Monkey Memory: *Same/Different* Concept Learning, Serial Probe Acquisition, and Probe Delay Effects

Anthony A. Wright and Héctor C. Santiago
University of Texas Health Science Center at Houston
Graduate School of Biomedical Sciences

Three rhesus monkeys were trained and tested in a *same/different* task with six successive sets of 70 item pairs to an 88% accuracy on each set. Their poor initial transfer performance (55% correct) with novel stimuli improved dramatically to 85% correct following daily item changes in the training stimuli. They acquired a serial-probe-recognition (SPR) task with variable (1–6) item list lengths. This SPR acquisition, although gradual, was more rapid for the monkeys than for pigeons similarly trained. Testing with a fixed list length of four items at different delays between the last list item and the probe test item revealed changes in the serial-position function: a recency effect (last items remembered well) for 0-s delay, recency and primacy effects (first and last list items remembered well) for 1-, 2-, and 10-s delays, and only a primacy effect for the longest 30-s delay. These results are compared with similar ones from pigeons and are discussed in relation to theories of memory processing.

The primacy effect, good memory of the first list items, has been recently demonstrated for two monkey species (Roberts & Kraemer, 1981; Sands & Wright, 1980a, 1980b) and a chimpanzee (Buchanan, Gill, & Braggio, 1981). The finding of animal primacy effects, as well as the more common recency effects (remembering the last list items well) is important because it demonstrates the continuity between human and nonhuman primate memory processing. Prior to the discovery of a nonhuman primate primacy effect, serial-position functions for monkeys revealed only recency effects (Davis & Fitts, 1976; Devine & Jones, 1975; Gaffan, 1977). Recency effects typically show a general monotonic decay as the retention interval is lengthened in much the same way that memory for single items shows monotonic decay with retention interval (Cox & D'Amato, 1982; D'Amato, 1973; Medin, Reynolds, & Parkinson, 1980; Overman & Doty, 1980). Indeed, any memory performance based upon a single underlying mechanism would be expected to show a monotonic decrease with retention interval. The U-shaped serial-position function may be different. A dual process is likely operating. It has been suggested that two storage mechanisms are related by a rehearsal process (e.g., Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). Others argued that two different interference mechanisms combine to produce the U-shaped serial-position function (e.g., Postman, Stark, & Fraser, 1968; Postman & Underwood, 1973).

Recall experiments with human subjects have shown the primacy effect to be invariant over a wide variation in the retention interval, whereas the recency effect rapidly dissipates as retention interval is lengthened (e.g., Glanzer & Cunitz, 1966; Glanzer & Cunitz, 1966; Postman & Phillips, 1965). These different time courses for the primacy and recency effects argue for a dual-process, serial-position function.
One purpose of the experiments reported in this article was to determine whether or not dual processes are indeed responsible for the monkey's U-shaped serial-position function and whether or not the monkey too has different time courses for its dual processes. Varying the retention interval (called probe delay) in a serial-probe-recognition (SPR) task is a better method than varying retention interval in a recall task. It would, of course, be impossible to conduct a free-recall experiment with monkeys even if this were the more desirable procedure. In the SPR task, a list of items is presented and followed (after a probe delay) by a single test (probe) item. The subject then indicates whether or not the probe was in the list. In a recall task, on the other hand, the subject has a free-recall period where as many items as possible are recalled. Thus, more time has elapsed when the last items are recalled (effectively longer retention interval) than when the first items were recalled. This time difference due to recall order may be particularly important at short retention intervals where it may be a substantial fraction of the retention interval itself. The SPR task takes more time to generate a serial-position function, but the retention interval is precisely controlled because only one item is tested with each trial.

The monkeys used in this study had not previously participated in memory experiments, and in order to test their serial-position functions at different probe delays, they had to be trained in several steps to perform the serial-probe-recognition task. The first experiment reports training in a same/different task which is the first step. In addition, conditions are reported that promote same/different concept learning, which by the nature of being a relational strategy shares common properties with the strategy required to perform the SPR task.

Experiment 1: Same/Different Acquisition

The first step in training the monkeys to perform a complicated task like the SPR task is to make it as simple as possible. In this case, it means using only one list item and presenting it simultaneously with the probe item so that they overlap in time. The monkeys can look back and forth between the two before making their choice response. This is a same/different task: If the items are identical a same response is correct, and if they are not identical a different response is correct. Only a small change is necessary to make the same/different task an SPR task. If the two items do not overlap in time, the first item becomes the list item, an SPR task with one list item. It is then a simple matter to display more than one list item; if the probe item matches any one of the list items, a same response is correct; otherwise, a different response is correct. Monkeys trained to use a same/different concept may be more likely to learn the SPR task rapidly than they would without this concept.

One of the purposes of the first experiment was to test for development of any same/different concept by testing the degree to which they would transfer performance from a learned set of items to a completely new set of items. Regarding a same/different concept, the subjects make their choice response based upon the relation of the probe item to their memory of the list item. This is a form of memory scanning (cf., Sternberg, 1966) whereby the list item is stored during its presentation, retrieved when the probe item is presented, and compared with the probe item. If the memorial representation matches the probe item, then they are judged to be the same; otherwise, they are judged to be different. The alternative to memory scanning, a relational response, is an item-specific response. The correct response to individual items of the pair are memorized. We felt that if we could encourage relational responding, as demonstrated by good transfer in the same/different task, then the subjects might more rapidly acquire the subsequent SPR task than if they were item-specific responding in the same/different task. These were the purposes of Experiment 1; the first phase involves same/different training with six different sets of items, and the last phase involves transfer with additional training intervening between the transfer tests.

Method

Subjects

The subjects were 3 four-year-old experimentally naive rhesus monkeys (Macaca mulatta), Joe, Linus, and Max.
The monkeys obtained their entire daily ration of food and liquid as reinforcers in the experimental chamber. They were reinforced with a 5-cc squirt of Tang orange juice or a 1-gm Noyes banana pellet for correct responses; the particular reinforcer was determined on a random basis for correct responses. This deprivation schedule was maintained during the 5 working days. On Friday afternoons, they were given free access to Purina monkey lab chow and water. On Saturdays, they were given 500 cc of water and 20 Purina monkey chow pellets. On Sundays, they were deprived of food and water in order to get them ready for the next week’s experimental training and testing.

**Apparatus**

The monkeys worked the experimental sessions in a standard primate chair with movement somewhat restricted by a collar. The monkeys viewed two rectangular (12 cm × 9 cm) rear-projection screens arranged vertically and separated 16 cm center to center. A Carousel slide projector (Kodak 760H) was positioned 61 cm behind each screen. The images were viewed at a distance of approximately 61 cm, and subtended visual angles of 12 degrees vertically and 20 degrees horizontally. Solenoid-operated shutters, constructed “in house,” controlled presentation times of the slide items. A three-position (left, right, and down) lever was placed within easy reach of each monkey’s right hand. A Cromemco Z-2D microcomputer controlled the experiment, and collected and analyzed the data.

**Same/Different Procedure**

Trials began by the simultaneous projection of two slides, one on each of the two screens. On *same* trials, identical slides were displayed on top and bottom screens; on *different* trials, nonidentical slides were displayed. The monkeys’ task was to classify the slides as either *same* or *different* by moving the response lever to the right or left, respectively. Correct responses were followed by a tone (500 Hz), a reinforcer (either a banana pellet or a 5-cc squirt of Tang orange juice), and a 2-s intertrial interval. Incorrect responses were followed by illumination of the house light for 10 s and no reinforcer. Each session consisted of 70 trials; 35 were *same* trials and 35 were *different* trials, intermixed with one another in a pseudorandom sequence. A correction procedure was used whereby a trial was repeated until the correct response was made. This helped remove any response biases that might have developed. However, all the data presented in this article included only the first-trial performances, and the correction procedure was used only during this initial set of six acquisitions. Generally, three 70-trial sessions were conducted daily; the same 70-trial session was repeated until the monkeys performed at 88% correct or better on two sessions (not necessarily consecutive). They were then trained on the second 70-trial session and so on until they had mastered all 6 sessions.

The stimuli were 210 distinctly different 35 mm color slides of fruit, flowers, animals, people, and other natural and man-made objects. All items were used in the first two 70-trial sets of items. For the third and fourth sets, these items were reorganized, and in some cases items that were used on *same* trials were used on *different* trials and vice versa. For the fifth and the sixth acquisitions, these 210 different items were reorganized again. The method of organization and reorganization of the items for the six different acquisitions is shown in Table 1. As an example, notice the group of items designated by numbers 1 through 70. Half of them were shown with their identical copies as *same* trials in the first acquisition and the other half in the second acquisition. Pairs of items from this group were used in the third and fourth acquisitions as *different* trials. Similarly, the group of items identified by the numbers 141 through 210, which initially were used on *different* trials in the first and second acquisitions, were used (along with their identical copies) on *same* trials in the third and fourth acquisitions. The reason for this organizational scheme was to break down any item-specific associations that the monkeys might attach to individual items; it was hoped that they would learn to attend exclusively to relations between items of each pair. Good transfer to novel items would be evidence that the monkeys were using a relational strategy.

**Transfer Procedure**

The procedures used for transfer testing and the training between transfer tests were similar in many respects to the previously described acquisition procedures. However, there were three differences: The items were presented successively rather than simultaneously, no correction procedure was used, and the monkeys initiated each trial. Trials began with a faintly audible clicker (5 Hz). Each monkey pressed down on the three-position lever, which displayed the first item in the upper screen. The first item was displayed for 1 s, and 1 s after its removal 10 s elapsed, whichever occurred first. An incorrect response or no response during the choice response period (abort) resulted in a 10-s time-out period with the house light illuminated. Baseline trials used in these transfer tests were those of the sixth and final acquisition. Transfer tests were conducted on five consecutive sessions; each session was composed of 50 baseline trials plus 20 transfer trials. Twenty of the original 70 baseline trials were removed from the session, and 20 pairs of novel stimuli were substituted; the particular substitutions varied among test sessions. These novel test stimuli were shown only.

### Table 1

<table>
<thead>
<tr>
<th>Acquisitions</th>
<th>1–70</th>
<th>71–140</th>
<th>141–210</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 and 2</td>
<td>Same</td>
<td>Different</td>
<td>Different</td>
</tr>
<tr>
<td>3 and 4</td>
<td>Different</td>
<td>Different</td>
<td>Same</td>
</tr>
<tr>
<td>5 and 6</td>
<td>Different</td>
<td>Same</td>
<td>Different</td>
</tr>
</tbody>
</table>
once (by definition of novelty), and correct choice responses were reinforced. Previously, we had discovered that when transfer test trials were conducted in extinction, the monkeys learned that responses to novel stimuli were not rewarded (Wright, Santiago, Urcuioli, & Sands, 1984). This discrimination was learned even with partial reinforcement for correct responses on training trials. The extinction test, consequently, proved to be an invalid test of the subjects' transferability.

The two transfer tests were separated by training. The amount of training varied somewhat from monkey to monkey: Joe received 4 days, Max 8 days, and Linus 12 days. Four new sets of 70 trials composed of novel items were used in this training. Each day a particular set was repeated three times, and then the next day a different set was used. If training progressed beyond 4 days, then it was repeated on the four sets of stimuli in the same order. The training order was Sets A, B, C, D for Joe and Linus and Sets D, A, C, B for Max.

Results

Same/Different Acquisition

Individual acquisition results for 2 monkeys are shown in Figure 1. One monkey, Max, participated in only three out of six acquisitions. Consequently, its acquisition data, although similar to the other monkeys' data, are not presented for this phase of the experiment. The main result shown in Figure 1 is that the first acquisition took over twice as long as subsequent acquisitions. Notice that the first acquisition for Linus was plotted in terms of 560 trials per block; Linus's first acquisition took approximately four times longer than the others. Another result is that the even-number acquisitions began at about 70% correct, whereas the odd-number acquisitions (except for Joe on the fifth acquisition) began at about 50% correct (chance performance). An examination of Table 1 may help to explain this difference in acquisition. On the second acquisition, all the slides were completely novel to the monkeys and as such represented a transfer test. Their first-session performance on the second acquisition (not shown in Figure 1 because it is an average of two sessions) was 72% correct. This was moderately good transfer performance. The third acquisition began at chance performance, which was not completely surprising because a substantial number of previously seen items were rearranged into new combi-

![Figure 1](image-url)
nations; some had been used to make up same trials and were now used on different trials and vice versa. During the first session of the third acquisition, performance on trials where the role had changed (from same to different or from different to same) averaged 31% correct, whereas on trials where there was no role change (for slide numbers 71 through 140) performance was perfect (100% correct). None of the items in the third acquisition was used in the fourth acquisition. Like the third acquisition, the fourth acquisition was made up of items from both of the first two acquisitions. The same arguments used to explain the initial performance on the third acquisition might be expected to apply to the fourth acquisition. First-session performance on the fourth acquisition, however, revealed that the monkeys performed at 67% correct on items where the role had changed and 67% correct on items where the role had not changed. This trend in the first-session performance may represent the beginning of concept development by these monkeys. Indeed, their first-session performance never again fell below chance performance as it did on the third acquisition for items where there was a role change. On the fifth and sixth acquisitions the monkeys’ performance to items with a role change was 63% correct in both cases, again better than chance performance. However, 63% was not very good transfer and was even a slight decrease in transfer performance, rather than an improvement which might be expected if the monkeys had continued to develop the same/different concept.

Transfer

Transfer results from the first transfer test are shown in the left-hand pair of histograms for each monkey in Figure 2. Performance on the 100 transfer trials is shown by the hatched histograms, and performance on the 250 baseline trials is shown by the unfilled histograms. There was no significant transfer on the first test for Joe and Max, \( t(5) = 1.03, p > .1 \) ns; \( t(5) = 2.3, p > .05 \) ns, respectively, and although Linus’s transfer was significantly different from chance, \( t(4) = 3.58, p < .05 \), it was less than chance performance.\(^1\) All transfer performances were significantly different \( (p < .05) \) from their baselines, \( F(1, 8) = 8.2, F(1, 8) = 123.8, F(1, 8) = 23.2 \) for Joe, Linus, and Max, respectively.

\(^1\) These are significance tests for a single mean because the performance is tested against a fixed value—50% or chance performance (see Hays, 1963, p. 311).
Prior to the second transfer test, the monkeys were given training on four new sets of items. Their individual session performance is shown in Table 2. On the three daily sessions of the third set of items, Joe showed a dramatic improvement in performance, possibly an example of insight learning. This improved performance was maintained through the fourth set of items and was stable thereafter. As shown in Figure 2, performance on the second transfer test for Joe improved dramatically. The other two monkeys were trained somewhat more than Joe on the four sets of items, and their performance is shown in Table 2. Their transfer performance improved also. Transfer by all three monkeys did not differ from their baseline performances, $F(1, 8) = 2.4, p > .20; F(1, 8) = 0.0, p > .75; F(1, 8) = 0.5, p > .75$ for Joe, Linus, and Max, respectively, and was significantly above chance, $t(4) = 11.4, p < .002; t(4) = 11.6, p < .002; t(4) = 9.8, p < .002$ for Joe, Linus, and Max, respectively.

**Discussion**

Previous studies of acquisition and transfer of the *same/different* concept showed that great apes (chimpanzees and orangutans) can acquire and correctly apply this concept (King, 1973; Robinson, 1955, 1960). The experiment of this article showed that rhesus monkeys can also learn and correctly apply the *same/different* concept. Indeed, the 85% correct transfer by the monkeys of this experiment was even better than that shown by the apes. The apes' transfer performance was less than 80% correct. In one experiment by King (1973), their transfer performance was less than 70% correct. The apes' transfer performance also was typically less than their training item performance. By contrast, our monkeys showed no deficit in transfer performance relative to training item performance. Thus, with regard to a *same/different* concept, monkeys were not shown to be inferior to apes in whether or not they could learn the concept or in the degree to which they could correctly apply it.

A companion experiment to the one presented here with monkeys was conducted with pigeons (Santiago & Wright, 1984). Pigeons were trained and tested with identical items and with procedures virtually identical to those used with monkeys. Pigeons, however, did not transfer as well as monkeys, $F(1, 23) = 44.6, p < .001$ for individual sessions for each species as replications. Figure 3 shows average pigeon transfer and average monkey transfer for comparison. Although pigeons did not transfer as well as monkeys, they did show transfer nonetheless, $t(9) = 13.8, p < .002$ for individual sessions as replications. To our knowledge this is the best reported transfer by pigeons. The pigeons' baseline performance was less than the monkeys' although at one time both species had

**Table 2**

*Percent Correct Performance on Slide Sets Separating Transfer Tests*

<table>
<thead>
<tr>
<th>Slide set</th>
<th>Monkey</th>
<th>Linus</th>
<th>Max</th>
<th>Joe</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>55, 60, 49</td>
<td>55, 54, 68</td>
<td>66, 71, 77</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>65, 68, 71</td>
<td>64, 68, 71</td>
<td>64, 75, 65</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>72, 76, 75</td>
<td>77, 71, 72</td>
<td>84, 87, 96</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>51, 59, 49</td>
<td>61, 65, 55</td>
<td>86, 89, 94</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>57, 72, 72</td>
<td>83, 78, 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>81, 71, 74</td>
<td>75, 83, 86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>70, 87, 84</td>
<td>84, 86, 94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>54, 61, 64</td>
<td>70, 81, 88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>77, 78, 75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>80, 74, 80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>75, 77, 88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>74, 74, 71</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Figure 3. Comparison of monkey and pigeon transfers on their final transfer test. (Range markers specify the total range of performance for three monkeys or two pigeons.)](image-url)
performed about equally well on these baseline trials. Pigeons had met the 88% performance criterion and had maintained better than 80% correct for several baseline sessions before transfer. The decrement in the pigeons' baseline performance indicates that the transfer test itself may have been disruptive. Pigeons are generally more disrupted by procedural changes than are monkeys. Possibly continued training with daily item changes interspersed with occasional novel items might eventually produce pigeon transfer as good as monkey transfer.

The monkeys' transfer performance was 85% correct and as good as their baseline performance. This means that they had fully developed the same/different concept. The monkeys could deal with new items in the same way they dealt with old items. Decisions were made on the basis of the relation between the items of each pair. Because they had never before seen any of the items, they had no opportunity to memorize the correct responses to individual items. This was not always the case, as shown by their poor initial transfer. Apparently, repeating the same trials over and over again promoted development of item-specific strategies to the detriment of relational ones.

Premack (1983) raised the question of whether transfer, such as reported in this article, is actually based upon a concept or whether it is based upon the animal making a familiarity judgment, choosing "the item that it has seen before or seen most recently" (p. 356). Premack did not present evidence that any transfer results were based upon familiarity judgments or that they were not based upon a same/different concept. Indeed, he did not even identify what the proper same/different task strategy should be in order to qualify as a concept, what the familiarity strategy actually is in this task, or how they differ. For purposes of this discussion, a same/different concept is defined as one where (a) the subject compares the second item (or its memorial representation) with its memory of the first item, (b) a similarity criterion is used to make the response decision, and (c) the subject responds same if this comparison meets the criterion and different if does not.

Good transfer performance, and hence evidence of good concept development, necessitates that the subject: (a) use the relational strategy as outlined above, (b) use a sufficiently strict criterion for judging the pair of items, and (c) adequately learn the proper identification responses of same and different.

If subjects showing good same/different transfer could be separated into those that use same/different concepts and those that use familiarity judgments, then these two groups would probably be distinguished by differences in their criterion (b above) rather than by any basic strategy differences (a above). Such a distinction would be one of degree (quantitative), not kind (qualitative). Premack (1983), on the other hand, seemed to argue for a basic strategy difference, a qualitative difference.

The common use of the term familiarity implies that the comparison extends backward in time, possibly to all previous items seen during the session. Perhaps if the comparison were restricted to the single previous item, then this strategy would be closer to a concept for those who distinguish between concepts and familiarity judgments. The basic relational strategy would remain the same in both cases. Support for this argument comes from a monkey memory-scanning experiment (Sands & Wright, 1982). The procedure (in Experiment 2) had been changed by previewing (in prior lists) some of the probe items on different trials. Thus, for the monkey to be correct the relational judgment had to be limited to list items of the current trial. The memory-scanning results revealed that this was indeed what the monkey did. The monkey quickly adapted to this procedure change, an unlikely adjustment if the monkey had had to learn a completely different strategy.

Another common use of the term familiarity judgment is that it is a snap judgment, one that is imprecise or superficial. For example, a probe item might be judged (on the basis of a snap, familiarity judgment) to be the same as the previous item, if both contained large expanses of blue sky notwithstanding differences in land, buildings, and so on in the lower portions of the pictures. Training may have laid the foundation for such a strategy. If all pictures had previously been very different from each other, then this snap-judgment strategy should work well. Training establishes how similar the two items
(images) must be in order to be correctly judged identical. Once again this is a difference in criterion, not a basic strategy difference. Tests of this argument could be conducted. Lures (probe items on different trials) could be selected to be more similar to their respective preceding item. The subjects would then have to make detailed comparisons in order to be correct, not just gross or snap judgments. Properly selected pairs of test items, tested before and after this procedure change, would reveal changes in the degree of comparison and changes in similarity criterion.

Thus, if subjects were trained to compare the test item (or its memorial representation) only to the previous item, make a point-by-point comparison of these memorial representations, and respond same only when they were identical in all respects, would not they be said to be responding on the basis of a same/different concept? Even the most skeptical concept theorist should agree. All of these can be accomplished by restricting the judgment criterion of what is commonly thought of as a familiarity judgment. It does not involve a different strategy.

Finally, the strategy used in the SPR task can also be argued to be based upon a same/different concept. The task is similar to a same/different task, except that there is a whole list of items preceding the test item rather than a single item. The subject retrieves the list items one at a time from memory (cf., Sternberg, 1966) and compares each with the test item (or memory of the test item). If any list item matches the test item in all aspects, then the subject responds same. If none matches, the subject responds different. (Again, the requirement that the list item match the test item “in all aspects” is a criterion adopted by the subject and imposed by the experimenter through the selection of the stimuli and the training conditions.) This same behavior and strategy is performed whether the items are new (transfer test), old, or just in new combinations. We regard such behavior to be evidence for, and based upon, a same/different concept. Performing the same/different task on the basis of the same concept used in the SPR task (same/different concept) was possibly an important step in the monkeys' acquisition of the SPR task.

The monkeys' acquisition of the SPR task is presented in the next experiment.

Experiment 2: Serial-Probe-Recognition Acquisition

Experiment 2 reports the acquisition of SPR performance. The monkeys began SPR training immediately after demonstrating excellent transfer in Experiment 1. They could thus be expected to perform well on one-item, list-length trials, but how would they do with longer list lengths? By presenting them with list lengths that varied from one to six items, how much training would be required before they learned all list lengths? These were some of the questions explored in Experiment 2.

Method

Subjects and Apparatus

The monkeys were the same three monkeys that participated in the previous experiment. The apparatus used was identical to that described in the first experiment.

Procedure

Trials began by the monkey pressing down on the three-position lever in the presence of a clicker (5 Hz) signal. A downward press of the lever initiated presentation of the list items which could include 1, 2, 3, 4, 5, or 6 items. Trials of different list lengths were intermixed in a pseudorandom order within the session. Each list item was presented for 1 s, with a 1-s delay between items. Following the last list item, there was a 1-s delay before presenting the probe item. The probe item was presented on the lower screen and remained in view until the monkey made a choice response. Movements of the lever to the right (single-choice response) were correct if the probe item matched any list item. Movements to the left were correct if it matched none of them. Correct responses were rewarded with either a 5 cc squirt of orange juice or a banana pellet, the type of reinforcer was determined randomly. Incorrect responses initiated a 10-s time-out with the house light turned on. A 3-s intertrial interval followed reinforcement for correct responses. The next trial began after the intertrial interval or the time-out period.

There were 216 unique trials used in this experiment constructed from 864 uniquely different items (108 same trials and 108 different trials). The Kodak Carousel projectors that were used held only 140 items, so the trials were distributed into six different pairs of Carousel trays (one tray held the list items and the other held the probe items). The monkeys generally performed in all 216 trials daily. The items were selected from a 3,000-item color-slide pool; they were novel items and had not been previously seen by the monkeys. Care was taken in item selection and composition of the trials so that items
Results

The monkeys' acquisition of the variable-list-length task is shown in Figure 4. The three monkeys acquired the variable-list-length SPR task at about the same rate. They gradually acquired the task beginning at about 60% correct and improved their performance to about 85% correct. Strangely, none showed any acquisition differences among the different list lengths. Acquisition differences were tested over the first 21 sessions (blocked by 3 sessions), where differences in acquisition might be expected to be greatest; there were no differences, $F(5, 36) = 0.8, p > .5$ ns; $F(5, 36) = 3.0, p > .05$ ns; $F(5, 36) = 0.6, p > .5$ ns for Joe, Linus, and Max, respectively. Average performance over the first 21 sessions was: 74.5%, 71.0%, 69.7%, 65.9%, 64.5%, 68.8% correct for one- through six-item list lengths, respectively. At no other time during their acquisition did the subjects show list-length performance differences either. Prior to this, they showed 85% transfer to one-item list lengths. Thus, we thought it likely that they might immediately perform well with one-item list lengths (equivalent to a same/different task), rapidly acquire the two-item list length trials, and more gradually acquire the others, in order of list length. This was not so. There must have been some leveling effect. All list-length performances began somewhat above chance performance and remained at about the same level throughout acquisition.

Discussion

It is not clear to us why the shorter list lengths were not more rapidly acquired than the longer ones. We had expected, from the same/different transfer results, about 85% correct transfer to the one-item list lengths. Pigeons too had been disrupted in their SPR performance when the list length varied (Wright et al., 1984). Pigeons performed much better when list length was fixed within blocks of trials than when it varied. Pigeons and monkeys generally are disrupted by the same sorts of things, only the disruption is usually more pronounced for pigeons than for monkeys. Pigeons were trained on the same variable list-length task with the same items as the monkeys. Pigeons, however, showed no signs of acquiring the task. Figure 5 shows the average monkey and average pigeon SPR performance on the one- to six-item variable-list-length task. Monkey acquisition, although slow, began well above pigeon performance and showed a gradual increase. The pigeons' performance remained at chance level throughout the experiment. Pigeons apparently are more disrupted by changing list length than are monkeys.

The mechanism of this disruption is not clear, but human subjects can have their memory processing disrupted when the post-
stimulus interval varies or when the exposure duration itself varies (Proctor, 1983). Proctor hypothesized that the human subjects would not rehearse when they were uncertain about onset and offset of the stimulus items. Finally, human subjects do not typically reveal primacy effects when the list length varies (Sternberg, 1966). Possibly this too is a disruption of some memory process. Thus, there are a number of variables that probably influence SPR performance. We mentioned three of them: list length, interstimulus interval, and item viewing time. All three are likely to have different effects depending upon whether they are blocked or variable. In the next experiment we deal with a fourth variable—probe delay. This is the time interval between the last list item and the probe (test) item. In this experiment we used a blocked design to investigate the effects of probe delay.

Experiment 3: Serial Position Curves as a Function of Probe Delay

This experiment was conducted after the monkeys had acquired the SPR task with variable list lengths of one to six items. They transferred well to fixed three-item list-length SPR task (average of the first five sessions for the three monkeys was 80% correct), and later to a four-item list-length SPR task (five-session average of 74% correct). Their maintained good performance in the SPR task allowed the retention interval to be increased without the risk of a floor effect. The effect of a variable such as probe delay (retention interval) can be measured only if performance is high enough to leave room for decrements to be visible. Our purpose was to obtain serial-position functions with four-item list lengths and see how the form of the serial-position function changed with probe delay. We were encouraged in this exploration by our previous results showing a U-shaped monkey serial-position function (Sands & Wright, 1980a, 1980b), indications in the human memory literature that the serial-position function results from dual processes, and the previous finding (Santiago & Wright, 1984) that the pigeons' primacy and recency effects change in separate ways with probe-delay changes.

**Method**

**Subjects**

The subjects were 2 monkeys, Joe and Linus, from the previous experiment. They were tested in this experiment immediately after they had acquired the SPR task with the variable list lengths (1–6 items) and had shown good performance with fixed list lengths of three and four items.

**Apparatus and Procedure**

The apparatus was the same as that used in the previous experiments. The procedure was similar except that the list length was fixed at four list items for each trial. The monkeys were trained and tested with two 20-trial sessions, each composed of 90 novel pictures which the monkeys had not previously seen. Each of six probe delays (0, 1, 2, 10, 20, and 30 s) was tested on four 20-trial sessions. The order of probe delay testing was: 2, 10, 0, 20, 30, 10, 2, 1, 10, 30, 20, 2, 30, 10, 2, 0, 1, 20, 0, 1, 30, 10, 1, 20, and 0 s.

**Results**

Figure 6 shows the results for the 2 monkeys tested at different probe delays. The
serial-position functions are from performance on same trials. Triangles are from performance on different trials (where the probe item did not match any of the list items). The 2 monkeys revealed the same trends in their serial-position function changes: The shortest probe delay (0 s) showed performance generally increasing from the first to the fourth item, revealing a recency effect and no primacy effect. The next three probe delays (1, 2, and 10 s) show prominent primacy effects as well as recency effects, revealing a U-shaped serial-position function. The 20-s probe delay shows the beginnings of a change from a U shape to monotonic decreasing shape. The longest probe delay (30 s) showed a monotonic decreasing serial-position function, revealing a prominent primacy effect and no recency effect.

These serial-position function changes were tested by a series of weighted contrast tests (Keppel, 1973, p. 94) and are shown in Table 3. At 0-s probe delay, Serial Position (SP) 4 was significantly greater than SP1 for both monkeys. In addition, Joe's SP1 was significantly less than its other three SPs, and Linus's SP4 was significantly greater than its other SPs. At 1-s probe delay, SP2 was significantly less than other SPs for both monkeys, demonstrating the U shape of the SP function. Linus also showed this same result at 2- and 10-s probe delays. Joe's trough moved from SP2 to SP3 at 2- and 10-s probe delays, and SP3 was significantly less than other SPs. At 20 s the SP functions are transitional. For both monkeys at 20 s, SP4 was significantly less than SP1 demonstrating the downward trend, but they both show inversions at SP3, which are probably the remains of the U shape. At 30-s probe delay, SP4 was significantly less than all other serial positions, demonstrating the decreasing trend of the functions.

In addition, the absolute levels of primacy and recency effects were compared. Changes in SP1 were tested with a series of contrast tests (Keppel, 1973). SP1 at 0-s probe delay was significantly less ($p < .05$) than SP1 at all other probe delays except 10 s. Among all other probe delays there were no significant differences for SP1. These tests support the SP1 trends shown in Figure 6. There was no hint of a primacy effect at 0-s probe delay. There was a prominent primacy effect at 1-s probe delay, and this primacy effect remained fairly stable out to the longest probe delay of 30 s. Contrast tests were also performed on SP4. There were no significant differences among the SP4 comparisons for 0-, 1-, 2-, and 10-s probe delays. However, SP4s at 20-
and 30-s probe delays were significantly different \((p < .01)\) from SP4s at the shorter probe delays. Performances for SP4 at 20 and 30 s did not differ significantly from each other. These validate the trends in the recency effects (SP4) shown in Figure 6. The recency effect was high and stable for 0-, 1-, 2-, and 10-s probe delays, but dissipated at 20- and 30-s probe delays. Finally, the absolute levels of the primacy and recency effects were not different. The mean SP1 performances for both monkeys at 1, 2, 10, 20, and 30 s were pooled, because these values were shown to be similar. They were compared with pooled SP4 values for 0, 1, 2, and 10-s probe delays because these values also had been shown to be similar. There was no significant difference, \(F(1, 16) = 0.5, p > .50\). This means that the primacy effect rose to the same level as the recency effect. In both cases the asymptotic level was very high: a mean of 96% for the primacy effect and 98% for the recency effect. Likewise, the low performance on SP1 was not significantly different from the low performance on SP4. Pooled SP4 mean performance at 0-s probe delay was compared with that for SP4 at 20 and 30 s, \(F(1, 4) = 2.1, p > .20\). The primacy and recency effects covered a very broad range because they had low mean values of 70% and 60%, respectively.

### Discussion

Only recently have primacy effects been shown for animals (Buchanan et al., 1981; Kesner & Novak, 1982; Roberts & Kraemer, 1981; Sands & Wright, 1980a, 1980b; Santiago & Wright, 1984). Figure 5 shows that only intermediate probe delays will produce the characteristic U-shaped function, revealing both primacy and recency effects. If the probe delay is too short, a primacy effect will not be found. If it is too long, a recency effect will not be found. Although pigeons show the same changes in their serial-position functions as do monkeys, a long probe delay (10 s) for pigeons is only a moderate probe delay for monkeys. Thus, for each individual species the particular probe delay to reveal both primacy and recency effects may have to be empirically determined.

Probe delay had a powerful effect on the

## Table 3

### F Values for Contrast Tests Between All Pairs of the Four Serial-Position Performances

<table>
<thead>
<tr>
<th>Probe delay/monkey</th>
<th>Serial position</th>
<th>Overall serial-position effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 vs. 3</td>
<td>4 vs. 2</td>
</tr>
<tr>
<td>0 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td>0.0</td>
<td>5.2</td>
</tr>
<tr>
<td>Linus</td>
<td>16.2**</td>
<td>44.0**</td>
</tr>
<tr>
<td>1 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td>2.1</td>
<td>79.5**</td>
</tr>
<tr>
<td>Linus</td>
<td>28.8**</td>
<td>141.7**</td>
</tr>
<tr>
<td>2 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td>248.1**</td>
<td>0.0</td>
</tr>
<tr>
<td>Linus</td>
<td>1.4</td>
<td>36.0**</td>
</tr>
<tr>
<td>10 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td>115.6**</td>
<td>3.6</td>
</tr>
<tr>
<td>Linus</td>
<td>2.6</td>
<td>22.0**</td>
</tr>
<tr>
<td>20 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td>19.4**</td>
<td>3.8</td>
</tr>
<tr>
<td>Linus</td>
<td>25.8**</td>
<td>0.0</td>
</tr>
<tr>
<td>30 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td>48.0**</td>
<td>200.9**</td>
</tr>
<tr>
<td>Linus</td>
<td>19.9**</td>
<td>7.2*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom for pairwise serial position tests and overall serial-position effect are 1, 9 and 3, 9, respectively. *4 = last list item. *p < .05. **p < .01.
serial-position function. It was somewhat surprising to find that initial increases in probe delay produced increases, rather than decreases, in performance; the primacy effect, totally absent at 0-s delay, emerged with the initial increases in probe delay. As probe delay was further increased, the recency effect began to dissipate, eventually leaving only the primacy effect intact. These dynamic serial-position function changes argue for dual memory processes. They are, at the same time, evidence against single-process explanations of the U-shaped serial-position function such as a network of associations and inhibitions (Hull, 1935; Lepley, 1934), or end-point distinctiveness (Bower, 1971; Murdock, 1960). The dynamic serial-position function changes show that two processes are operating quite independent of one another. The primacy effect emerges before there is any apparent change in the recency effect. The primacy effect is quite stable thereafter, and the recency effect begins to dissipate with further delay increases.

Dual memory processes have been proposed previously. The primacy portion and the medial portion of the serial-position function have been argued to represent a long-term memory store (LTS), whereas the recency portion has been argued to represent a short-term memory store (STS; see, for example, Glanzer, 1972). The recency portion may actually be a combination of LTS and STS. Techniques have been developed to parcel out the separate contributions of LTS and STS to the various portions of the serial-position function. This separation is made possible by comparing and analyzing results from immediate recall and delayed (with distractor task) recall. Delayed recall is free of any STS contribution (so it is argued) and represents pure LTS memory. The STS contribution is calculated by subtracting the delayed-recall, serial-position function (pure LTS) from the immediate-recall function (combination of LTS and STS). This subtraction is possible as long as delayed recall is equal to or less than immediate recall at all positions of the serial-position function. Otherwise, the result will be negative at those serial positions, an impossibility from the standpoint of memory. However, this is precisely the finding that we have shown in this article for the initial increases in probe delay. Immediate recognition of the first list items is poor. After about 1-s delay, performance has substantially risen revealing a primacy effect. This emergence of a primacy effect with delay is a problem for all theories that calculate STS in this manner (e.g., Glanzer, 1972; Raymond, 1969; Waugh & Norman, 1965).

The literature on human memory does not afford many comparisons to the probe-delay experiments in this article. Human retention interval (the analogous variable to probe delay) studies have largely been confined to paired-associate procedures or free-recall procedures. Paired-associate procedures do not yield serial-position functions. Recall procedures do not provide adequate control of the effective delay time; there is a longer delay for items that are recalled last than for those that are recalled first. One study (Jahnke & Erlick, 1968), however, did use an SPR procedure (delay is precisely controlled in SPR), and the probe delay was varied. The problem with this study was that delay was not varied over a sufficiently wide range. Only three intermediate delays were used: 4, 8, and 12 s. All the resulting serial-position functions were U shaped. Tests were not performed with the delay values (0 s and >30 s), which might have produced the more interesting shaped serial-position functions, monotonic increasing, and monotonic decreasing.

Several studies have used free-recall procedures and have shown dissipation of the recency effect with probe delay (Gardiner, Thompson, & Maskarinec, 1974; Glanzer & Cunitz, 1966; Postman & Phillips, 1965; Roediger & Crowder, 1975). These studies used delays of up to 30 s. Glanzer and Cunitz (1966) and Postman and Phillips (1965) used several delay values, whereas the other researchers used only 0-s and 30-s delay values. The results were consistent from these four studies. As delay was increased, the recency effect...
effect dissipated. The degree to which it dissipated varied somewhat among the studies. Gardiner et al. (1974) and Roediger and Crowder (1975) showed an actual negative-recency effect with auditory presentation of the items. That is, performance continued to fall throughout the list with the terminal-item performance being less than the medial-item performance. Gardiner et al. (1974) did not show a negative-recency effect with visual presentations. Terminal-item performance did not fall below medial-item performance. Perhaps they would have found a negative-recency effect with visual items had they used even longer probe delays. We showed a negative-recency effect for monkeys with visual items at 30-s delay, whereas only 10 s was required to reveal one for pigeons. Humans might require a delay considerably longer than monkeys to reveal their negative-recency effect with visual stimuli. The probe-delay experiments with pigeons and monkeys demonstrated that the negative-recency effect was not a separate phenomenon; it formed one end point on a continuum of effects that probe delay had on the serial-position function.

None of the four studies of humans found monotonic increasing functions at 0-s delay as we have shown for pigeons and monkeys. The reason for their failure to find monotonic increasing functions at 0-s probe delay may be due to the nature of the free-recall procedure itself. The free-recall procedure allows a block of time (e.g., 60 s) for the subject to write down all of the list items that are remembered. Items recalled at the end of the recall period are delayed more than those recalled at the beginning of the recall period; "the S's own responses act as a delay task" (Glanzer, 1972, p. 163). The probe-delay results with pigeons and monkeys show that suppression of the primacy effect lasts only about 1 s. Even with humans the transition to a U-shaped serial-position function should take place within the first few seconds. Thus, the finding of a monotonic-increasing function could have easily been lost during the comparatively long recall period. If the first list items are not immediately recallable, perhaps they will be after some of the later items are recalled. Analysis of output gives support to this argument: "The terminal items of a series of unrelated words have the highest probability of recall and also are usually recalled first" (Postman & Phillips, 1965, p. 133).

One recall study used a procedure that controlled the time from the end of the list to the item recalled (Ellis & Hope, 1968, Experiment IV). They presented a single probe (former list item) following the list. The subjects had to identify the list position (by pressing the proper button) of the item. Their performance at a 0-s probe delay showed no primacy effect and a strong recency effect, a result similar to the ones we showed for pigeons and monkeys. At 10-s probe delay, the serial-position function revealed primacy as well as recency effects, the familiar U shape. The recency effect had started to dissipate slightly relative to the 0-s delay condition. Thus, this study demonstrates similar trends to those we showed for pigeons and monkeys. It indicates that if the interval from end-of-list to test is precisely controlled, then humans too will show an initial monotonic increasing function with a recency and no primacy effect, transition to a U-shaped function with longer delays, and finally the beginning of dissipation of the recency effect with further delay increases. Discussing the transition from the monotonic increasing function to the U-shaped one, Ellis and Hope said that the emergence of the primacy effect resulted "perhaps by providing opportunity to rehearse during the delay interval" (p. 617). There seems to be an inconsistency in such an argument. If the subject has the ability to retrieve the first items in order to rehearse them during the delay interval, then why can this same retrieval not serve as the basis for the correct response when tested immediately? This finding of emergence of the primacy effect with increasing probe delay is a problem for all theories that use rehearsal to explain the primacy effect (Atkinson & Shiffrin, 1968; Raffel, 1936; Waugh & Norman, 1965; Welch & Burnett, 1924).

Inadequacy of rehearsal as a theory to explain emergence of the primacy effect leads to the question of what is an adequate theory to explain it. There is one theory that seems adequate in this regard. It is interference theory, specifically the mechanism of release-from-retroactive-interference, also known as...
spontaneous recovery. Retroactive interference (RI) is interference from memory items, which follow the items being tested. Initially RI is large, but it dissipates comparatively rapidly with time (Abra, 1969; Briggs, 1954; Ceraso & Henderson, 1966; Forrester, 1970; Kamman & Melton, 1967; Koppenaal, 1963; Martin & MacKay, 1970; Melton & Irwin, 1940; Postman et al., 1968; Postman, Stark, & Henschel, 1969; Shulman & Martin, 1970; Silverstein, 1967; Slamecka, 1966; Underwood, 1948a, 1948b). Large initial RI accounts for the lack of a primacy effect when the initial items are tested immediately. The later list items interfere (retroactively) with memory of the first ones. Rapid dissipation of RI allows their memory to be recovered, resulting in emergence of the primacy effect. Turning now to dissipation of the recency effect with probe delay, this too can be accounted for by interference theory. Proactive interference (PI) is interference from memory items that precede the items being tested. “PI is virtually absent at first but increases steadily with time” (Postman & Phillips, 1965, p. 123). A number of studies have shown that PI is absent initially and slowly grows, and its growth is slower than RI dissipation (Keppel & Underwood, 1962, Postman, 1962; Postman et al., 1968). Initial absence of PI and its slow growth accounts for the recency effect being present well after the primacy effect has emerged. The eventual increase in PI accounts for the eventual dissipation of the recency effect.

These different RI and PI effects and their separate time courses have been worked out with paired-associate procedures. It is unfortunate that some interference theorists did not use an SPR procedure and investigate the effects on the serial-position function. The SPR procedure seems ideally suited and very sensitive to the different interference effects with which the previously mentioned theorists have been concerned. The time courses, however, are somewhat different, being considerably faster in the SPR procedure than in the paired-associate procedure. In the SPR procedure, RI dissipated, revealing the primacy effect in about 1 s. By contrast, in the paired-associate procedure, first-list recall (index of release-from-RI) increased over 30 min in some studies (Forrester, 1970; Kamman & Melton, 1967; Martin & MacKay, 1970; Postman et al., 1968; Postman et al., 1969; Shulman & Martin, 1970) and as much as 24 hr or more in others (Abra, 1969; Briggs, 1954; Ceraso & Henderson, 1966; Silverstein, 1967; Underwood, 1948a, 1948b). The possibility that a neurophysiological consolidation mechanism (cf., Hebb, 1949) is relevant to these results has been previously excluded because of the slow time course demonstrated by these paired-associate procedures and because an influential study (McGeoch, 1942) showed that the same RI effect could be obtained when the second list (interfering list) was presented 6 weeks after the first one. Any first list consolidation would have certainly run its course long before 6 weeks! The much shorter time course demonstrated here with the SPR procedure once again opens up the possibility that consolidation does play a role; the time values are of a proper order of magnitude for neurophysiological consolidation.

Concluding Remark

These changes in the serial-position function with probe delay for both monkeys and pigeons demonstrate that these animals have primacy and recency effects just as humans do. Their similar dynamic changes in serial-position functions indicate that underlying memory processes are similar if not identical in these two species. This is an important step in establishing continuity of species in terms of their cognitive ability, which may eventually lead to establishing that animals process information and think in ways basically similar to humans.

References


Kamman, R., & Melton, A. W. (1967). Absolute recovery of first-list responses from unlearning during 26 minutes filled with an easy or difficult information processing task. Proceedings of 75th Annual Convention of the American Psychological Association, 2, 63–64.


Robinson, J. S. (1955). The sameness-difference discrimi-

Received October 21, 1983
Revision received May 21, 1984