	.01
Without	.32	.23	.18	.27

Forward Errors

Each subject made more forward errors during the first two sessions of training on lists on which four items were displayed from the start of training than during the first two sessions of $A \rightarrow B \rightarrow C \rightarrow D$ training on any of the previous lists. The relevant data are shown in Figure 7. At Position 1, forward errors were defined as responses to B, C, or D; at Position 2, they were defined as responses to $A \rightarrow C$ or $A \rightarrow D$ sequences; and at Position 3, they were defined as responses to $A \rightarrow B \rightarrow D$ sequences. A two-way repeated measure ANOVA (Training \times Position) of the probability of a forward error during the first two sessions of $A \rightarrow B \rightarrow C \rightarrow D$ training on lists trained with and without the incremental method yielded a significant effect of training, $F(1, 2) = 34.625$, $p < .05$, but not of position, $F(1, 2) = 1.73$, $p > .36$.

As noted previously, subjects responded to the last item at a high level of accuracy on lists on which three items were displayed from the start of training (cf. Figure 6). This tendency was also evident in Franklin's and Rutherford's performance on lists on which training began with four items, but not for Dudley. On List 15, the first list on which Dudley was trained with four items present from the start of training, Dudley made frequent backward errors. These added considerable variability to the transitional probability functions. When the frequency of backward errors is low, transitional probability functions provide a valid index of a subject's knowledge of the ordinal position of each item. For example, if the probability of the backward error $A \rightarrow B \rightarrow A$ is low, a high value of $p(A)$ and a low value of $p(B | A)$ would indicate that the subject had not differentiated Items C and D. If the probability of the backward error $A \rightarrow B \rightarrow A$ is high, then the trial ends without the subject having an opportunity to choose between Items C and D.

Because of the high frequency with which Dudley made backward errors on List 15, and because he was only trained on two lists on which four items were presented from the start of training, our analysis of transitional probabilities will focus on each subject's second list on which four items were introduced at the start of training. These are shown in Figure 8. For Rutherford and Franklin, transitional probabilities are shown for the first 12 blocks of training (20 trials per block); however for Dudley, transitional probabilities are shown for the first 25 blocks. As can be seen in the upper two panels, the values of $p(R_D | R_A \rightarrow R_B \rightarrow R_C)$ for Rutherford and Franklin were the highest of the four transitional probability functions. By contrast, the values of $p(R_D | R_A \rightarrow R_B \rightarrow R_C)$ for Dudley vacillated between zero and 1.0 through the

first 12 blocks.³ A value of zero for $p(R_D | R_A \rightarrow R_B \rightarrow R_C)$ means that a backward error occurred on all trials on which a correct response occurred to Items A, B, and C. One-way repeated measure ANOVAs of the conditional probabilities for the first 12 blocks of training were significant for each subject [Rutherford: $F(11, 3) = 15.74, p < .001$; Franklin: $F(11, 3) = 2.86, p < .05$; Dudley: $F(11, 3) = 3.09, p < .05$].

Discussion

The results of this experiment provide clear evidence that monkeys can develop expertise at using a trial-and-error strategy to learn arbitrary lists. Such expertise is not surprising when considered in the perspective of human cognition. A sizeable literature exists on the development of expertise for particular skills in humans, for example, solving logical and mathematical problems (Chi, Glaser, & Rees, 1983), playing chess (Simon & Chase, 1973), telegraphy (Bryan & Harter, 1897), and motor learning (Fitts, 1964). By contrast, the literature on animal expertise is virtually nil.

The paucity of research on animal expertise is particularly puzzling in light of recent research on the ability of animals to learn various cognitive tasks, for example, matching to sample (Wright, Santiago, & Sands, 1984), concept formation (Herrnstein, Loveland, & Cable, 1976; Wasserman & Bhatt, 1992), timing (Gibbon & Church, 1990), and serial learning (Sands & Wright, 1980; Terrace, 1987, 1999). In fact, the neglect of animal expertise antedates the recent revival of interest in animal cognition. Harlow (1949) identified two major impediments to the study of animal expertise. One was the practice of using naive subjects. The other was the practice of terminating an experiment once a subject reached an asymptotic level of performance. Harlow also noted a basic contradiction between the practice of using naive subjects to satisfy a single performance criterion and claims that the principles of animal behavior could be used to explain human behavior: "The behavior of the human being is not to be understood in terms of the results of single learning situations but rather in terms of changes which are affected through multiple, though comparable, learning

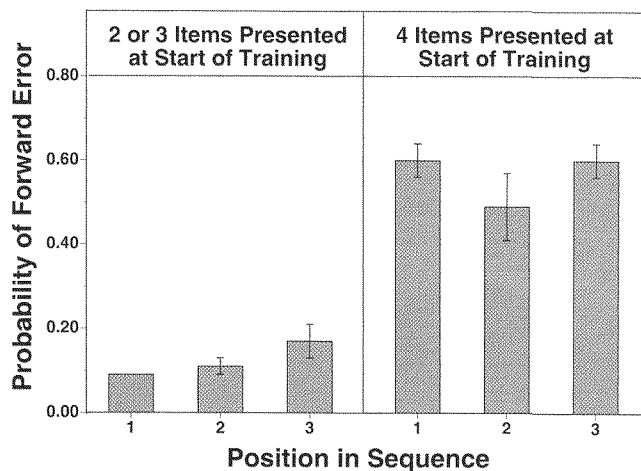


Figure 7. The probability of a forward error during Phase 4 on lists on which training began with 2, 3, and 4 items.

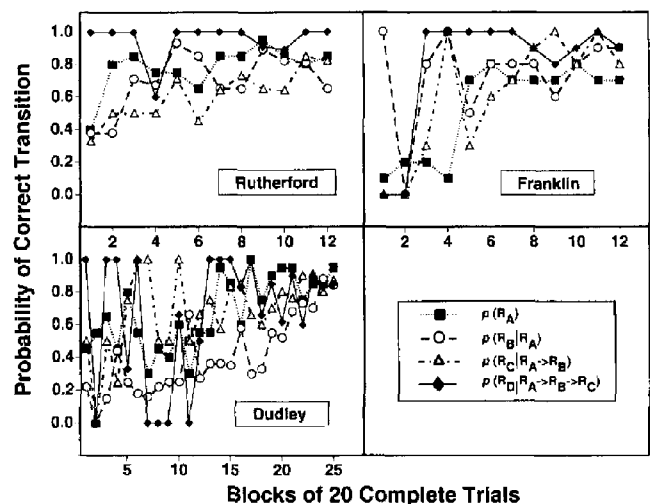


Figure 8. The probability of a correct response on complete trials at each transition of the second four-item list on which training began with four items. Block size is 20 completed trials.

problems" (Harlow, 1949, p. 51). The development of serial expertise in the present experiment shows that the same is true in the case of monkeys trained to learn arbitrary lists.

The Development of Serial Skills in Monkeys

It is not possible to specify which aspects of our subjects' list-learning experience were crucial to their ability to master a novel list on which all items were present from the start of training. We do not know, for example, whether it was necessary to train our subjects to master as many as eight novel four-item lists before eliminating the incremental method or whether the incremental method could have been eliminated in a single step (rather than in three steps, cf. Figure 2). To answer that type of question it would be necessary to perform control experiments that would systematically vary the training histories of different groups of subjects. It should, however, be noted that the interpretation of performance of human subjects on serial tasks poses the same problem.

With human subjects, it is especially difficult to separate the influence of prior training from that of the independent variables. The literature on human serial learning is based almost entirely on experiments performed on educated adults in which the typical subject has mastered countless pre-experimental lists by virtue of learning a language and acquiring a formal education (see Crowder, 1976, and Murdock, 1974, for reviews). Indeed, experiments on human serial learning take for granted the very serial skills that are needed to learn novel lists. The few experiments on serial learning that have been conducted with children also take such skills for granted (Flavell, 1970). At issue in those experiments is

³ One factor that contributed to the high degree of variability of Dudley's conditional probability functions was his tendency not to complete trials. Many trials on which Dudley responded correctly to A, B, and C were terminated by the time limit of each trial (20 s). Because the functions shown in Figure 9 are based only on completed trials, they underestimate Dudley's knowledge of the ordinal position of the first three items.

when verbal rehearsal strategies develop, as opposed to the more basic skills of learning to produce a set of arbitrary items in a particular sequence.⁴

Because it is impractical to conduct experiments on the necessary and sufficient conditions for the mastery of various kinds of serial skills, we opted to address the more tractable problem of identifying specific skills that contribute to the development of a monkey's serial expertise during the mastery of arbitrary lists. The results of this and previous experiments on serial learning by monkeys suggest three candidate skills: learning to eliminate backward errors, learning the start and end items, and encoding lists items with respect to their ordinal position.

Elimination of Backward Errors

One way to characterize the development of serial skills is to describe the systematic elimination of particular types of error. A monkey who starts training on a new four-item list with the knowledge that lists should be executed in a forward manner will avoid making any of the three possible types of backward errors ($A \rightarrow B \rightarrow A$, $A \rightarrow B \rightarrow C \rightarrow A$, and $A \rightarrow B \rightarrow C \rightarrow B$). By not making backwards errors, the probability of executing the list correctly by chance is increased by more than one order of magnitude (from $.25 \times .33 \times .33 \times .33 = .009$ to $.25 \times .33 \times .25 \times 1.0 = .04$).

Swartz et al. (1991) observed a reduction in the number of backward errors as monkeys learned 8 four-item arbitrary lists, particularly during the first few lists on which they were trained. Backward errors rarely occurred in the present experiment, with the exception of the first four-item list on which Dudley was trained with all of the items present at the start of training (his 15th four-item list). Dudley made substantially more backward errors on that list than on any previous list. It would appear that the challenge of determining the ordinal position of four new items resulted in a regression to habits that were eliminated on lists trained by the incremental method.

Unlike backward errors, which are avoidable, some forward errors are necessary to determine the ordinal position of items on a new list, for example, starting a trial with a response to B, C, or D to identify the first item of a new list, skipping from A to C and/or A to D to identify B, and making an $A \rightarrow B \rightarrow D$ error to identify C. As monkeys learn new lists, they should become more proficient at remembering the consequences of each type of forward error and at using that knowledge to avoid repeating particular forward errors. Although no evidence of that type of expertise was observed in the present experiment, a trend in that direction was observed in a recent experiment on the acquisition of seven-item lists by rhesus macaques (Terrace, Son, & Brannon, 2000).

In What Order Do Subjects Learn List Items?

Subjects have to learn that A is the first item before they can learn the identity of any of the other items. Having learned A, the question arises, do subjects learn that D is the last item and then differentiate B and C, or do they first differentiate B and C and then respond to D by default? Both hypotheses are plausible. Prior to learning four-item lists on which all items were present from the start of training, the subjects of the present experiment learned more than a dozen lists on which D was added incrementally, that

is, only after subjects learned to respond in the correct order to Items A, B, and C. That history could have biased subjects to respond to D by default. At the same time, D is a highly discriminable stimulus because a correct response to D produces reinforcement. Subjects could therefore learn to save D for last after executing a few correct sequences. Either strategy simplifies the problem of responding to A, B, C, and D on a new list in the correct order. Once subjects have learned not to make a backward error, the correct identification of D or B increases the probability of a correct response from 0.18 to 0.5.

The transitional probabilities shown in Figure 8 are consistent with both hypotheses. The transitions from A to B and from B to C gave rise to the largest number of errors. To argue that subjects learned D before they learned B or C, the probability of $A \rightarrow D$ errors would have to be substantially lower than the probability of $A \rightarrow C$ errors. Analyses of error distributions failed to reveal a significant difference in the probabilities of those errors. Accordingly, the results of the present study cannot adjudicate between the two explanations of how subjects learn four-item lists. The hypothesis that subjects learn to respond to the last item before differentiating the middle items could be tested by training monkeys to learn longer lists on which all items were present from the start of training. For example, there are 10 types of forward error a subject can make after responding to A on a six-item list ($A \rightarrow C$, $A \rightarrow D$, $A \rightarrow E$, $A \rightarrow F$, $A \rightarrow B \rightarrow D$, $A \rightarrow B \rightarrow E$, $A \rightarrow B \rightarrow F$, $A \rightarrow B \rightarrow C \rightarrow E$, $A \rightarrow B \rightarrow C \rightarrow F$, and $A \rightarrow B \rightarrow C \rightarrow D \rightarrow F$). To the extent that subjects learn to avoid making errors that include responses to the last item after responding correctly to early list items (e.g., $A \rightarrow F$, $A \rightarrow B \rightarrow F$, and $A \rightarrow B \rightarrow C \rightarrow F$), it could be argued that subjects have learned that F is the last item and that correct responses to F do not occur by default.

Representation of Ordinal Position on Simultaneous Chains

The literature on human serial learning provides evidence that subjects form two types of representation while memorizing a list: item-item and item-position associations (Crowder, 1976; Ebenholz, 1963, 1972; Young, Patterson, & Benson, 1963). Evidence of item-item and item-position associations after training on relatively few lists has also been reported in a variety of experiments on serial learning by monkeys. Evidence for item-item associations has been obtained from two-item subset tests that were administered after subjects mastered a four-item list (D'Amato & Colombo, 1988; Ohshima, 1997; Swartz et al., 1991). Evidence for item-position associations has been obtained from experiments on derived lists (Chen et al., 1997) on lists composed of two-item discrimination problems (Buffalo, Gaffan, & Murray, 1994) and lists on the ability of monkeys to fill gaps on a list with a "wild-card" (D'Amato & Colombo, 1989). Both types of associations contribute to the development of serial expertise.

Item-Item Associations

A subset test presents all possible pairs of items from an n -item list. There are, for example, 10 possible subsets that can be derived

⁴ An interesting exception is a recent experiment, showing that infants as young as 6 months have robust serial memories (Gulya, Rovee-Collier, Galluccio, & Wilk, 1998).

from a five-item list: AB, AC, AD, AE, BC, BD, BE, CD, CE, and DE. Monkeys respond to all such subsets at a uniformly high level of accuracy (D'Amato & Colombo, 1988). Of equal significance, the latencies of their responses to the first and second items of a given subset pair varies systematically with the location of those items on the original list.

An internal associative chain ($a \rightarrow b \rightarrow c \rightarrow d \rightarrow e$) can account for both aspects of a monkey's performance on a subset test. The monkey starts at the beginning of a linear representation of the list and moves through it until he locates the first item of the pair that is displayed on a particular trial. Having responded to the first item, he begins a similar search for the second item. Assuming a constant increment in search time for each item, the monkey's latency of responding to the first item of each subset should increase linearly as a function of the item's ordinal position on the original list. Similarly, latency of responding to the second item should increase linearly as a function of the number of items from the original list that intervene between the first and the second items of the subset pair. Precisely, such functions were obtained from cebus monkeys, who were trained to produce a five-item list (D'Amato & Colombo, 1988; McGonigle & Chalmers, 1996) and from rhesus monkeys, who were trained to produce four-item lists (Swartz et al., 1991). The subjects of the latter experiment were Franklin and Rutherford, who also served in this experiment.

Item-Position Associations

Evidence that monkeys learn item-position associations was obtained in a replication of a classic experiment on human participants concerning their knowledge of the ordinal position of list (Ebenholtz, 1963). In this instance, Franklin and Rutherford also served as subjects items (Chen et al., 1997).⁵ Both subjects were trained to produce four derived lists that were composed of items from the first four lists on which they were trained ($A1 \rightarrow B1 \rightarrow C1 \rightarrow D1$, $A2 \rightarrow B2 \rightarrow C2 \rightarrow D2$, $A3 \rightarrow B3 \rightarrow C3 \rightarrow D3$, and $A4 \rightarrow B4 \rightarrow C4 \rightarrow D4$). Derived lists contained one, and only one, item from each position of Lists 1-4. The ordinal position of each item was maintained on two of the derived lists ($A2 \rightarrow B4 \rightarrow C1 \rightarrow D3$ and $A3 \rightarrow B1 \rightarrow C4 \rightarrow D2$). On the other two derived lists, the original ordinal position of each item was changed ($B3 \rightarrow A1 \rightarrow D4 \rightarrow C2$ and $D1 \rightarrow C2 \rightarrow B3 \rightarrow A4$). All derived lists were trained with all items present from the start of training.

Each of the derived lists should be equally difficult if a monkey's knowledge of the four original lists was limited to item-item associations. If, however, a monkey acquired knowledge of the ordinal position of the items of the original lists, the two derived lists on which the ordinal position of each item was maintained should be acquired more rapidly than the two lists on which the ordinal position of each item was changed. Both subjects rapidly acquired the two derived lists on which the original ordinal position of each item was maintained. By contrast, they had difficulty learning the lists on which the ordinal position of each item was changed. These results show that monkeys acquire knowledge of an item's ordinal position while learning to execute a list.

Evidence that monkeys encode the first item of a list more readily than other items was provided by an experiment on lists composed of four two-choice discriminations (Buffalo et al., 1994). In that study, the list position of a particular pair determined whether a stimulus was positive or negative. For example, A was

rewarded when it was a member of the first pair of stimuli but not if it was a member of any other pair. Similarly, B was rewarded if it was a member of the second pair but not if it was presented at any other position. Subjects showed a marked primacy effect, indicating that they encoded A as a member of the first pair more strongly than they did items of other pairs.

Another study that provided evidence that monkeys acquire knowledge of the ordinal position of particular lists items tested subjects on a "wild card" test after they were trained to produce a simultaneous chain (D'Amato & Colombo, 1989). On the wild card test, one of the original list items was omitted on approximately half the trials. A novel stimulus, referred to as a wild card (W), was made available as a substitute. Which of the original list items was omitted on a particular trial was varied randomly. In effect, the wild-card procedure created as many new lists as there were items on the original list. For example, following training on the five-item list $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$, each of the following lists were presented on wild-card trials: $W \rightarrow B \rightarrow C \rightarrow D \rightarrow E$, $A \rightarrow W \rightarrow C \rightarrow D \rightarrow E$, $A \rightarrow B \rightarrow W \rightarrow D \rightarrow E$, $A \rightarrow B \rightarrow C \rightarrow W \rightarrow E$, and $A \rightarrow B \rightarrow C \rightarrow D \rightarrow W$.

The overall level of accuracy on wild-card trials exceeded that predicted by chance by a considerable margin on three-, four-, and five-item lists. Of greater importance was the uniformly high level of responding on each type of wild-card trial. The high level of accuracy on wild-card trials and the fact that there was no relationship between accuracy and wild-card trial type suggest that monkeys relied on their knowledge of the ordinal position of a missing item when responding to a wild card.

Implications for Future Research

This experiment showed that rhesus macaques can use a trial-and-error strategy to determine the ordinal position of items on four-item lists after training on fewer than two dozen lists. That ability could not be predicted from previous experiments on serial learning in which monkeys were trained by an incremental method that did not require a trial and error strategy. By contrast, the performance of Franklin, Rutherford and Dudley in the present experiment suggests that, with more list-learning experience, monkeys could master longer lists and determine the ordinal position of items on new lists with fewer errors. Their performance also suggests that it is only a matter of time before monkeys will be used in experiments that address the formidable array of questions about the nature of serially organized behavior that, for more than century, could only be addressed in human participants.

⁵ The derived list experiment was conducted after Franklin and Rutherford mastered the first two lists on which all items were presented from the start of training.

References

- Bryan, W. G., & Harter, N. (1897). Studies in the physiology and psychology of telegraphic language. *Psychological Review*, 4, 27-53.
- Buffalo, B., Gaffan, D., & Murray, E. A. (1994). A primacy effect in monkeys when list position is relevant. *Quarterly Journal of Experimental Psychology*, 47B, 353-369.
- Chen, S., Swartz, K., & Terrace, H. S. (1997). Knowledge of the ordinal

- position of list items in rhesus monkeys. *Psychological Science*, 8, 80–86.
- Chi, M., Glaser, R., & Rees, E. (1983). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence*. Hillsdale, NJ: Erlbaum.
- Chomsky, N. (1957). *Syntactical structures*. The Hague: Mouton Publishers.
- 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*, 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*, 15, 252–261.
- D'Amato, M. R., & Colombo, M. (1990). The symbolic distance effect in monkeys (*Cebus apella*). *Animal Learning and Behavior*, 18, 133–140.
- Ebenholtz, S. M. (1963). Serial learning: Position learning and sequential associations. *Journal of Experimental Psychology*, 66, 353–362.
- Ebenholtz, S. M. (1972). Serial learning and dimensional organization. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 267–314). New York: Academic Press.
- Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press.
- Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsett (Eds.), *Advances in child development and behavior* (pp. 181–211). New York: Academic Press.
- Gibbon, J., & Church, R. M. (1990). Representation of time. *Cognition*, 37, 23–54.
- Gulya, M., Rovee-Collier, C., Galluccio, L., & Wilk, A. (1998). Memory processing of a serial list by young infants. *Psychological Science*, 9, 303–307.
- Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, 56, 51–65.
- Herrnstein, R. J. (1985). Riddles of natural categorization. *Philosophical Transactions of the Royal Society of London*, 308B, 129–144.
- Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 2, 285–302.
- McGonigle, B., & Chalmers, M. (1996). The ontology of order. In L. Smith (Ed.), *Critical readings on Piaget* (pp. 279–311). New York: Routledge.
- Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.
- Ohshiba, N. (1997). Memorization of serial items by Japanese monkeys, a chimpanzee, and humans. *Japanese Psychological Research*, 39, 236–252.
- 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.
- Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, 61, 394–403.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Swartz, K. B., Chen, S., & Terrace, H. S. (1991). Serial learning by rhesus monkeys. I: Acquisition and retention of multiple four-item lists. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 396–410.
- 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. 115–138). Cambridge, MA: Ballinger.
- Terrace, H. S. (1987). Chunking by a pigeon in a serial learning task. *Nature*, 325, 149–151.
- Terrace, H. S. (1999). The comparative psychology of serially organized behavior. In S. Fountain, J. H. Danks, & M. K. McBeath (Eds.), *Biomedical implication of model systems of complex cognitive capacities*. New York: Sage.
- Terrace, H. S., Son, L., & Brannon, E. (2000). The development of serial expertise by rhesus macaques. *Nature*, (under review).
- Wasserman, E., & Bhatt, A. (1992). Conceptualization of natural and artificial stimuli by pigeons. In W. Honig & K. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 203–223). Hillsdale, NJ: Erlbaum.
- Wright, A. A., Santiago, H. C., & Sands, S. F. (1984). Monkey memory: Same/different concept learning, serial probe acquisition, and probe delay effects. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 513–529.
- Young, R. K., Patterson, J., & Benson, W. M. (1963). Backward serial learning. *Journal of Verbal Learning and Verbal Behavior*, 1, 335–338.

Received December 22, 1998

Revision received December 27, 1999

Accepted January 10, 2000 ■