Research Article

SERIAL EXPERTISE OF RHESUS MACAQUES

Herbert S. Terrace,[1] Lisa K. Son,[1] and Elizabeth M. Brannon[2]

[1]*Department of Psychology, Columbia University, and* [2]*Center for Cognitive Neuroscience, Duke University*

Abstract—*Here we describe the development of serial expertise in 4 experimentally naive rhesus monkeys that learned, by trial and error, the correct order in which to respond to 3-, 4-, and 7-item lists of arbitrarily selected photographs. The probabilities of guessing the correct sequence on 3-, 4-, and 7-item lists were, respectively, 1/6, 1/24, and 1/5,040. Each monkey became progressively more efficient at determining the correct order in which to respond on new lists. During subsequent testing, the subjects were presented with all possible pairs of the 28 items used to construct the four 7-item lists (excluding pairs of items that occupied the same ordinal position in different lists). Subjects responded to pairs from different lists in the correct order 91% of the time on the first trials on which these pairs were presented. These features of subjects' performance, which cannot be attributed to procedural memory, satisfy two criteria of declarative memory: rapid acquisition of new knowledge and flexible application of existing knowledge to a new problem.*

Expertise is the ability to solve novel instances of a particular kind of problem rapidly and efficiently (Chi, Glaser, & Farr, 1988). A large literature describes differences between the skills used by human experts and novices (e.g., mathematics and physics—Chi, Feltovitch, & Glaser, 1981; chess—Simon & Chase, 1973; serial memory—Chase & Ericsson, 1981; and medicine—Lesgold et al., 1988). By contrast, the literature on animal expertise is limited to a small number of experiments showing how monkeys (Harlow, 1949), rats (Eichenbaum, Fagan, & Cohen, 1986), and korvids (Hunter & Kamil, 1971) "learn to learn" a simple rule for discriminating two novel objects.

Although there is much evidence of cognitive abilities in animals (e.g., delayed matching to sample—Roberts & Grant, 1976; concept formation—Wasserman, Kiedinger, & Bhatt, 1988; short-term serial memory—Wright, Santiago, Sands, Kendrick, & Cook, 1985; long-term serial memory—Terrace, 2000; formation of cognitive maps—Olton, 1979; numerical discrimination—Brannon & Terrace, 1998; and timing—Gibbon & Church, 1990), the extent of an animal's expertise at those tasks (if any) has not been investigated. Expertise presupposes cognitive ability, but the converse does not follow. To demonstrate expertise, it is necessary to show that the efficiency of performing a cognitive task increases as new exemplars of that task are learned.

The paucity of evidence regarding animals' expertise limits comparisons of animal and human cognition to the earliest stages of skill development. The few experiments on animal expertise that have been performed are also limited by their focus on tasks that require subjects to make but a *single* response (as do most experiments on animal learning and memory). By contrast, experiments on human expertise typically involve complicated sequences and presuppose an ability to plan and execute novel sequences (Miller, Galanter, & Pribram, 1960).

Here we describe the first experimental evidence of serial expertise in an animal. First, we showed that monkeys became progressively more efficient at determining, by trial and error, the correct order in which to respond on seven 3-, eleven 4-, and four 7-item lists, each composed of different arbitrarily selected photographs. Subjects were then tested on their ability to apply the serial knowledge they had acquired while learning 7-item lists on a new task. In this test, the monkeys were shown item pairs that could be derived from the four 7-item lists on which they had been trained. These pairs consisted of items from the same list and items from different lists. Subjects responded in the correct order on the vast majority of the trials on which these pairs were presented for the first time, whether the items were drawn from the same list or from different lists. This showed that subjects could compare representations of the ordinal position of each item from any of the 7-item lists they had learned and then apply that knowledge to solve a novel problem without any additional training.

METHOD

Subjects and Apparatus

The subjects were 4 experimentally naive male monkeys (Benedict, Macduff, Oberon, and Rosencrantz; *Macaca mulatta*) who were born in 1995. Since 1997, they were housed in the monkey colony of the New York State Psychiatric Institute in accordance with National Institutes of Health guidelines. The subjects were neither food nor water deprived.

All training and testing was conducted in each monkey's home cage with the aid of a touch-sensitive 15-in. video monitor (Brannon & Terrace, 2000). List training began immediately after a subject learned to touch the video monitor to obtain a small banana-flavored pellet (190 mg). Lists were composed of digitized color photographs (each 1.5 in. × 1.5 in.) that were selected as distinctive exemplars of natural and man-made objects (e.g., animals, people, scenery, cars, bridges). The photographs could appear on the monitor in any of 16 equally spaced positions defined by a 4 × 4 matrix. Subjects initiated each trial by touching a "start stimulus" in the center of the video monitor. Reaction times (RTs) of all responses were recorded on each trial.

Training Subjects to Produce Simultaneous Chains

As illustrated in Figure 1, lists were trained as *simultaneous chains* (Terrace, 1984). All of the list items were displayed simultaneously throughout each trial, and their positions were varied randomly from trial to trial. The monkey's task was to respond to each item in a particular order (Fig. 1a), regardless of its spatial position (Fig. 1b). Variation of spatial position prevented subjects from performing the required sequence as a fixed motor pattern or as a discrete set of responses to specific external cues (e.g., the choice points of a maze).

Address correspondence to Herbert Terrace, Department of Psychology, Columbia University, 1190 Amsterdam Ave., New York, NY 10027; e-mail: terrace@columbia.edu.

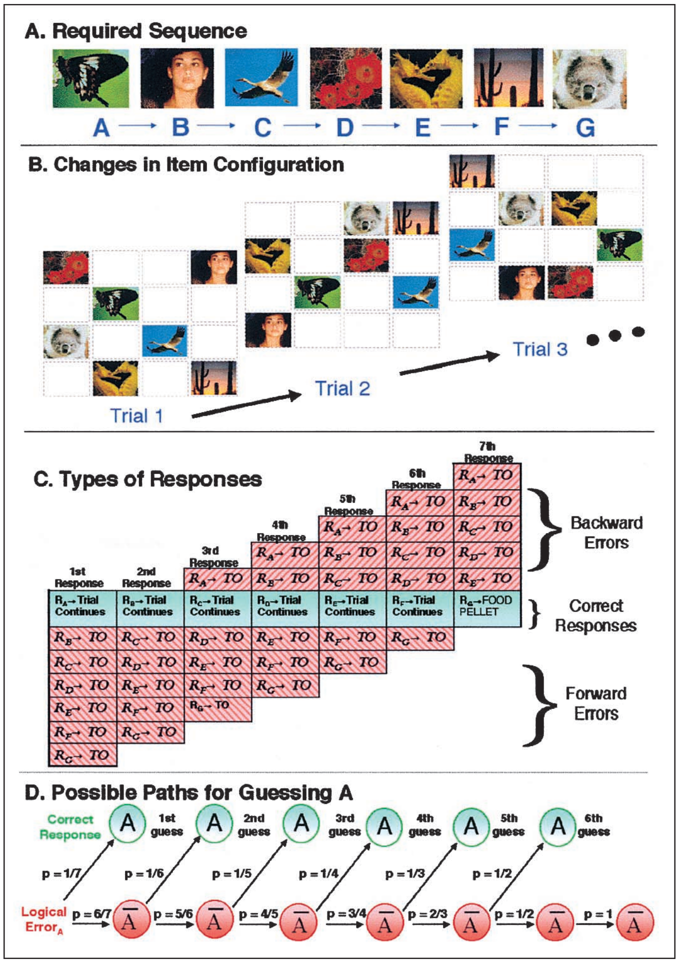

Fig. 1. The simultaneous chaining paradigm used in the study. An example of a 7-item list is shown in (a). During a trial, list items were presented on a touch-sensitive video monitor in a randomly configured display (b). The task was to touch the items in the prescribed order (A → B → C → D → E → F → G) irrespective of their positions on the monitor. The configuration on each trial was selected at random from the more than 5.8 million configurations that could be generated by presenting 7 items in any of 16 positions. The illustration in (c) shows the types of errors that subjects had to learn not to make in order to complete a trial correctly. See the text for additional details. Possible paths for determining the ordinal position of item A are shown in (d). The diagram shows the probabilities of guessing A at the start of the trial on the first, second, third, fourth, fifth, and sixth attempts by making logical errors to eliminate the remaining items. TO = time-out.

Instead, subjects had to represent their position in the sequence as they responded to successive items. The following thought experiment, which is based on a 7-item simultaneous chain, shows why.

Imagine trying to enter your seven-digit personal identification number (PIN)—say, *9-2-1-5-8-4-7*—at a cash machine on which the positions of the numbers were changed each time you tried to operate it. You could not enter your PIN 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 the button in the upper middle position to enter *2*, and so on. Instead, you would have to locate each number on the number pad and mentally keep track of your position in the sequence as you pressed different buttons. As difficult as this task may seem, it would be far more difficult to deduce your PIN by trial and error. Only correct responses would allow a trial to continue. Any error would terminate a trial and result in a new trial on which the digits were displayed in a different configuration. To determine your PIN, you would have to recall the consequences of the correct responses you made at each position and any of the different types of 21 logical errors you might have made while attempting to produce the required sequence (see the next paragraph). Further, you would have to determine the first six digits without getting any money from the cash machine. This is precisely the type of problem the monkeys had to solve at the start of training on each of the four 7-item lists on which they were trained. Instead of numerals, the monkeys had to respond to photographs. Instead of cash, they were given banana pellets.

The bottom half of Figure 1c shows the different types of *logical errors* a subject can make while determining the ordinal position of each item at the start of training on a new 7-item list. A logical error is the first incorrect guess a subject makes to a particular item at a given position of the list (e.g., a response to G at the second position). Although logical errors are necessary for discovering the ordinal position of an item, repetitions of the same error are not. Logical errors are made to obtain information by virtue of their consequences (G cannot be the 2nd item because the trial was terminated). *Repetitive errors* occur because the subject has forgotten the consequences of an earlier logical error.

The actual number of logical errors a subject will make while learning a new 7-item list can vary between 0 and 21. A subject who makes no logical guesses on the first trial of training on a new list guesses the ordinal position of each item correctly (an event that would occur, on average, once every 7!, or 5,040, trials). A subject who makes 21 logical errors responds to all incorrect items at each position. As shown in Figure 1c, a subject can make 6 different incorrect logical guesses at the first ordinal position before identifying A by default, 5 different incorrect logical guesses at the second position, 4 at the third, 3 at the fourth, 2 at the fifth, 1 at the sixth, and none at the seventh.

On average, the number of logical errors needed to guess the ordinal position of an item is the product of the sum of the number of different logical errors that can be made at a given position in the sequence and the probability of guessing the correct item at that position. For example, the sum of the number of different logical errors needed to determine A on a 7-item list is 21 (6 + 5 + 4 + 3 + 2 + 1), and the probability of selecting A with a correct guess is 1/7. Thus, the expected number of logical errors needed to determine A is 3 (21/7). Similarly, the expected number of logical errors needed to determine B is 2.5 (15/6), the expected number to determine C is 2 (10/5), and so on. The value of the expected number of logical errors at each position decreases linearly, 0.5 guesses at each position, until it reaches a value of 0 at item F.

As in the PIN example, any error in this experiment terminated the trial and initiated an 8-s time-out (TO) during which the video monitor was dark. Two types of error were distinguished: Responses that skipped one or more items were called *forward* errors; responses to an item to which a subject had responded previously were called *backward* errors. All instances of both types of error can be seen in Figure 1c. A correct response allowed the trial to continue and also produced brief visual and auditory feedback (0.3 s). Visual feedback was provided by a green border around the correctly selected item; auditory feedback by a 1000-Hz tone. A banana pellet was provided only after the subject responded in the correct order to each of the n list items. To learn the first $n - 1$ items of each list, subjects had to rely exclusively on the secondarily reinforcing consequences of correct and incorrect responses (brief feedback following a correct response and a TO following an error). For example, on a new 7-item list, food reward was provided only after subjects responded to items A, B C, D, E , F, and G in the correct order. Each subject was trained on 22 lists in the following order: seven 3-item lists (Lists 1–7), eleven 4-item lists (Lists 8–18), and four 7-item lists (Lists 19–22). All list items were presented from the start of training on each list. Each session consisted of 60 trials.

Criteria for Introducing New Lists

The goal of 3- and 4-item training was to develop a learning set for guessing the correct order in which to respond to a new set of list items by trial and error, rather than achieving a high level of accuracy on any particular list (Harlow, 1949). An exception was made for the first 3- and 4-item lists so that we could be sure that subjects learned to execute at least one 3- and one 4-item list at a high level of accuracy. Subjects were therefore trained until they completed 65% of the trials correctly in a single session on their first 3- and 4-item lists. In the case of the remaining 3- and 4-item lists, subjects were advanced to a new list after 3 days of training on a particular list or each time they completed at least 65% of the trials correctly during a daily session. Training on each of the four 7-item lists was continued until a subject completed 65% of the trials correctly during a single session.

Evaluating Knowledge of Ordinal Position of List Items

Previous experiments have shown that monkeys acquire knowledge of the ordinal position of list items following training on a single 5-item list (D'Amato & Colombo, 1988) or training on multiple lists containing 3 or 4 items (Chen, Swartz, & Terrace, 1997; Orlov, Yakovlev, Hochstein, & Zohary, 2000). In the present experiment, knowledge of ordinal position was evaluated by a 2-item-subset test that was derived from all 28 of the photographs used to construct the four 7-item lists.

Following the acquisition of their fourth 7-item list, subjects were retrained on each 7-item list to a criterion of completing 80% of the trials correctly during one session. They were then given a 2-item-subset test composed of all the items used in the four 7-item lists: 84 within-list subsets and the 252 between-list subsets. Within-list subsets were composed of items from a particular list (e.g., from List 3, the subsets $A_3B_3, A_3C_3, \ldots, A_3G_3; B_3C_3, B_3D_3, \ldots, F_3G_3$). Between-list subsets were composed of items drawn from different lists (e.g., the subsets A_2B_4 from Lists 2 and 4, C_3F_5 from Lists 3 and 5, E_1G_3 from Lists 1 and 3). For the purpose of analysis, these subsets could be divided into six types on the basis of the distance between their ordinal positions on the original lists (e.g., pairs of items separated by a distance of 1: the subsets $A_1B_1, B_2C_2, \ldots, A_1B_2; B_2C_3, C_3D_4, \ldots, F_3G_4$; a distance of 2: $A_1C_1, B_2D_2, \ldots, A_1C_2; B_2D_3, C_3E_4, \ldots, E_3G_4$; a distance of 6: $A_1G_1, A_2G_1, \ldots, A_1G_2, A_2G_2, A_2G_3, \ldots, A_4G_4$). Excluded from the subset test were the 21 pairs that could be composed of items occupying the same ordinal positions in the original lists (e.g., the subsets $A_1A_2, A_1A_3, \ldots, G_3G_4$). Within- and between-list pairs were interspersed randomly throughout the subset test. Subjects were rewarded if they responded in the order specified by the items' ordinal positions on the original list (or lists). For example, subjects were rewarded for responding to items C and F in that order, whether they came from the same or different lists (e.g., C_3F_3 or C_2F_6).

RESULTS

Accuracy Levels and Trials to Criterion

Compelling evidence of serial expertise was observed at each stage of training. Accuracy of responding increased progressively during the course of learning new 3-, 4-, and 7-item lists, as did subjects' efficiency in determining the ordinal position of items on each new list.

3- and 4-item lists

Figures 2a and 2b show the mean accuracy of responding on successive 3- and 4-item lists during the first and last sessions of training on each list. None of the subjects exceeded the chance level of accuracy (.17) during the first session of training on the first 3-item list. However, all subjects exceeded that level of accuracy during the first session of training on at least five of the six remaining 3-item lists. Benedict and Oberon exceeded the chance level of accuracy (.04) during the first session of training on all 4-item lists; Rosencrantz and Macduff, on 10 out of 11 of those lists. Subjects not only exceeded chance levels of accuracy during their first session of training on 3- and 4-item lists, but also responded at progressively higher levels of accuracy on successive lists. The slopes of the regression lines shown in Figures 2a and 2b for first-session accuracy during successive 3- and 4-item lists differed significantly from zero: $t(4) = 2.94$, $p < .05$, for 3-item lists; $t(9) = 3.16$, $p < .05$, for 4-item lists; the slope of the regression line for accuracy during the last session of training on 4-item lists also differed significantly from zero, $t(8) = 11.23$, $p < .05$.

7-item lists

The strongest evidence of serial expertise was subjects' mastery of 7-item lists (on which chance performance was 1/5,040). Each monkey learned all four of the 7-item lists on which it was trained and required progressively fewer sessions to satisfy the accuracy criterion on each new list. As shown in Figure 2c, subjects needed, on average, 31.5, 17.5, 13, and 12.25 sessions, respectively, to satisfy the accuracy criterion on their first, second, third, and fourth 7-item lists (ranges: 21–55, 11–25, 11–19, and 7–17, respectively). Similar reductions in the amount of training needed to satisfy an accuracy criterion have been observed in experiments in which adult human subjects were trained to memorize successive lists of arbitrarily selected words (Keppel, Postman, & Zavortink, 1968). Unlike the monkeys of this experiment, who learned only eighteen 3- and 4-item lists prior to their training on 7-item lists, the typical human subject learns thousands of lists before serving in an actual experiment.

Accuracy on Partially Completed Trials

The functions shown in Figure 2 underestimate subjects' serial knowledge because they are based entirely on correctly completed trials. The conditional probability of responding correctly at each position of a list is a more sensitive measure of serial knowledge because it provides credit for partially correct trials. In contrast to an overall measure of accuracy, which assigns a single value to the outcome of each trial, conditional probabilities assign an equal weight to each correct response on each trial.

Figure 3 shows three representative conditional probability functions, each of which provides additional evidence of subjects' serial

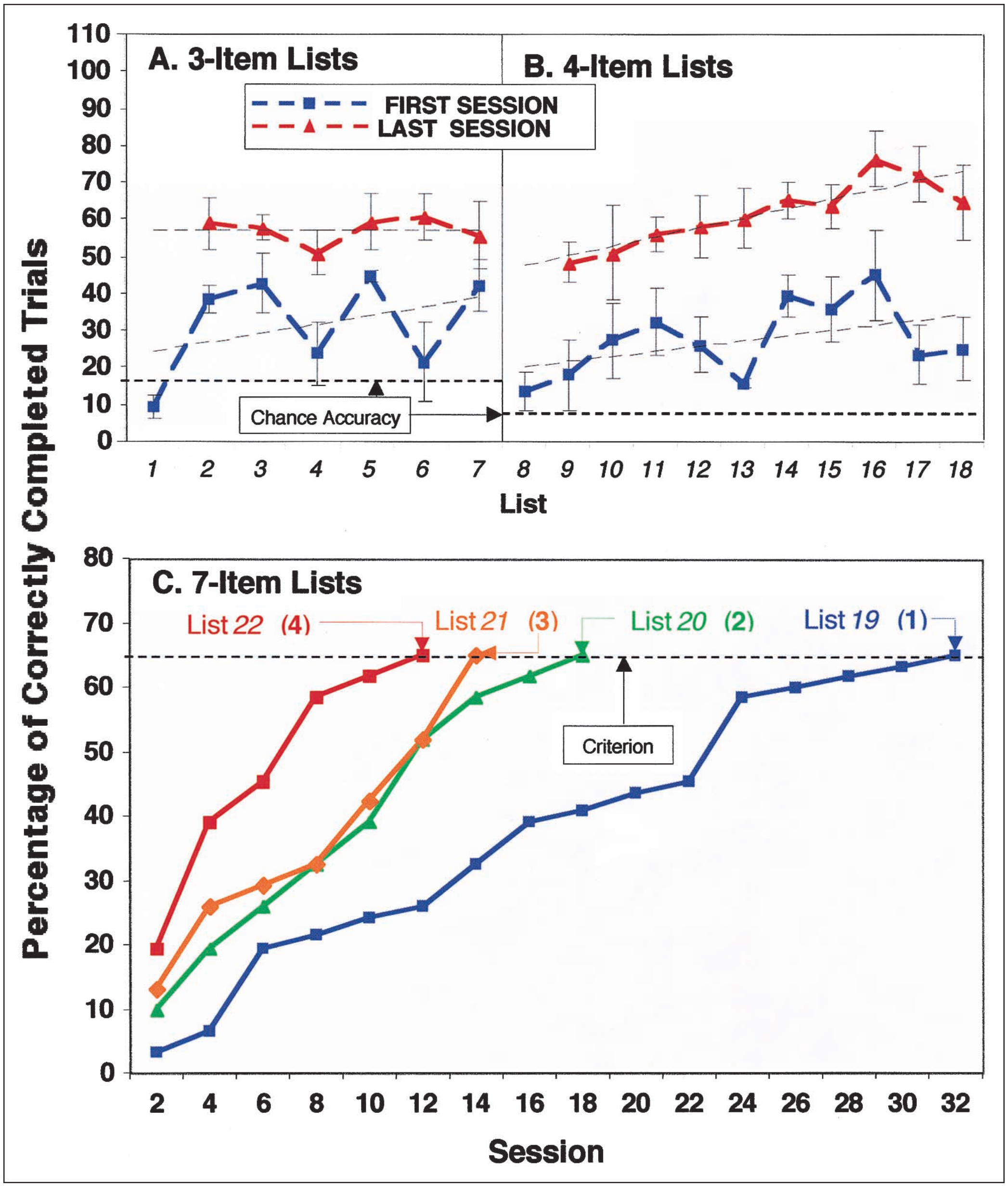

Fig. 2. Learning curves for the 3-, 4-, and 7-item lists. The percentage of correctly completed trials during the first and the last (lower and upper functions, respectively) sessions of training is shown for each 3-item (a) and 4-item (b) list. Data from the last session of training on the first 3- and 4-item lists are not included because those lists were trained to an accuracy criterion of 65% correctly completed trials. Also shown are regression lines. The mean accuracy of responding on each 7-item list during even-numbered sessions is shown in (c). The probability of executing a new 7-item list correctly by chance, assuming no backward errors, is 1/5,040 (1/7!). Note that the abscissa in (c) is Session (not List).

expertise. The functions are based on subjects' performance during the first session of training on three new lists: the last 4-item list and the first and the last 7-item lists. The probability of responding to all of the items correctly on a particular trial is the product of the conditional probabilities of responding correctly to each item. For example, the conditional probabilities of a correct response to A, B, C, D, E, F, and G during the first session of training on the first 7-item list were, respectively, .79, .62, .49, .46, .32, .27, and .29. However, the relative frequency of correctly completed trials was only .003.

A comparison of subjects' accuracy on each of those lists revealed a clear primacy effect. Of greater significance were the high absolute levels of accuracy at positions A and B at the start of training on the last 7-item list. A comparison of those levels with the levels that would be attained by an "ideal list learner" showed that subjects identified positions A and B with close to the minimum number of logical errors. An ideal list learner remembers the consequences of each error at a particular position and does not repeat that error while learning a new list. If an ideal list learner does not guess an item's ordinal position correctly with its first response to that item (see Fig. 1d), its best strategy would be to make logical errors until it encounters the correct item. The extent to which the number of logical errors a subject makes approximates the number an ideal list learner makes is a measure of the subject's expertise at learning new lists.

The light dotted line at the top of Figure 3 shows the expected conditional probabilities of an ideal list learner. On a 7-item list, an ideal list learner would need, on average, 3 logical errors to identify A, 2.5 logical errors to identify B, 2 logical errors to identify C, and so on. Given 60 trials/session, and an average of 3 logical errors at A, the conditional probability of a correct response by an ideal ordinal-position detector at position A would be .95 (57/60); given an average of 2.5 logical errors at position B, the conditional probability of a correct response by such a list learner at position B would be .96 (54.5/57). The serial expertise of subjects in this experiment compared favorably with that of an ideal list learner at positions A and B at the start of training on their final 7-item list. On average, subjects made only 4 errors at position A (accuracy = .92) and 5 errors at position B (accuracy = .88) during the first session of training on that list.

The functions shown in Figure 3 also provide evidence of a recency effect and suggest two factors that contributed to its development: the salience of the *n*th item (because it was followed by a reward) and the subject's working memory of the first $n - 1$ items to which it responded on a particular trial. To execute the entire sequence correctly, the subject had to retain in working memory each item to which it responded in order to avoid returning to any of those items (a backward error). On this view, a recency effect would not be expected to occur until the subject executed a new list correctly.

All subjects learned to execute new 4-item lists correctly during the session in which such lists were introduced. In each instance, a recency effect was observed during the first session of training. Only 2 of the 4 subjects completed any trial correctly during the first session of training on the first 7-item list. As can be seen in Figure 3, there was no recency effect during that session. A recency effect was observed during the first session of training on subsequent 7-item lists, once subjects were able to complete a trial correctly.

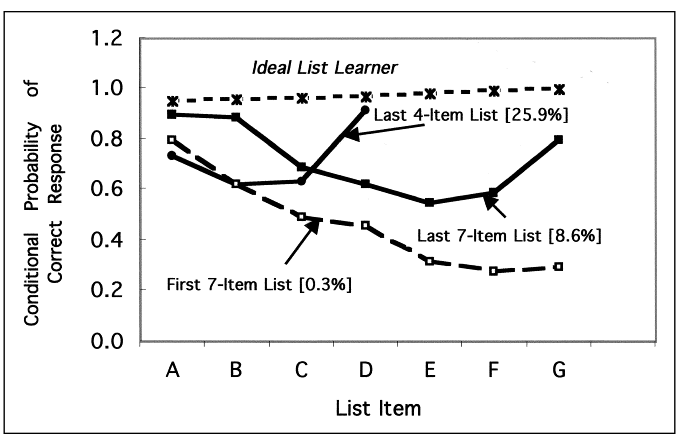

Fig. 3. Serial position functions: average conditional probability of a correct response at each position during initial training on the last 4-item list (List 18) and the first and last 7-item lists (Lists 19 and 22). For each list, the percentage in the brackets indicates the mean total accuracy of the entire list: the product of the conditional probabilities at each position. The dotted line at the top of the graph shows the expected conditional probabilities of an ideal list learner, who does not repeat errors while learning a new list.

Knowledge of Ordinal Position

On average, subjects responded correctly on the first presentation of each subset type on 94% of the within-list subset trials (range across subjects: 92–96%) and on 91% of the between-list subset trials (range across subjects: 89–96%).[1] Indeed, accuracy never fell below 70% for any subset. These data add significantly to current understanding of the monkey's long-term and working memory systems. They show that each monkey represented, in long-term memory, the ordinal positions of items from each of the four 7-item lists learned and that the monkey could then compare, in working memory, the ordinal positions of any 2 items from any of the 7-item lists, without any training to make such comparisons.

Distance and Magnitude Effects

Two other features of subjects' performance on subset tests have been referred to in the human literature as *distance* and *magnitude* effects (Banks, 1977; Holyoak & Patterson, 1981; Moyer & Landauer, 1967). A distance effect is an improvement in performance that results from an increase in the ordinal distance between items. A magnitude effect is a decrement in performance as the ordinal position of the first item is advanced. The accuracy functions in the upper portion of Figure 4 provide evidence of a distance effect; the RTs in the lower panel provide evidence of a magnitude effect.

The distance between the items of a given subset pair is defined as the difference between the ordinal position numbers of those items on the lists from which they were drawn. Thus, all subsets containing the items C and D (e.g., C_4D_3, C_2D_1, C_1D_1) are separated by a distance of 1, all subsets containing the items B and D (e.g., B_2D_2, B_3D_1) are separated by a distance of 2, and so on. Associative models of serial learning predict a decrease in the accuracy of responding to subset pairs as the distance between items increases because associative strength between items decreases with distance. Spatial models of serial learning (e.g., Holyoak & Patterson, 1981) predict an increase in accuracy as the distance between items increases because larger distances are more discriminable than smaller distances.

Even though accuracy of responding was nearly perfect for many subset types, there is clear evidence of a distance effect in the accuracy functions in Figure 4 (shown in blue). The mean level of accuracy was 83% at distance 1 and 92% at distance 2. For distances greater than 2,

1. Accuracy on subsets composed of nonadjacent items that contained no end items (BD, BE, BF, CE, CF, and DF) did not differ from accuracy on subsets containing an end item (A or G) or subsets composed of two adjacent items. The mean accuracy on subsets composed of nonadjacent items that contained no end items was 92%; mean accuracy on subsets containing an end item was 90%.

the average level of accuracy was 99% for all subsets. Were it not for this ceiling effect, it is likely that distance would have exerted a stronger influence on accuracy. A one-way analysis of variance (ANOVA) yielded a significant effect of distance on accuracy, $F(5, 35) = 19.21$, $p < .01$. Scheffé tests showed that distance 1 differed from distances 2 through 5, and that distance 2 differed from distances 3 and 5 ($p < .05$ in each instance), but that the distances 3 through 5 did not differ among themselves.

RTs of correct responses to the first item on subset tests increased with the magnitude of the first item at distance 1 but not at larger distances. The red functions in the lower panel of Figure 4 show the mean RTs to the first item of each subset type at distances 1 through 6. A one-way ANOVA yielded a significant effect of magnitude at distance 1, $F(5, 423) = 3.65$, $p < .01$. This ANOVA was based on all RTs less than or equal to 6,000 ms (96% of RTs). Analogous RT functions have been obtained from human subjects in experiments on the discriminability of immediately adjacent list items (e.g., letters of the alphabet such as *cd*, *kl*, or *rs*, as opposed to *dk*, *hs*, or *nz*; Hamilton & Sanford, 1978) and abstract geometric items from 5-item simultaneous chains (AB, BC, CD, EF, and FG, as opposed to AF, CG, etc.; Colombo & Frost, 2001).

The distance and magnitude effects shown in Figure 4 suggest two processes for deciding which item of a subset pair came first. At distance 1, it is difficult to perceive the ordinal position of each item directly. Accordingly, subjects may have used an iterative process to compare each item with a representation of the first item on its original list. The more advanced the position of the first item of the subset, the longer it would take to determine its ordinal position. For distances greater than 1, that decision was equally easy at all positions, and hence equally rapid.

DISCUSSION

The sequences that Rosencrantz, Macduff, Benedict, and Oberon learned are by far the most difficult lists that have been mastered by nonhuman primates, including those trained in experiments on the linguistic and numerical abilities of apes (Matsuzawa, 1985; Premack, 1976; Rumbaugh, 1977) and in recent experiments on sequence production (Colombo, Eickhoff, & Gross, 1993) and sequence recognition (Carpenter, Georgopolous, & Pellizzer, 1999; Orlov et al., 2000) by rhesus macaques. It is doubtful, however, that the sequential skills described here reflect the upper limit of monkeys' serial expertise. The ease with which the subjects learned 7-item lists and the steady decrease in the number of sessions they needed to master new lists suggest that monkeys could learn such lists even more rapidly and that they could also master longer lists. Indeed, at the start of training on their fourth 7-item list, subjects identified the first two list items as rapidly as was logically possible. Subjects' serial expertise was hardly limited to the efficient learning of new 7-item lists. They were also able to apply their knowledge of 7-item lists to determine the correct order in which to respond to within- and between-list subsets on the first occasion on which each type of subset was presented.

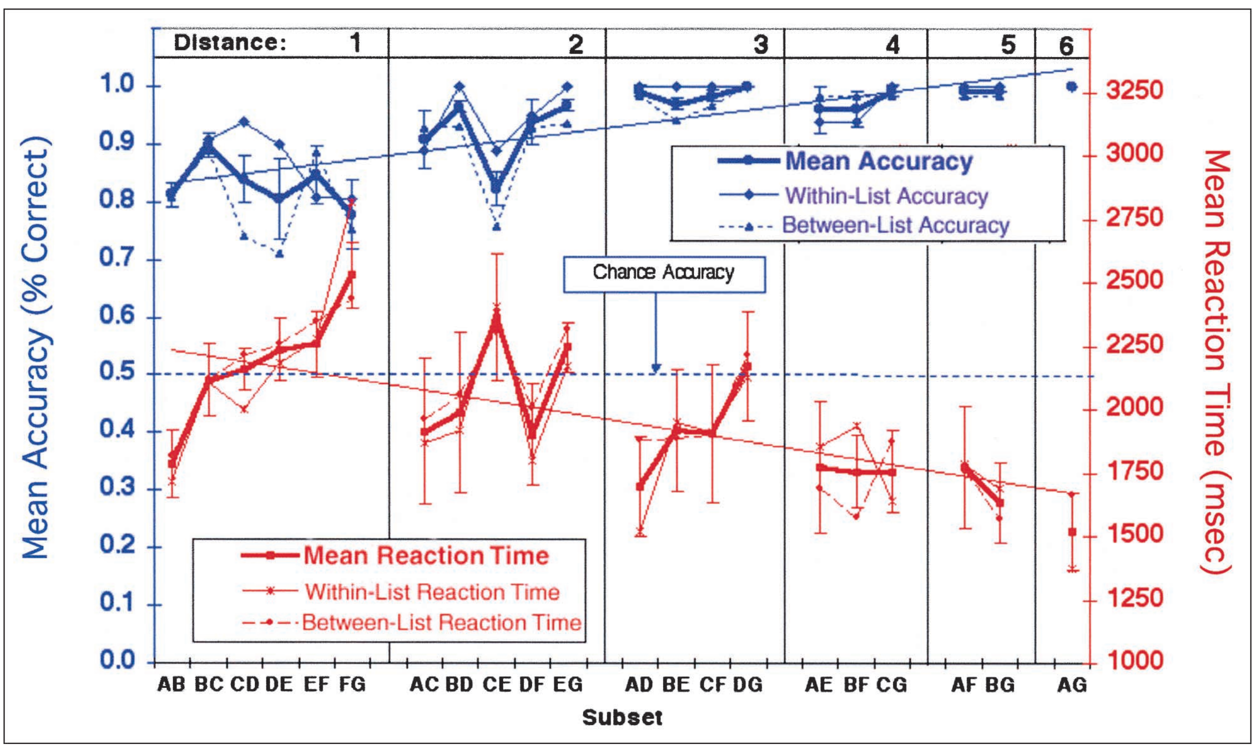

Fig. 4. Accuracy and reaction times on the subset test. The blue functions in the upper panel show mean accuracy to each type of within- and between-list subset at distances 1 through 6. The red functions in the lower panel show the mean reaction time to each type of subset at distances 1 through 6. The entries on the abscissa are generic in that they refer to the types of within- and between-list subsets that are represented at each position. For example, AB refers to A_1B_1, A_2B_2, A_3B_3, and A_4B_4 in the case of within-list subsets and to A_1B_2, A_1B_3, . . . , A_4B_1 subsets in the case of between-list subsets.

What kind of knowledge does an expert monkey use? Philosophers have distinguished between two types of human knowledge: "knowing how" to execute motor skills (e.g., riding a bicycle) and "knowing that" certain facts are true (e.g., that an elephant is bigger than a dog and that a dog is bigger than a fly; Ryle, 1949). Cognitive psychologists have drawn an analogous distinction in their definitions of procedural and declarative knowledge (Squire, 1994). Procedural knowledge, which is characterized as inflexible and unconscious, is acquired slowly through repetitive training on a particular problem and is not mediated by the hippocampus. Declarative knowledge, which is acquired rapidly, is flexible, is expressed consciously to declare particular facts, and is mediated by the hippocampus (Tulving & Schacter, 1990).

Procedural knowledge cannot explain the serial expertise of the monkeys in this experiment because (a) a different motor sequence was required on every trial, (b) serial expertise increased as new lists were introduced, and (c) the monkeys were able to use previously acquired serial knowledge to perform a new procedure correctly on the first trial on which they were tested.

Does it follow that the subjects of this experiment relied on declarative knowledge to learn the correct order in which to respond on new lists or to respond correctly to novel within- and between-list pairs on the subset test? The answer is no if the criteria for declarative knowledge include the ability to declare that knowledge. But given the mismatch between the breadth of subjects' serial expertise and the definition of procedural knowledge, it seems reasonable to broaden the definition of declarative knowledge to include nonverbal as well as verbal knowledge (e.g., Eichenbaum, 1999).

From an evolutionary perspective, definitions of declarative knowledge that are restricted to linguistic propositions make little sense. A considerable body of evidence suggests that declarative knowledge evolved before language (Bickerton, 1995; Reber, 1993). For example, a monkey's ability to judge the relative social rank of other monkeys in their living groups requires declarative knowledge (Harcourt & de Waal, 1992). Such judgments presuppose the ability to represent the ordinal position of other monkeys, in different combinations, in a manner similar to that observed on the subset tests administered in this experiment. Another nonverbal precursor of declarative knowledge is suggested by an animal's ability to represent spatial information and to use that information to solve problems (Gallistel, 1992). By encoding an item's ordinal position spatially during list learning, a monkey could compare the ordinal positions of items from various lists with respect to preexisting spatial coordinates. That hypothesis is supported by the RT data obtained in the present experiment for within- and between-list subset tests.

One candidate for a nonverbal component of declarative knowledge is a knowledge system in which information is encoded analogically as images (Kosslyn, 1980; Lashley, 1951). Analogical theories of human cognition have been criticized on the grounds that images convey less information than propositional representations (Pylyshyn, 1981). That criticism does not apply to monkeys. Indeed, it is the utter irrelevance of propositional representations in monkeys that is the strongest argument for analogical representations as the basis of their serial expertise.

Acknowledgments—This research was supported by National Institute of Mental Health (NIMH) Grant MH40462 to H.S. Terrace, NIMH Grant MH50884 to H. Lisanby and H.S. Terrace, and a National Science Foundation graduate fellowship to E.M. Brannon. We wish to thank Jennifer Mangels, Janet Metcalfe, Michele Miozzo, Harold Sackeim, Bennett Schwartz, and Eugene Winograd for valuable comments and Eric Anderson and Shaofu Chen for help in conducting the research.

REFERENCES

Banks, W.P. (1977). Encoding and processing of symbolic information in comparative judgments. *The Psychology of Learning and Motivation*, *11*, 101–159.

Bickerton, D. (1995). *Language and human behavior.* Seattle: University of Washington Press.

Brannon, E.M., & Terrace, H.S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science*, *282*, 746–749.

Brannon, E.M., & Terrace, H.S. (2000). Representation of the numerosities 1–9 by rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*, 31–49.

Carpenter, A.F., Georgopolous, A.P., & Pellizzer, G. (1999). Motor control encoding of serial position in a context-recall task. *Science*, *283*, 1752–1757.

Chase, W.G., & Ericsson, K.A. (1981). Skilled memory. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–191). Hillsdale, NJ: Erlbaum.

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.T.H., Feltovitch, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152.

Chi, M.T.H., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Colombo, M., Eickhoff, A.E., & Gross, C.G. (1993). The effects of inferior temporal and dorsolateral frontal lesions on serial-order behavior and visual imagery in monkeys. *Cognitive Brain Research*, *1*, 211–217.

Colombo, M., & Frost, N. (2001). Representation of serial order in humans: A comparison to the findings with monkeys (*Cebus appela*). *Psychonomic Bulletin & Review*, *8*, 262–269.

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

Eichenbaum, H. (1999). The hippocampus and mechanisms of declarative memory. *Behavioural Brain Research*, *103*, 123–133.

Eichenbaum, H., Fagan, A., & Cohen, N.J. (1986). Normal olfactory discrimination learning set and facilitation of reversal after medial-temporal damage in rats: Implications for an account of preserved learning abilities in amnesia. *Journal of Neuroscience*, *6*, 1876–1884.

Gallistel, C.R. (1992). *Animal cognition*. Cambridge, MA: MIT Press.

Gibbon, J., & Church, R.M. (1990). Representation of time. *Cognition*, *37*, 23–54.

Hamilton, J.M.E., & Sanford, A.J. (1978). The symbolic distance effect for alphabetic order judgments: A subjective report and reaction time analysis. *Quarterly Journal of Experimental Psychology*, *30*, 33–43.

Harcourt, A.H., & de Waal, F.B.M. (1992). Coalitions and alliances: Are primates more complex than non-primates? In A.H. Harcourt & F.B.M. de Waal (Eds.), *Coalitions and alliances in humans and other animals* (pp. 444–469). New York: Oxford University Press.

Harlow, H.F. (1949). The formation of learning sets. *Psychological Review*, *56*, 51–65.

Holyoak, K.J., & Patterson, K.K. (1981). A positional discriminability model of linear-order judgments. *Journal of Experimental Psychology*, *7*, 1283–1302.

Hunter, M.W., III, & Kamil, A.C. (1971). Object discrimination learning set and hypothesis behavior in the northern blue jay (*Cyanocitta cristata*). *Psychonomic Science*, *22*, 271–273.

Keppel, G., Postman, L., & Zavortink, B. (1968). Studies of learning to learn: VIII. The influence of massive amounts of training upon the learning and retention of paired-associate lists. *Journal of Verbal Learning and Verbal Behavior*, *7*, 790–796.

Kosslyn, S. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

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: John Wiley & Sons.

Lesgold, A., Robinson, H., Feltovitch, P., Glaser, R., Klopfer, D., & Wang, Y. (Eds.). (1988). *Expertise in a complex skill: Diagnosing X-ray pictures*. Hillsdale, NJ: Erlbaum.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature*, *315*, 57–59.

Miller, G.A., Galanter, E., & Pribram, K.H. (1960). *Plans and the structure of behavior.* New York: Holt, Rinehart and Winston.

Moyer, R.S., & Landauer, T.K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519–1520.

Olton, D.S. (1979). Mazes, maps, and memory. *American Psychologist*, *34*, 583–596.

Orlov, T., Yakovlev, B., Hochstein, S., & Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. *Nature*, *404*, 77–80.

Premack, D. (1976). *Intelligence in ape and man*. Hillsdale, NJ: Erlbaum.

Pylyshyn, Z. (1981). The imagery debate: Analogue versus tacit knowledge. *Psychological Review*, *88*, 16–45.

Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Roberts, W.A., & Grant, D.S. (1976). Studies of short-term memory in the pigeon using the delayed matching-to-sample procedure. In D.L. Medin, W.A. Roberts, & R.T. Davis (Eds.), *Processes of animal memory* (pp. 79–98). Hillsdale, NJ: Erlbaum.

Rumbaugh, D.M. (1977). *Language learning by a chimpanzee: The Lana project.* New York: Academic Press.

Ryle, G. (1949). *The concept of mind*. New York: Barnes & Noble.

Simon, H.A., & Chase, W.G. (1973). Skill in chess. *American Scientist*, *61*, 394–403.

Squire, L.R. (1994). Declarative and nondeclarative memory. In D.L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 204–231). Cambridge, MA: MIT Press.

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. (2000). The comparative psychology of serially organized behavior. In R. Cook (Ed.), *Avian cognition* (chap. 19). Medford, MA: Comparative Cognition Press. Available www.pigeon.psy.tufts.edu/avc/

Tulving, E., & Schacter, D.L. (1990). Priming and human memory systems. *Science*, *247*, 301–306.

Wasserman, E.A., Kiedinger, R.E., & Bhatt, R.S. (1988). Conceptual behavior in pigeons: Categories, subcategories, and pseudocategories. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 235–246.

Wright, A.A., Santiago, H.C., Sands, S.F., Kendrick, D.F., & Cook, R.G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science*, *229*, 287–289.

(Received 11/12/01; Revision accepted 3/12/02)