Levels of representation in the rhesus macaque: An examination of linear ordering mechanisms

Dustin J. Merritt
The University of Montana

Follow this and additional works at: https://scholarworks.umt.edu/etd
Let us know how access to this document benefits you.

Recommended Citation
https://scholarworks.umt.edu/etd/4932

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
The University of MONTANA

Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

** Please check "Yes" or "No" and provide signature **

Yes, I grant permission
No, I do not grant permission  

Author's Signature  

Date 5/24/99

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.
Levels of Representation in the Rhesus Macaque: An Examination of Linear Ordering Mechanisms

by

Dustin J. Merritt

B.A. The University of Montana, 1994

presented in partial fulfillment of the requirements for the degree of Master of Arts

The University of Montana

1999

Approved by:

[Signatures and dates]

Chairperson

Dean, Graduate School

Date
Recent accounts of performance on tasks examining linearity in nonhuman species have posited both cognitive and noncognitive interpretations. Given the standard preparation (A+B-, B+C-, C+D-, D+E-), Value Transfer Theory (VTT) explains transitive choice in terms of a proportional transfer of value (dependent on composite value of the stimulus) from the reinforced stimulus to the nonreinforced stimulus in adjacent-item pairings, with the amount of transferred value functioning as the decisive factor in subset (both adjacent and nonadjacent) tests. This explanation is in direct conflict with the idea that adjacent-pair training leads to representation of linear order. In order to examine the competing hypotheses, 2 male rhesus macaques were trained to solve a five-item simultaneous chain (Experiment 1), after which, list items were substituted at random with a wildcard item (Experiment 2). During the third experiment, Abe was trained to perform on a linear arrangement of 5 stimuli, while Bob was trained to perform a nonlinear task composed of 8 stimuli arranged in 4 pairs. For 2 of the pairs, the S+ was reinforced 100% of the time, and for the remaining 2 pairs, the S+ was reinforced 50% of the time. During the fourth experiment, both monkeys were presented with a wildcard task utilizing the same stimuli from Experiment 3 in order to determine whether any transfer would occur from Experiment 3 to Experiment 4. The icons for both monkeys were arranged in a series of descending stimulus values (as determined by VTT). It was hypothesized that the linear representation of the list would transfer from Experiment 3 to Experiment 4 for Abe, and thus facilitate performance during the fourth experiment relative to the second experiment, while Bob’s performance was not expected to change during Experiment 4 relative to Experiment 2. Results: Abe showed much poorer performance in Experiment 4 relative to Experiment 2, while Bob showed no changes in performance in Experiment 4 relative to Experiment 2. Although differences were opposite in relation to the predicted differences, there is some evidence to suggest that the linear representation established in Experiment 3 (for Abe) hindered performance in Experiment 4. Mechanisms of linear representation are discussed in addition to the potential for representational analogue processes.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>EXPERIMENT 1</td>
<td>19</td>
</tr>
<tr>
<td>Method</td>
<td>19</td>
</tr>
<tr>
<td>Subjects</td>
<td>19</td>
</tr>
<tr>
<td>Apparatus</td>
<td>19</td>
</tr>
<tr>
<td>Procedure</td>
<td>19</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>20</td>
</tr>
<tr>
<td>EXPERIMENT 2</td>
<td>30</td>
</tr>
<tr>
<td>Method</td>
<td>30</td>
</tr>
<tr>
<td>Subjects</td>
<td>30</td>
</tr>
<tr>
<td>Apparatus</td>
<td>30</td>
</tr>
<tr>
<td>Procedure</td>
<td>30</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>31</td>
</tr>
<tr>
<td>EXPERIMENT 3</td>
<td>47</td>
</tr>
<tr>
<td>Method</td>
<td>47</td>
</tr>
<tr>
<td>Subjects</td>
<td>47</td>
</tr>
<tr>
<td>Apparatus</td>
<td>47</td>
</tr>
<tr>
<td>Procedure</td>
<td>47</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>49</td>
</tr>
<tr>
<td>EXPERIMENT 4</td>
<td>56</td>
</tr>
<tr>
<td>Method</td>
<td>56</td>
</tr>
<tr>
<td>Subjects</td>
<td>56</td>
</tr>
<tr>
<td>Apparatus</td>
<td>57</td>
</tr>
<tr>
<td>Procedure</td>
<td>56</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>57</td>
</tr>
<tr>
<td>GENERAL DISCUSSION</td>
<td>63</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Error patterns during the first three sessions and the last three sessions of training. 79
2. Percentage correct responses at each wildcard position 80
3. Transition probabilities during the criterion session of the five item wildcard phase. 81
4. Average transition probabilities during the middle two sessions of the five-item wildcard phase 82
Transitive inference, often defined with the present example: If A is greater than B, and B is greater than C, then A must be greater than C given a linear system, has been examined utilizing an number of different species, where animals (or humans for that matter) are essentially integrating information from two relationships to draw inferences about a third. Although some researchers have made the argument that transitivity is language dependent, transitive inference has been widely studied in nonhuman animals (see Wynne, Delius, & Staddon, 1992; Markovits & Dumas, 1992). Gillan (1981) demonstrated that chimpanzees are capable of transitive inference. The chimpanzees were trained using the following five-term series procedure: A-B+, B-C+, C-D+, D-E+, and in the test trial, the chimpanzees were presented with B and D. Gillan hypothesized that if the chimpanzees were able to transively infer the value of B against D, the chimpanzees would choose D. In fact, one chimpanzee did choose D over B at above chance levels. Gillan suggests that this is evidence that chimpanzees are able to mentally organize these pairs into an ordered series and thus make the appropriate comparisons with regard to non-paired items.

However, McGonigle and Chalmers (1977) contend that deductive reasoning is not required in order to choose the correct stimulus in the B D pair. Reinforcement may be able to account for transitive inference. McGonigle and Chalmers argue that stimuli are encoded in triads. For example, BCD (B-C+, C-D+) has three possible pairings: BC, CD, and BD. If each of these pairs is presented an equal number of times, then the probability of any given pair being presented is .33. For each pair, a stimulus value of 1.0 is given to the positive (rewarded) stimulus and a value of 0.0 is given to the negative
(non-rewarded) stimulus. In order to obtain absolute stimulus values, the probability of presentation is multiplied by the initial stimulus value (see example below).

<table>
<thead>
<tr>
<th>Choice</th>
<th>Stimulus value (+)</th>
<th>Stimulus value (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D+C-</td>
<td>(1.0)(.33)=.33</td>
<td>(0.0)(.33)=0</td>
</tr>
<tr>
<td>C+B-</td>
<td>(1.0)(.33)=.33</td>
<td>(0.0)(.33)=0</td>
</tr>
<tr>
<td>D+B-</td>
<td>*(0.5)(.33)=.17</td>
<td>*(0.5)(.33)=.17</td>
</tr>
</tbody>
</table>

*B and D have never been paired together, so stimulus values for this pair are going to be equal.

The total stimulus value for D is (.33+.17)=.50; the total stimulus value for B is (0.0+.17)=.17. The total stimulus value for C is .33. In test trials, when B is compared with D, the stimulus C would have to be inferred as a referent. Because half of the C choices during training rule out responses to B and the other half of the choice response confirm D directly; all of the stimulus value of C gets transferred to D during the BD choice phase. Thus, the new stimulus values for B and D are B=.17, D=.83. These data coincide with the findings obtained by McGonigle and Chalmers. However, the model has been criticized by Zentall and Sherburne (1994) as being unparsimonious because the triad midpoint must be inferred.

In a recent publication, Wynne (1995) outlined several different models based on the Rescorla-Wagner theory. The Bush-Mosteller model can be summarized as follows:

\[
V(X)_{i+1} = V(X)_i + UB(1-V(X)_i) \text{ for reward,}
\]

\[
V(X)_{i+1} = V(X)_i - DB \times V(X)_i \text{ for nonreward}
\]
\[ V(X) = \text{the stimulus value of } X \text{ on a given trial}. \]
\[ DB = \text{the growth parameter for nonreward}. \]
\[ UB = \text{the growth parameter for reward}. \]

The Bush-Mosteller model will generate the typical negatively accelerated curve (similar to the Rescorla-Wagner model) as the stimulus value approaches asymptotic value. In a binary choice situation, choice will be determined by the relative values of the stimuli presented (i.e. Herrnstein’s matching function). For example:
\[ P(X|Y) = \frac{V(X)}{V(X) + V(Y)} \]

Couvillon and Bitterman (1986) proposed a model that utilizes the output from the formula above and adapts it. Although the model worked well for predicting choice in bees (Couvillon & Bitterman, 1986, 1992) the original formula is simpler and much more predictive when applied to data obtained from pigeons. Wynne generated predictions using the Bush-Mosteller model and compared the predicted values against the data obtained from Ferson, Wynne, Delius, and Staddon (1991) in their study of transitive inference in pigeons. Predictions from the model were consistent with the data from Ferson et al. The model accurately predicted percentage of correct responses for each pair during the training phase. In addition, for tests of transitive inference, the model showed which stimulus would be selected in a nonadjacent pairing. The model was also able to account for the symbolic distance effect (SDE). However, there are some problems with the Bush-Mosteller model. Although the model is able to accurately predict transitive choice and choice during the training phase; the model loses its ability to predict if the order of stimulus presentation is reversed. For instance, the
typical presentation is as follows: A+B−, B+C−, C+D−, D+E−. The reversed presentation is: E−D+, D−C+, C−B+, B−A+. An important point to keep in mind is that the value for growth rate parameter for nonreinforcement is much larger than the value of the growth rate parameter for reinforcement. Consider what is occurring when the stimuli are presented in reversed order. During the E−D+ presentation, the animal learns very quickly to choose D+ and to never choose E−. In this first pairing, D+ is always reinforced and E− is never reinforced. Thus, the D+ stimulus attains a high positive value. During the second stimulus pairing (C+D−) the animal should learn to choose C+. However, during the first few trials, the animal will choose D− because of its former high positive value. Also, because the growth rate parameter for nonreinforcement is so large, the D− loses associative strength very quickly. This process continues until the animal is presented with the final stimulus pair (A+B−). Initially, the animal will choose B− because of its positive value from the previous stimulus pairing. During these first few trials, the associative strength for B− will drop to the point where essentially B = C = D = E. Thus, the model predicts that responding will be random during the transitive choice phase. In fact, the only choice the animal will consistently make is the A+ during an A+B− pairing. The actual data obtained from Ferson et al. indicate that pigeons are capable of accurately making transitive choices when the order of stimulus presentation is reversed.

Another theoretical model that has been used to explain transitive inference is the value transfer theory (VTT). Fersen et al. (1991) explain that a given stimulus value is dependent on the sum of its direct and indirect values. During the initial training procedure, the
animal is presented with four pairs of stimuli (AB, BC, CD, DE). The
anchor points (A and E) do not have the same direct stimulus values as
B, C, and D because A is always reinforced and E is never reinforced.
However, B, C, and D are each reinforced during 50% of the pairings.
Therefore, B, C, and D all have the same direct value. According to VTT,
when stimuli are paired during training, a small portion of the positive
value associated with the reinforced stimulus is transferred to the
nonreinforced stimulus (indirect value).

\[ V_i = V_i+1 + a*V_i+1 \]

1. \[ V(A) = 2.0 \]
2. \[ V(B) = 1.0 + 2a \]
3. \[ V(C) = 1.0 + a + 2a^2 \]
4. \[ V(D) = 1.0 + a + a^2 + 2 \]
5. \[ V(E) = 0.0 + a + a^2 + a^3 + 2a^4 \]

\( a < 0.5 \)

\( V_i = \text{Composite stimulus value} \)

\( R_i = \text{Direct stimulus value} \)

\( V_i+1 = \text{Composite stimulus value of the rewarded stimulus when presented with stimulus } i. \)

(formula taken from Fersen et al. 1991)

Note that according to the above equations, B, C, and D all have the
same direct value, but the indirect values for B, C, and D differ
(B>C>D). Because the amount of indirect value transfer is dependent on
the composite value of the rewarded stimulus; the amount of transfer
will decrease as the distance from the first-item anchor increases. (see formula) For example, the value transfer from A to B is much larger than the value transfer from C to D. Thus, value transfer theory can predict accurately for transitive choice. The composite value of B (direct + indirect value) is larger than the composite value for D. Value transfer theory can also accurately predict choice during the training phase. If large differences in stimulus value are easier to discriminate than small differences, then it would be expected that response accuracy should increase as differences in paired stimulus values increase.

<table>
<thead>
<tr>
<th>Value difference</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>V(A) - V(B) =</td>
<td>2</td>
</tr>
<tr>
<td>V(B) - V(C) =</td>
<td>3</td>
</tr>
<tr>
<td>V(C) - V(D) =</td>
<td>4</td>
</tr>
<tr>
<td>V(D) - V(E) =</td>
<td>1</td>
</tr>
</tbody>
</table>

The largest stimulus value difference is between D and E. Thus, this pair is the easiest to discriminate and consequently, learning for this pair occurs very quickly. The values for C and D are similar, making this pair more difficult to learn. As can be seen from the value obtained by subtracting the nonrewarded stimulus value from the rewarded stimulus value, each pair can be ranked in terms of discriminability. Rank values based on stimulus differences easily predict for the serial position effect that is often observed during training. Value transfer theory can also easily predict for the SDE as stimulus value differences continue to increase as the number of items that separate the stimulus pair increase.
According to VTT, animals do not need to seriate items or develop logical inferences to solve transitivity problems. Zentall and Sherburne (1994) developed an experimental procedure to test whether previously nonpaired stimuli could be judged relative to one another in the absence of a linear hierarchy. Stimulus A was paired with stimulus B, and A was reinforced 100% of the time and B was never reinforced. Stimulus C was paired with stimulus D, and stimulus C was reinforced 50% of the time and stimulus D was never reinforced. What should the animal choose if stimulus B is paired with stimulus D? If the animal normally solves problems such as this by mentally ordering relative stimulus values into a linear hierarchy, then it would be expected that the animal respond in a random fashion. On the other hand, value transfer theory would predict that stimulus B should be chosen preferentially over stimulus D. Even though B and D are never reinforced (the direct values are the same); positive transfer from A to B should be larger than the positive transfer from C to D (different indirect values). The results supported value transfer theory (see also Steirn, Weaver, & Zentall, 1995). Pigeons preferred stimulus B to stimulus D. Although the data support VTT, Zentall and Sherburne point out that the mechanism of transfer has not yet been identified.

One explanation is that generalization may be occurring. That is, when the S+ gets reinforced, the S- (which is part of the context) is also associated with reinforcement. A second possibility is that higher-order conditioning is taking place. For example, when S+ and S- are presented to the subject simultaneously, the animal will see the S+ first on some trials, but on other trials the animal will see the S- first. Thus, the S- becomes a predictor for the presence of the S+.
Zentall and Sherburne suggest that one could test these two hypotheses by devaluing the S+. If context is responsible, then there will be no immediate effect on the S-. However, if value transfer is the result of higher-order conditioning, then preference for S- should be reduced because the source of its value is its link to the S+. In a recent publication, Zentall, Sherburne, Roper, and Kraemer (1996) tested these hypotheses by including a stimulus devaluation procedure in which stimulus A+ could be devalued relative to stimulus B- following the initial training procedure. During the devaluation procedure, stimulus A+ was presented singly for a period of 6 seconds. Responses to stimulus A+ were not reinforced, thus decreasing the direct value of stimulus A+. If an association between the reward and stimulus B- is responsible for preferential responding to B in a B-D pairing, then it would be expected that the devaluation of A+ should have no effect on the B response. However, if B- becomes associated with A+ during training, then preferential responding to B during a B-D pairing should be reduced following the devaluation of A+. In fact, Zentall et al. (1996) show that when A+ is devalued, preferential responding in a B-D pairing is eliminated. It is then suggested that higher-order conditioning may be the mechanism by which indirect value is transferred from the rewarded stimulus in an adjacent stimulus pairing. While this explanation may fit well with choice preference described by Zentall et al., it does not explain choice preferences when stimuli are trained in a linear fashion. For example, when A+ is paired with B- and B+ is paired with C-, B becomes predictive of both an S+ and an S-. Thus, it seems unlikely that an association with A+ could account for
preferential responding to B (in a B-D test pairing) if training occurs with linearly arranged stimuli.

A problem with evaluating the validity of VTT is the inherent difficulty in developing an experiment in which a linear representation and VTT yield different outcomes. In a recent publication, Treichler and Van Tilberg (1996) tested the ability of rhesus macaques to arrange two 5-item lists that had been linked to form a single 10-item linearly arranged list. First, Treichler and Van Tilberg trained six rhesus macaques on a list of four conditional pairs in a manner that allowed for a linear formation (A-B+, B-C+, C-D+, D-E+). Similarly, the monkeys were also trained on a second list of four conditional pairs (F-G+, G-H+, H-I+, I-J+). After both lists had been trained, each animal was trained to discriminate between the end anchor of one list (E-) and the first-item anchor of the other list (F+) such that a single 10-item list was formed. A unique feature of this preparation is that several internal test items can be generated as opposed to the single test pair that is formed from the standard 5-item series. In this experiment, the monkeys performed at above chance levels when presented with intra and interlist test pairs. This poses an interesting problem for value transfer theory, especially when items in analogous positions on each list are compared (I-D). Theoretically, the stimulus values should be the same (or quite similar), yet the monkeys were still able to choose the correct item at above chance levels. It may be argued that the link training may have lowered the stimulus value on the first-item anchor of the second list enough to account for the difference. However, the only way that this would account for the preference is if stimulus value difference resulted in an all-or-none pattern of responding in favor of
the stimulus with the higher value. As pointed out by Fersen, Delius, & Staddon (1991), stimulus preference occurs along a continuum that is dependent on the degree to which stimulus values differ.

Transitive inference has been found in a number of different species including rats (Davis, 1992; Roberts & Phelps, 1994), pigeons, (Fersen, Wynne, Delius & Staddon, 1991), and monkeys (McGonigle & Chalmers, 1977, 1992; Treichler & Van Tilberg; 1996). It seems logical that species with abstract representational capabilities may show different response patterns with regard to transitive inference tasks than do species that are deemed less likely to possess these capabilities. However, this does not seem to be the case. Monkeys and pigeons both show similar response patterns during training as well as similar response patterns during transitive choice phases. Thus, many of the mathematical models that are currently being used to predict choice are equally predictive for both pigeons and monkeys. This has led some researchers to conclude that similar response patterns during training and test phases may result from similar mechanisms (see Couvillon & Bitterman, 1992; Fersen, Wynne, Delius & Staddon, 1991; Wynne, 1995). Although the evidence provided in the transitive inference literature does not make it easy to discount this assumption; there is considerable evidence that similar response patterns across species may reflect insensitivity that is inherent in the preparation.

Serial learning tasks have proven useful in identifying differences in representational ability across species. The serial learning tasks described here all utilize a simultaneous chain preparation in which all of the stimuli are presented simultaneously in a randomly arranged spatial configuration for each trial. During
training, the first item of the list is presented singly, and the animal is reinforced for responding to that item. Next, the animal is presented with the first two items of the list, and reinforcement is given for responding to the items in the correct order. For example, item A and item B are presented simultaneously, and the animal must respond first to A, and secondly to B before reinforcement is given. This process continues until the animal has learned to successfully complete the entire list in the correct order.

An important feature of the simultaneous chain preparation is that the random configuration across trials eliminates the possibility that the animal could be responding according to a reinforced chain of physical responses. Additionally, there are no physical cues inherent in the stimuli that provide the animal with any information as to the current location within the chain. For example, when responding to A, there are no physical cues present that suggest a response to B is the next correct choice (Straub & Terrace, 1981; Terrace, 1993).

When pigeons are trained using a simultaneous chain preparation, they are capable of learning to respond to a 5-item sequence correctly. That is, they are able to accurately place the five items in the appropriate serial positions (A-B-C-D-E). However, when the pigeons are presented with a two-item subset that is selected from the original list, the pigeons' performance begins to break down. Generally, pigeons respond at uniformly high levels across all subsets with the exception of the interior pairs. That is, pigeons are able to respond at above chance levels if the subsets contain at least one anchor item (A or E), but if both subset items are selected from the interior, then response accuracy falls to chance levels.
Terrace, Chen and Newman (1995) explain that pigeons do not form linear representations of serial lists. Rather the pigeons performance (on an intact 5-item list) can be explained in terms of directional associations. For example, response to item A elicits a response to item B which in turn elicits a response to item C and so on. Although this may be an adequate explanation of performance on an intact 5-term series, it does not explain the pigeons performance on subset pairings. Terrace (1991, 1993) suggests that pigeons response patterns on subset pairings can be accounted for by a few simple decision rules:

1. Always respond to item A first.
2. Respond to item E (or whatever the end-anchor point happens to be) last.
3. Respond to any other item(s) by default.

As can be seen, subset performance can be fully accounted for by these decision rules. Pigeons will respond accurately to subsets that contain beginning or endpoint anchors, yet response to interior items will not exceed chance levels. This suggests that pigeons do not form linear representations of 5-item lists.

Further evidence that pigeons may utilize simple decision rules to organize linear data was provided by Terrace (1991); (see also Terrace & Chen, 1991a; 1991b). Pigeons were shown the typical 5-term series in which stimuli were presented simultaneously (simultaneous chain). However, unlike previous serial learning tasks, Terrace (1991) arranged the list in a manner that would facilitate chunking of the list in memory. For example, pigeons were presented with items A-B-C-D′-E′
where unprimed letters represent colored icons and primed letters represent achromatic geometric forms. As can be seen, the unidimensional stimuli are arranged such that icons sharing similar dimensional properties are chunked in a group of three (for colors) and a group of two (for achromatic geometric forms). This arrangement allows for utilization of Terrace's simple decision rules that were outlined earlier. Consistent with this notion, Terrace shows that pigeons not only perform well on subsets containing anchor points, but they also perform well on subsets containing interior items. Terrace suggests that the pigeons are using the same decision rules as discussed earlier, but they are amended to fit the inclusion of a second two-item list (achromatic geometric forms) following the presentation of the three-item list (colored icons). Terrace offers the following rules to explain subset performance:

1. Locate and respond to the first item in chunk1.
2. Respond to the end-item anchor in chunk1 last.
3. Respond to any other item in chunk1 by default.
4. Locate and respond to the first item in chunk2.
5. Respond to the end-item anchor in chunk2 last.
6. Respond to any other item in chunk2 by default.

Once again, these simple decision rules predict the higher than chance performance on all subset pairs, including interior subset pairs. Thus, a linear representation of the data is not necessary to account for the pigeons performance on subset pairs.

There is some evidence however, that pigeons do have some knowledge of ordinal position; they just may not have the memory span
necessary to maintain more than three items at a time in working memory (although Terrace's study regarding chunked lists seems to suggest otherwise). Terrace (1986) conducted a study in which pigeons were required to learn a three-item serial list presented in a simultaneous chain format (A-B-C). Once pigeons had reached criterion on the first list, they were then trained with two new items (X-Y) and one item included from the first list. The item from the first list was then placed in three different serial positions (for example: A-X-Y, X-A-Y, X-Y-A). If pigeons possessed some knowledge of ordinal position, it would be expected that there should be negative transfer when the ordinal position of the item from list 1 was altered when presented in list 2. In fact, Terrace did find that negative transfer occurred when an item from list 1 had its ordinal position changed in list 2. This finding is difficult to explain using Terrace's simple decision rules, particularly for negative transfer with the B item. Supposedly, the B item should get a default response which would imply that knowledge of ordinal position is not needed.

Similar findings have been found in studies that employ a wildcard as a substitute for one (or more) item(s) in a serial list. A wildcard is a stimulus that has no specific serial position but can occupy the position of any item in a serial list: W-B-C, A-W-C, A-B-W. If the animal has some knowledge of ordinal position, then the animal should be able to substitute the wildcard for the missing item. Terrace (1995) found that pigeons presented with a wildcard in a three-item series can successfully position the wildcard in the sequence. However, Terrace explains that the pigeons performance on wildcard trials is probably due to new associations forming between the wildcard and the original list.
items. Although pigeons perform well on a three-item series, performance drops to below chance levels on four and five-item lists except when the wildcard replaces the last item in the series. Terrace argues that for three-item lists, the pigeons are only required to form one new association for W-B transition; no associations for A-W-C (the wildcard separates two salient anchor points); and no associations for A-B-W as the wildcard can receive a response by default. However, for lists that contain more than three items, more than one association needs to be formed (A-B-W-D) except in cases where the wildcard occupies the position of either anchor. The pigeons responded at above chance levels on four and five-item lists when the end-item anchor was replaced with a wildcard. When the first-item anchor was replaced with a wild-card, responding did not exceed chance levels, but response accuracy was higher than with wildcards that replaced interior items.

Several differences between monkeys and pigeons have been found with regard to performance on serial learning tasks. D'amato and Colombo (1988) trained Cebus monkeys (Cebus appella) to perform a serial learning task in which the monkeys were required to place five serial items in the correct order. Not surprisingly, the monkeys were able to perform the serial task at 75% accuracy criterion in an average of 36 sessions (40 trials per session). The monkeys were then tested with all of the subsets that could be generated from the five-term series. Unlike the pigeons (see Terrace, 1993; Terrace & McGonigle, 1994), the monkeys were able to perform at above chance levels on all subsets, including subsets drawn from interior items. In addition to subset differences, there were also differences with regard to response patterns. In monkeys, first-item response latencies increased
monotonically as distance from the first-item anchor increased. Monkeys showed similar latency increases as the distance between the items of the subset increased. For example, response latencies between the first response (B) and the second response (C) tend to increase as the number of intervening items increase (B-D as opposed to B-C). Pigeons on the other hand do not show any of these latency patterns. D'amato and Colombo suggest that the latency differences may reflect differences in the representational abilities between the two species. The latencies that monkeys show may be the result of an internal linear representation of the data. D'amato and Colombo (1989) argue that the latency data is consistent with the notion that monkeys form an internal associative chain that can be accessed when monkeys are presented with subsets of larger lists. For example, when monkeys are presented with a subset pair (B-D), they can access their internal analogue until they have a match (a-B-c-D). The seek time associated with this type of mental processing should follow a monotonic latency function, which is exactly what occurs.

One problem with the notion that an internal associative chain may be responsible for subset performance is that it implies that monkeys have no knowledge of ordinal position. That is, a elicits b which elicits c and so on. This process does not require the same type of abstract reasoning power that would be expected for a monkey with knowledge of ordinal position (NOTE: lower case letters refer to internal representations). D'amato and Colombo (1989) trained Cebus monkeys to perform a five-term series using a simultaneous chaining preparation. Next, monkeys were trained with wildcard trials in which one baseline item was replaced with a wildcard item such that five new
lists were created (example: W-B-C-D-E, A-W-C-D-E, A-B-W-D-E, A-B-C-W-E, A-B-C-D-W). An associative chain account would have difficulty explaining above chance performance on wildcard trails because it seems unlikely that new associations would be able form for each new wildcard position. In fact, the monkeys were able to perform at above chance levels at each wildcard position. Furthermore, the pattern of errors observed is not consistent with the pattern of errors expected from an associative chain account. It would be expected (based on an associative chain account) that A-B-C-D-W or W-B-C-D-E would be the easiest sequences to learn because only one new association is required for each, yet this does not seem to be the case.

Does this mean that monkeys do not organize lists according an associative chain? Not necessarily, there is some evidence that suggests that monkeys will utilize the simplest strategy available to them when organizing linear data. Therefore, they may encode the data in terms of ordinal position in some instances or they may represent the data in terms of an associative chain in other instances (see D'amato & Colombo, 1990). When examining the latency data, it is interesting to note that latency patterns change for monkeys depending on the type of training procedure used. Monkeys show an SDE when trained on a serial learning task using a simultaneous chain. Rather than having shorter latencies as the distance between subset items increase; response latency tends to increase which suggests that the monkeys may be using an internal associative chain to represent the data. On the other hand, monkeys also show an SDE during subset tests when they are forced to discriminate between adjacent pairs during training. These latency differences may reflect representational differences that are occurring
in each instance. For example, when monkeys are trained using a simultaneous chain, the simplest strategy needed to solve all internal subset pairs is an internal associative representation. Knowledge of ordinal position is not necessary. On the other hand, when monkeys are given a conditional discrimination task on adjacent pairs, they are forced to infer the ordinal position of each item in the list. This can easily account for the SDE in that monkeys are able to extract positional information directly from each item in the list rather than having to follow an associative chain until the correct items are found.

The argument presented here is that monkeys are capable of representing serial lists in a linear fashion. Although mathematical models are capable of predicting transitive choice outcomes, this does not mean that the assumptions inherent in the models are correct. As discussed earlier, pigeons and monkeys exhibit similar response patterns when they are trained on adjacent-item subsets of a larger (linear) list. However, this does not mean that pigeons and monkeys represent their lists in the same fashion. There is considerable evidence that monkeys are able to utilize different list learning strategies depending on the task demands (D'amato & Colombo, 1990). Also, it has been shown that monkeys develop new list learning strategies as their level of experience with serial lists increases (Swartz, Chen, & Terrace 1991). Based on this evidence, it is hypothesized that value transfer theory is not an adequate means of explaining transitive choice in monkeys. Specifically, it is predicted that prior linear ordering experience with a given set of stimuli (TI training) should transfer to other linear ordering tasks, thus resulting in facilitation of performance (i.e. on a wildcard task). On the other hand, prior experience with a set of
stimuli that do not possess any inherent qualities that suggest a particular linear arrangement (regardless of their stimulus value arrangement) should not facilitate performance for other linear ordering tasks.

Method

Experiment 1

The current experiment was designed to examine whether the subjects were able to attain stable, above-chance performance on a five-item simultaneous chain. The second objective was to determine the potential underlying mechanisms responsible for performance on a five-item chain; especially in comparison to the mechanisms postulated for simultaneous chain performance in pigeons.

Subjects Two experimentally naive 5 year-old male rhesus macaques (Macaca mulatta) were used in the present experiment. The monkeys were pair-housed for the duration of the experiment and were supplied with food and ad-libitum access to water.

Apparatus The pair-housing unit served as the testing chamber for each animal (61cm X 92cm X 61cm). Attached to the front of each chamber was a cart containing a video monitor, response apparatus, feeder, and video camera. The monitor was placed approximately 15cm from the front of the chamber. The response apparatus was a Groovytouch EZscreen touchscreen, on which five computer-generated icons were presented. All icons were white and they consisted of an hourglass shape, a circle, a "plus" sign, a square, and a triangle. A Plexiglas template was placed over the touchscreen to prevent "accidental" responses to the wrong icons as sometimes occurs simply by dragging a limb across the screen inadvertently. A feeder dispensed reinforcers to a food cup located
near the bottom of the monitor. The monkeys were monitored using a video camera that was mounted on the top of the cart.

**Procedure**

The icons were presented in one of eight randomly selected locations on the screen. Monkeys were initially shaped to touch the first icon (item A). Following the shaping phase, monkeys were required to respond first to icon A and second to icon B. After one 200 trial session of at least 80% accuracy, monkeys were trained to respond first to item A, second to item B, and third to item C until 80% criterion was met. This process continued until the monkeys were able to reproduce the entire five item list at 80% criterion during a single session. Reinforcement was not delivered until the terminal icon was selected. Each trial was separated by a 15 second intertrial interval (ITI), and incorrect responses were punished with an additional 5 second time-out. As a feedback to indicate a correct response, the icon flashed when the monkey selected the correct icon, and when an error was committed, the screen flashed white for 5 seconds, which was followed by the usual 15 second ITI. An incorrect response occurred when monkeys either made a backward error (ABA) or a forward error (ABD); however, repeat responses were not penalized (ABBBC). During the testing phase, all possible subset pairs (two icons) and triplets (three icons) were selected from the list and presented to the monkeys. Dependent measures included response latency and accuracy.

**Results and Discussion**

As expected, both Bob and Abe were able to respond at a criterion level of 75% on the five-item simultaneous chain. These data are similar to the data reported by Terrace (1991) in which pigeons were
able to accurately solve a five-item simultaneous chain. However, analyses of subset data from the series reveal substantial differences in performance, especially with respect to interior non-anchor items. During both four (Straub & Terrace, 1981) and five-item (Terrace, 1991) phases, pigeons responses did not exceed chance levels with interior subset items. The monkeys on the other hand (from the present experiment), performed at near ceiling levels on all possible subset pairs. Both monkeys received 12 trials for each of the possible 10 pairs for a total of 120 trials, and both monkeys responded accurately on 119 of the 120 possible trials (Bob's only error was with pair AB, and Abe's only error was with pair BC, representing a 91.7% level of accuracy for the respective pairs).

Figure 1 represents the monkeys' performance on triplet subsets from the five-item series. Similar to the subset-pair data, both monkeys performed at well above chance levels on the triplet items (90% for Abe and 95.8% for Bob). Terrace's pigeons (1981) were given similar tests involving triplet subsets drawn from a four-item series. Although the pigeons performed at levels that were beyond the levels expected by chance (ranging from 39% to 66%); it is clear that the monkeys performance was better than the performance demonstrated by the pigeons. Based on these figures, it is tempting to argue that the difference between the pigeon and monkey performances is simply a matter of magnitude as opposed to differences occurring at the representational level. However, an important difference between the two procedures is that Terrace's pigeons were tested with subsets from a four-item series, whereas the monkeys were tested with subsets from a five-item series. This distinction is important in that a five-item series eliminates the
possibility that the utilization of simple decision rules is responsible for high levels of performance on all triplet subsets. Consider Terrace’s pigeon data; based on the decision rules Terrace has previously defined (and as described earlier in this manuscript), it is a simple matter to derive a set of predictions regarding pigeon performance on triplet items. It can be predicted that triplet subsets ABD and ABC would yield the highest level of performance based on the decision rules. Selection of the icon A yields an associative response to icon B, which in turn will yield an associative response to C. Alternatively, the triplet ABD would also be easily solved by applying the rules: respond to item A first, and respond to item D last, and respond to anything else by default. In this case it seems that performance would be exceptionally high given that the only internal item is a direct associative link from item A (unlike subset ACD, which can still be solved by utilization of the decision rules, albeit not as efficiently due to the associative break). It seems that the most difficult subset to solve is the BCD pair, mostly because there is no associative cue indicating which icon should receive the initial response. Unlike the other three triplets, whose first item A may be primed for a response due to an association between trial onset and item A, there is no similar association eliciting a response to B. Hence, it would be predicted that performance on triplet BCD would be extremely low relative to the other triplet items. Examination of Terrace’s data supports these predictions strongly in that a ranking of performances shows that subset ABD was the most easily solved (65.94%), followed by subsets ABC (53.57%), ACD (41.73%), and BCD (39.13%).
Examination of the present data show that monkeys do not demonstrate the same response patterns shown by pigeons on triplet item subsets, which suggests that perhaps monkeys represent serial order in a fundamentally different manner than do pigeons. If monkeys were to utilize similar decision rules when presented with novel triplet subsets, it would be expected that accuracy would be highest with triplets containing a first-item anchor, followed by triplets containing an end-item anchor. However, the data do not support this. An analysis of Abe's data shows that the majority of all errors occurred on triplet subsets containing a first-item anchor (75%). Additionally, performance on subsets beginning with the second item in the series (item B) was higher than performance on subsets containing a first-item anchor (93.7% versus 87.5%). A possible explanation for this pattern may be that the monkeys are correctly choosing item A in the triplet subsets, yet show preferential degradation on the latter subsets due to the increases in subset distance associated with first-item anchor triplets. If this were the case, it would be expected that selection errors would be biased toward second and third item selection errors following the correct selection of item A. However, error patterns show that the incorrect selection of the second item was primarily responsible for errors of commission for triplets beginning with item A, accounting for 83% of all errors. It is also interesting to note that there was no evidence for preferential responding to the end-item anchor during errors of commission for triplets containing item E. In fact, three of the five errors were the result of incorrect responses to the second item in the triplet, whereas a premature response to the end-item accounted for the remaining two errors.
Probably the most compelling evidence to suggest representational differences between the pigeons and monkeys is demonstrated by performance on interior subsets versus subsets containing both anchor points. Based on the mechanistic rules postulated for pigeons, it can be predicted which subsets would present the most difficulty, and which subsets would present the least difficulty. Because subset ABE contains both end-item anchors in addition to a direct association between A and B; this subset pair would represent the simplest of the subsets according to Terrace's decision rules. By contrast, pair BCD would represent the most difficult subset of the series because it does not contain a last-item anchor, nor does it contain a first-item anchor. The missing first-item anchor is especially significant in that there is no cue suggesting where in the list the animal should initially respond. Thus, performance on the BCD subset should be much lower than any other triplet in the series, and performance on the ABE subset should be higher than any other triplet in the series. These predictions are incongruent with the data obtained from the monkeys. Both Bob and Abe performed at higher levels on subset BCD than on subset ABE (combined scores for subset ABE were 83.3% as compared to 95.8% for subset BCD). These general error patterns are completely inconsistent with the explanations offered for pigeon performance, and strongly suggest an alternative representational structure to account for monkey performance on serial lists.

One manner in which serial order representation has been conceptualized is in terms of an internal associative chain (D'amato & Colombo, 1988). This is essentially a representational analogue of an associative chain in which all stimuli are physically present during the
problem solving task. Take for example, a serial order task in which a series of arbitrary stimuli (labeled ABCDE for simplicity) are trained in a serial fashion. As described earlier, the initiation of the trial signals a response to A, and the selection of A signals a response to B and so on. However, the internal analogue of this process suggests that a response to each item is not contingent on the stimulus itself, but rather the representation of the stimulus. For instance, the representation of item A signals the representation of item B and so on down the list (a-b-c-d-e)².

It is clear from the displayed error patterns that pigeons do not form an internal analogue of an associative chain—but do monkeys? The high level of performance shown by monkeys on interior subset items (both pairs and triplets) suggests an internal representational structure, but the error patterns alone do not provide any real evidence as to the form of the associated representational structure. However, given the inherent properties of an associative chain, some predictions regarding latency can be generated to examine the possibility of an internal analogue. In addition to the prediction of superior performance on interior subset items (relative to pigeons), it can also be predicted that varying distances between subsets will yield differences in response latency to subset items derived from the series. Specifically, it can be predicted that as distance from the first-item anchor increases, so will the latency to respond to the first item in the subset. The logic behind this prediction is that the initiation of the trial produces a representation of the first item (item A), at which point the monkey must proceed through the list until a match is found between the stimulus and the represented item. The time needed to
travel down the list is what yields the associated increases in response latency for the first subset item when distance from the first-item anchor increases. It can also be predicted that similar increases in response latency will occur as the separation between subset pairs (or triplets) increases.

Analysis of latencies for subset pairs demonstrate that both Bob and Abe show increased response latencies as the distance between the first subset item and the first-item anchor point increase. Latencies for first-item subset responses and their concomitant distances are presented in figure 2: 1.49, 1.77, 2.56, 3.02 and 3.07, 4.92, 6.13, 10.76 seconds represent the latencies to respond to each item with increasing distances from the first-item anchor for Bob and Abe respectively (numbers in parentheses are the standard errors). Although Bob and Abe show differences in terms of the absolute values associated with their latencies, the differences between them are in magnitude only—both monkeys show the same pattern of increasing latencies as distance from the first-item anchor increases. In a similar fashion, Bob and Abe show steady increases in latency as the distance from the first item in the subset pairing and the second item in the subset pairing increase (figure 3): 1.05, 1.07, 1.31, 1.52 and 1.49, 1.41, 2.13, 2.50 seconds for Bob and Abe respectively. While these data are not quite as consistent as the first-item response latency data, they do show a steady increase in latency as distances between subset items widen. The data also provide a compelling argument for an internal associative representation in that 15 of the possible 16 means are consistent with the pattern predicted by an internal associative chaining account.
A more rigorous challenge to the internal associative chain account is presented by the triplet item subsets. For triplet items, there are three possible distance measurements (which will be referred to in this text as "steps"): the distance from the first-item anchor and the first item in the triplet subset (Step 1); the distance from the first triplet item and the second triplet item (Step 2); and the distance between the second triplet item and the third triplet item (Step 3). If monkeys are solving subset items via an internal associative analogue, then similar latency patterns should emerge from the triplet item subsets. The data are presented in figure 4. The latencies for Step 1 are presented in order of ascending distances (in seconds): 1.68, 2.34, 3.16 for Bob, and 3.62, 3.61, 6.70 for Abe. The latencies for Step 2: 1.07, 1.68, 2.45 for Bob and 1.23, 1.81, 3.12 (.212) for Abe. Finally, the latencies for Step 3 are as follows: .860, .900, 1.27 for Bob, and .98, 1.05, 1.09 for Abe.

These data provide strong evidence for the associative chain analogue explanation regarding the nature of the representational structure. At each step in the sequence, increases in subset differences also resulted in concomitant increases in reaction time. As can be seen from the provided data, 17 of the 18 means were consistent with the predictions posited by the internal associative chaining account. Terrace's pigeons on the other hand, demonstrate no such latency patterns with increasing distance. Although pigeons do respond more quickly to item A during subset tests relative to other subset items, latencies are similar across varying distances from the first-item anchor to the first subset item as well as for the varying distances between subset pair items. The favorable bias for item A (in
terms of latency) is probably due to the pigeons' predisposition to always choose "A" first during subset pairings.

It is interesting to note that as monkeys performance improves on a list, there are reliable shifts in error patterns, but before error patterns are discussed, it is important to clarify how errors will be described in this text. Errors are described as occurring in terms of the number of "steps" by which they can differ from the correct selection. For example, a one-step forward error occurs when the correct selection is skipped in favor of the item following the correct selection in the list (A-B-D for example). The same occurs with a one-step backward error (A-B-A for example). While the majority of all errors are of the one-step forward variety (probably due to remote forward associations), the number of multiple-step errors are frequent enough to warrant a comparative analysis at varying performance levels.

Analyses of error patterns are particularly interesting because there are certain predictions that can be derived based on the idea of a burgeoning associative chain that strengthens with continued list experience. First of all, whether internal or external (an internal analogue should possess similar properties), associations will develop between all items contained within the list according to the following principles: (a) obviously, directional associations will develop between adjacent items within the list, (b) associations will develop between remote items (nonadjacent items), but association strength will decrease with increased separation distance, (c) forward associations will generally be stronger than backward associations (see Slamecka, 1985). Thus, it might be expected that as experience with a particular list increases, the strength of the internal associations will continue to
increase as well, resulting in a set of predictable shifts in error patterns. Specifically, as the strength of the associative chain increases, there should be a systematic decrease in the proportion of backward errors (with a bias toward one-step backward errors in terms of the shift from two and three step backward errors), and an increase in the proportion of one-step forward errors, possibly at the expense of other forward error types.

In order to explore this possibility, error patterns from the first and last three days of training during the five-item phase of the simultaneous chain presentation were examined (see Table 1). As can clearly be seen, the number of one-step forward errors increased substantially in nearly every viable position (seven increases of a possible eight), while one-step backward errors also tended to decrease in all but one viable position (the one failure is probably due to a floor effect). The only shift that cannot be accounted for is the exceptionally high proportion of two-step backward errors by Abe at position four. Overall however, these data are consistent with the notion that an associative chain is developing with increased experience with the five-item list, and in combination with the latency data, suggest that perhaps the strengthening of the associative chain is occurring at the representational level.

It appears that while performance on complete serial lists is similar for monkeys and pigeons, this similarity is not reflective of similar representational mechanisms. Comparisons across the various subset tests suggest that performance on serial order tasks may be governed by an internal representation of that respective list, whereas pigeon performance may be the result of a series of discriminative
mechanisms suggesting to which icon next to respond. Furthermore, the
data also suggest that as monkeys acquire more experience with a given
list, there may be changes occurring internally that facilitate list
performance. A second experiment was conducted in order to further
elucidate the mechanisms responsible for serial order performance.

**Experiment 2**

In order to further examine some of the potential mechanisms
associated with performance during simultaneous chain presentations, a
wildcard item (an item that replaces any other items within the list at
random) was introduced into the list. Thus, if the monkeys are able to
form an internal representation of the list, then they should be able to
utilize information derived from the representation (either associative,
ordinal, or perhaps positional information) to accurately position the
wildcard within the sequence.

**Method**

**Subjects:** Same as in Experiment 1.

**Apparatus:** Same as in Experiment 1.

**Procedure:** The same serial list utilized in experiment 1 was used in
this second experiment. Both monkeys were retrained starting with the
three-item series from experiment 1 (ABC). Once performance had reached
an 80% criterion level, a wildcard was introduced (a white "W" shaped
icon). The wildcard replaced baseline items at random on about one-half
of all trials such that the sequences WBC,AWC, and ABW were possible.
In order to obtain reinforcement, the monkeys were required to respond
to all baseline items correctly, and during wildcard trials, the monkeys
were also required to replace the missing baseline item with the
wildcard. Once criterion had been reached, monkeys were retrained on
the four-item baseline list, and when an 80% criterion was reached (see results section regarding criterion values), a wildcard was again placed in the list to replace a baseline item at random. The same process was implemented for five-items as well. The reinforcement and time-out protocols were the same as in Experiment 1.

Results and Discussion

Although both Bob and Abe were able to satisfy the accuracy criterion of 80% during the three-item wildcard phase, neither was able to satisfy the 80% accuracy criterion for the four and five item phase. Because of this, new criteria were set in order to provide a standard by which future performance could be judged. At this point, it is important to reiterate that the function of this particular task was to: (a) determine whether the monkeys could perform at above chance levels on a wildcard task, and (b) to provide a standard by which the post-treatment wildcard performance could be judged (meaning that criteria standards were performance based). Therefore, satisfactory performance levels for the four-item phase were determined to be any combination of 80, 80, 70, and 40 percent on any of the wildcard positions for Bob, and any combination of 80, 80, 80, and 70 percent on any of the wildcard positions for Abe. Similarly, satisfactory performance levels for the five-item wildcard phase were determined to be 80, 60, 60, and 50 percent for any four of the five wildcard positions for Bob, and 80, 80, 80, 60, and 60 percent on any of the five wildcard positions for Abe. Before continuing with the analysis, it is important to point out that in order to perform at a satisfactory level, these criterion levels had to be maintained over the course of 200 trials (about half of which are wildcard trials). This means that
for each wildcard, performance had to be maintained for 2-5 times as many trials as required by the Cebus monkeys used by D'amato and Colombo (1989). This requirement was installed to ensure that performance was stable and that reliable comparisons could be made during the post-treatment wildcard phase.

The number of sessions necessary for each of the monkeys to satisfy the performance requirement for the various list lengths are as follows: for lists ABC, ABCD, and ABCDE, Bob required a total of 5, 12, and 11 sessions to meet criterion levels; Abe required 6, 8, and 6 sessions to meet criterion levels for each of the three respective lists. Table 2 shows wildcard performances at each position within the sequence, and it also shows overall baseline and wildcard performances during the criterion session. In order to accurately evaluate performance on wildcard trials, one must assume that the monkeys are not randomly guessing with respect to the baseline-item sequence, especially given their high levels of performance on baseline trials. Therefore, if baseline performances exceed levels predicted by chance (and clearly they do), and if the monkeys are simply guessing as to where to position the wildcard in the series, then it is simple enough to predict that wildcard performance will probably hover around 20 percent for each wildcard position. However, as can be seen from table 2, overall performance was well above chance levels. In addition, performance levels on each of the various positions exceeded chance levels as well. This suggests that the monkeys are not only able to respond accurately on baseline trials, but are able to accurately substitute a wildcard for a missing item within the list. The question of course, is whether a
prerequisite ordinal representation of the list is necessary for accurate wildcard performance.

D'amato and Colombo (1988) demonstrated that Cebus monkeys could accurately respond to interior (non-anchored) subset items from a five-item serial list. However, D'amato and Colombo also suggest that the mechanisms responsible for accurate performance may be an internal associative analogue of the list itself. Although this suggests an internal representation of the list, it also suggests that the animal has no knowledge of ordinal position. That is, in order to determine the appropriate response, the animal must first access the internal representation of the first item (item "a"), which will elicit a representation of the second item (item "b") and so on until each stimulus is matched with its' represented counterpart. This may provide an effective means for solving the task, but if this internal associative mechanism is responsible for subset performance, then a knowledge of ordinal position is not required in order to solve the task.

During a subsequent study, D'amato and Colombo (1989) introduced a wildcard into a five-item sequence in order to eliminate the possibility that monkeys were utilizing an internal associative representation to solve a serial order task. The idea was that a random substitution of a list item would render it nearly impossible to solve wildcard trials due to an excessive number of new associations needed for each wildcard position. Take for example, a five-item series represented as ABWDE, where the "W" represents the wildcard item replacing item "C". Because random substitutions would theoretically result in associations of equal strength between the wildcard and baseline items, remote forward
associations would be required in order to accurately position the wildcard. In the above example, accurate transitioning between the wildcard and the following baseline item would be dependent on the relative strengths of remote forward associations. That is, the strength of the BD association would have to be strong enough to prevent an inaccurate transition to item E. Additionally, the association between baseline items and the wildcard have to be strong enough relative to remote forward associations (BD in the above example) to ensure accurate baseline-wildcard transitions (B-W above); yet not too strong, otherwise the wildcard could interfere with baseline-baseline transitions (A-B above). In essence, these assumptions, in addition to the myriad new associations required, render the associative chaining account untenable. Thus, D'amato and Colombo conclude that the Cebus monkeys were able to solve wildcard trials by means of an ordinal representation of the list items.

The problem with the argument postulated by D'amato and Colombo is that despite their contentions, the associative chaining account can explain wildcard performance. Take for example nonadjacent subset performance; in this case, monkeys are able to access their representation of the list until the appropriate matches are found between the actual stimuli and their represented counterparts. If monkeys can learn to respond to specific items when a match is found, then why does it seem so implausible that monkeys can respond to a wildcard item when a match is not found? As part of their objection to the associative chaining account, D'amato and Colombo are assuming that the wildcard item is represented in the same way that other baseline items in the list are represented, and as such, the wildcard becomes
incorporated into the list. However, it is equally possible that the monkeys are treating the wildcard as a special case, and learn to respond to the wildcard when a match is not found between the represented item and the displayed stimuli. Thus, although the wildcard data provide additional evidence for an internal representation of serial order, the data do not permit a dismissal of the associative chain account in favor of an ordinal representation.

As mentioned earlier, the data are consistent with the data of D'amato and Colombo (1989) in that the rhesus monkeys were also able to accurately position the wildcard in each of the five positions at above chance levels. However, performance on wildcard lists was predictably lower than performance on baseline lists. Of particular interest, are the source of the wildcard errors. It seems reasonable that many of the errors may be the result of difficulties associated with accurately positioning the wildcard within the series of baseline items. It might be expected, based on an associative chain account, that wildcard items on either anchor of the list would be more easily solved due to fewer transitions between baseline and wildcard items. However, as can be seen in Table 2, this is not the case. In fact, it appears as if the fifth position wildcard item presented the most difficulty for both Bob and Abe, with the fifth position wildcard representing the worst performance of any position during 67% of all sessions for Abe, and 100% of all sessions for Bob.

It is of interest then, to determine whether the source of error on wildcard trials was due to difficulties associated with accurate positioning of the wildcard, or whether there were other major sources of error contributing to the unexpected error patterns. Of course,
errors seriating list items manifest themselves as incorrect transitions from one item to the next item, and as such, one can determine the probability that a correct transition will occur. Take for example, the transition from item B to item C; the probability of a correct transition from B to C is calculated given that a correct transition has previously occurred between items A and B. If positioning of the wildcard were the primary source of error, then it would be expected that most of the errors on wildcard trials stem from incorrect transitions both to and from the wildcard item. Examination of this possibility can be accomplished by combining the substituted transitional probability values associated with transitions to and from the wildcard, with transitional probabilities among the corresponding non-replaced items from the baseline trials. These probabilities can then be multiplied to arrive at the expected value that would result if the only source of error in the sequence was the direct consequence of inaccurate positioning of the wildcard.

As can be seen in Figure 5, both Bob and Abe show similarities between the expected and actual performance levels for the four-item series. Although this would seem to suggest that the most discernible source of error for the first three wildcard positions is directly related to positioning of the wildcard; the similarities between the actual values for wildcard performance in each of the positions and the overall baseline performance suggest that differential sources of error do not exist. Although this is especially true of Abe’s data, Bob on the other hand, shows slightly different error patterns. Bob’s performance level for the first position is clearly lower than baseline performance level, yet is nearly identical to the predicted value. This
suggests that most of the error for first-position wildcard trials stems from the inability to properly insert the wildcard into the correct location within the sequence. It is also interesting to note that Bob's performance on the third-position wildcard trials is lower than baseline and expected performance levels. Furthermore, the expected performance level is nearly identical to the combined mean of baseline and third-position wildcard trials, suggesting that difficulties for third-position wildcard trials were the result of inaccurate positioning of the wildcard in addition to disruption of baseline to baseline transitions.

Unlike the first three wildcard positions, both Bob and Abe's performance level for sequences containing a wildcard in the fourth position was much lower than the performance level for baseline sequences. Although performances for both monkeys show similar degradation during fourth-position wildcard trials, the sources of error during these trials appear to be slightly different. Analysis of Bob’s data show that actual performance during wildcard trials was much lower than the expected performance levels, yet the expected performance levels were nearly identical to performance during baseline trials. This suggests that nearly all of the error during fourth-position wildcard trials can be attributed to difficulties that occurred during baseline-baseline transitions prior to actually reaching the wildcard itself. Abe’s performance on the other hand, shows moderate deviation from both expected and baseline performances, yet the expected value is closer to baseline levels than to actual wildcard levels. This suggests that while both baseline transition errors and wildcard positioning errors are responsible for the sub-baseline performance during
fourth-position wildcard trials; the majority of the error can be attributed to earlier difficulties associated with transitions between baseline items.

Analysis of Bob's performance during the five item wildcard phase shows substantial deviations from expected performance levels in every wildcard position except for the second position. Conversely, in each of these positions, the deviation between expected wildcard performance levels and baseline performance levels are relatively small. This suggests that most of the error in these positions can be attributed to poor accuracy with regard to baseline-baseline transitions. Abe on the other hand, showed substantial deviations from baseline performance for wildcard positions three and four, and a slight deviation at position five. Curiously, while most of Bob's errors were attributable to disruption of baseline performance in all wildcard positions, almost all of Abe's errors in the third position were attributable to difficulties positioning the wildcard in the sequence. Furthermore, error patterns continued to change reliably when shifting from the third to the fourth position, and finally to the fifth position, with the number of baseline-baseline transition errors tending to increase relatively, while the number of wildcard positioning errors tended to decrease.

The question of course, is whether an internal associative chaining mechanism can accommodate the described data, or whether the ordinal representation posited by Colombo and D'Amato is the more appropriate characterization. As described earlier, D'Amato and Colombo argue that wildcard trials would be extremely difficult to accommodate given an associative chain interpretation, but if above chance performance were to develop (with the requisite assumptions regarding
new associations and their relative strengths), performance should be highest for the wildcard replacing the end-item anchor (item E).

Although it is argued here that Colombo and D’amato’s objections to the associative chain interpretation are unfounded, the same prediction (albeit different mechanisms) would seem to hold true regarding performance level on the end-item anchor. That is, given the postulates: (a) monkeys are able to internally represent serial order via an internal associative chain, and (b) monkeys are able to respond to a wildcard item when a baseline item cannot be found; it makes sense that the fewer number of baseline-wildcard (and vice versa) transitions required, the fewer possible occasions for error. For example, consider the series ABWDE. In order to accurately position the wildcard, the monkeys must proceed through their representational analog until a response is given to item B, at which point the monkeys should realize that item C is missing. Because a match cannot be found for the representational analogue of item C, a response to the wildcard must be made, and hence, a successful transition from item B to the wildcard. Accessing the representation of item C should then elicit the representation of item D, at which point a match is found, and the successful transition from the wildcard to item D occurs.

On the other hand, when considering the series ABCDW, there is only one transition between the baseline items and the wildcard item, meaning that there are fewer potent sources of error. Based on these arguments, there should be a serial position effect, with wildcard positions one and five (for a five-term series) representing the highest performance levels, and the three middle positions representing the lowest performance levels. However, this does not seem to be the case.
As mentioned earlier, performance level for the end-item anchor was lower than for the other four positions in most cases (67% for Abe, and 100% for Bob). So, if the error patterns displayed by Bob and Abe are inconsistent with the notion of an internal associative chain, how are they representing the list?

As discussed earlier, it is unlikely that the wildcard is being treated as simply another static icon within the series, and consequently, new associations are not required for each wildcard position. The other possibility (and as suggested by D'amato and Colombo (1989), is that the monkeys are representing the list ordinally. That is, the monkeys know which ordinal position each item occupies and can use this information to accurately position the wildcard within the sequence. This type of representation may also help explain why monkeys show preferential performance degradation for end-item anchors. For instance, the monkeys may be able to more precisely position the wildcard by locating baseline items that occupy ordinal positions directly before and directly after the wildcard item, but may show difficulty with anchors because of the decreased amount of available information that is derived from the baseline items. Although this interpretation is consistent with Bob’s data and partially consistent with Abe’s data (Abe did well on first-item anchors), it is still possible that the monkeys are accessing an internal associative chain during wildcard trials.

One of the problems with dismissing the associative chain account, is that given the meager evidence against it, one would have to assume that the nature of the representation is unidimensional, and that any problem-solving strategies employed do not deviate from the underlying
representational structure. However, it is entirely possible that wildcard performance occurs by means of a primary representational structure in combination with a set of subordinate rules. Perhaps some useful insight can be gained by examining some of the properties associated with wildcard trials that might contribute to additional rule-based responding. First of all, one of the major problems with D’amato and Colombo’s (1989) characterization of the wildcard trials was that they failed to consider the fundamental differences between baseline and wildcard items. Mainly, in their paper, their objection to the associative chaining account is based on the notion that wildcard items are treated in a similar fashion to baseline items. However, it seems perfectly reasonable that monkeys recognize: (a) that baseline items are static or ordinarily stable from trial to trial, (b) the wildcard item is dynamic in nature, and that its ordinal position changes from trial to trial, (c) the probability of error generally increases with each non-wildcard response, and relatedly, the probability of error increases with the passage of time following the initial response. That is, with regard to this last postulate, 80% (4/5) of all wildcard responses will occur prior to the fifth item, and because temporal and proprioceptive cues are related to the number of responses produced, they may also provide useful information as to which item warrants the next response in the series.

As suggested by the above characterization of the wildcard list, it may be that monkeys are able to use this information to develop a set of supplementary rules that are used in conjunction with their internal representation of the list. It seems plausible, that the monkeys may learn the following general rules: (a) respond to a specific baseline
icon only during a specific temporal or physiological (determined via proprioceptive feedback) window; (b) once an item has been given a response, do not respond to that item again; and (c) do not wait too long, or offer too many responses before selecting the wildcard. Obviously, appropriate application of rule (b) will be partially dependent on the precision of the information given by rule (a), which provides useful feedback for baseline items, but does not provide any substantive feedback for wildcard items.

So, according to these rules, specific predictions regarding error patterns can be proposed. Of course, given that the wildcard is not consistently located in any specific position, it is impossible to use temporal or proprioceptive cues as reliable indicators as to general position within the list. However, the probabilistic heuristic described in rule (c) suggests that baseline error patterns will be related to the position of the wildcard within the sequence. Because wildcard items located in the latter part of the list require several baseline-baseline transitions prior to reaching the wildcard, the majority of the error associated with these wildcard trials will be the result of premature selections of the wildcard. Additionally, it can be predicted that with each baseline-baseline transition that occurs prior to reaching the wildcard, the probability of prematurely selecting the wildcard will increase, thus resulting in a steady decrease in transition probabilities as movement occurs through the list toward the wildcard item.

As can be seen in Table 3, Bob shows a steady decrease in transitional probabilities as he travels through the list toward the wildcard item for each viable wildcard position (positions 3, 4, and 5).
during the criterion session. Further analyses were conducted on trials with fourth and fifth position wildcards in order to determine whether the major source of error could be attributed to premature wildcard selection. Indeed, 12 of Bob's 13 errors for the fifth position wildcard trials were the result of premature selection of the wildcard item, and all 8 of Bob's errors during fourth position wildcard trials were the result of premature selection of the wildcard. While Bob's data are in support of the subordinate decision rules, Abe's data are not. Examination of the error patterns produced by Abe show no decreasing probability values with forward transitions prior to reaching the wildcard, and neither does Abe show a substantial number of forward transition errors in general for fifth position wildcard trials (Abe showed three as opposed to Bob's twelve).

This suggests that perhaps the supplementary decision rules utilized by Bob may be specific to Bob, and that Abe may not be employing a similar strategy. This also suggests that perhaps the utilization of the indicated decision rules may be a relatively unstable phenomenon, and may not be shown by other monkeys in general. However, it is also possible that Abe is subject to the decision rules as well, but through continued experience has learned to avoid premature responses to the wildcard item (or to resist the urge to respond prematurely to the wildcard item). As stated earlier, Abe's performance levels on fifth-item wildcard trials was relatively poor overall, with the wildcard representing the poorest performance during 67% of all sessions. This is in contrast with Abe's performance during the criterion session (Figure 5) during which performance on fifth-item wildcard trials was higher than performance for both third- and
fourth-item wildcard trials. Similarly, Abe's fifth-item wildcard performance during the session immediately prior to the criterion session was higher than all other wildcard trials except for second-item wildcard trials. In light of the poor fifth-item wildcard performance for the first four sessions, and given an overall change in error patterns for the last two sessions, it is entirely possible that the performance shift demonstrated by Abe in the last two sessions was the result of a decreasing reliance on the supplementary decision rules relative to the degree that Abe was relying on the internal representation of the list, which in effect, was manifested in terms of decreasing precipitative responses to the wildcard.

If the observed shift in Abe's error patterns was the result of an attenuation of the rule-based component of the response system, it would be expected that Abe's error patterns prior to the performance shift would mirror both the earlier and later performance patterns demonstrated by Bob. Table 4 shows transition probabilities during the two middle sessions for the five-item wildcard phase (sessions three and four for Abe, and sessions five and six for Bob). As can be seen from the data, Abe's error patterns during the middle two sessions are similar to the error patterns demonstrated by Bob during the middle two sessions and during the criterion session. When the wildcard item occupied a position toward the end of the list, the probability of making an incorrect transition tended to increase with each successive baseline transition prior to selection of the wildcard. It is also interesting to note that while wildcard position affected baseline-baseline transition probabilities, transitions to and from the wildcard were unaffected by wildcard position (Figure 6). More
specifically, baseline-baseline transitions were preferentially affected
(relative to transitions to and from wildcard items), and the baseline
transition probabilities were not declining in coincidence with an
overall decrease in transition probabilities in general.

Tables 3 and 4 also demonstrate that the error pattern differences
shown by Abe during the criterion session were not restricted to
successive baseline transitions, but there were also differences in
overall error patterns (as can be seen graphically in Figure 6). That
is, during the criterion session, Abe did not show a steady decrease in
transition probabilities with each successive transition, and neither
did Abe show an overall decrease in the average baseline-baseline
transition probability with latter positioned wildcard items. This
supports the notion that Abe learned to avoid premature responses to
wildcard items, but based on the described evidence, it is difficult to
determine whether the error pattern shift was simply the result of a
decreased reliance on the supplementary decision rules, or whether the
shift was due to the development of a new strategy altogether. If Abe’s
performance during the criterion session resulted from an attenuation of
the decision rules, then it can be predicted that (a) there will be an
overall decrease in the number of errors for trials containing a latter
positioned wildcard item, and (b), of the remaining errors that do
occur, the majority of those errors will reflect some remnants of the
rule based component of the response system. Specifically, with regard
to this last statement, the remaining errors will manifest themselves in
terms of premature responses to the wildcard item.

In fact, during Abe’s criterion session, only three errors were
committed during the fifth-position wildcard phase, and all three of
them were the result of premature responses to the wildcard item. Similarly, only six errors were committed during four-item wildcard phase, three of which were the result of premature responses to the wildcard item. Thus, it seems that the observed changes from the previous error patterns are not from the resultant development of an additional strategy, but rather from the attenuation of a specific component in a multidimensional strategy.

At this point, it seems that a dismissal of D'amato and Colombo's (1988) initial postulation of an internal associative representation for wildcard list items is somewhat premature, especially given that the requisite assumptions necessary for such an argument may not hold up under scrutiny. Probably the most salient of the presented disagreements with respect to D'amato and Colombo's (1989) arguments is concerned with the nature of the wildcard list representation. Given the evidence presented by D'amato and Colombo (and in the present paper), it seems reasonable to conclude that the monkeys are either (a) representing the list in an ordinal fashion, or (b) representing the list as an associative chain, albeit with the addition of a few subordinate rules as a supplementation of the response system. Although D'amato and Colombo argue for an ordinal representation of the list; a more detailed analysis of the present data support the notion that some supplementary decision rules may contribute to overall response patterns, and that deviation from predictions proffered by an associative chain account may result from the consequent use of the supplementary decision rules concomitant with a superordinate internal associative representation. Does this mean that monkeys are incapable of ordinal representation? Not necessarily, it simply means that
monkeys may not represent serial items ordinally when an associative chain representation will suffice.

Experiment 3

Due to the difficulties associated with developing a test capable of demonstrating differential outcomes for cognitive versus noncognitive (namely Value Transfer Theory) mechanisms, both monkeys were trained utilizing a linear and a nonlinear array of icons. Given that reinforcement contingencies can be easily manipulated within the present preparation, it is a simple matter to create a series of icons that bear no linear relationship to one another, yet share similarities in terms of their stimulus value properties (the properties supposedly used to solve transitive inference problems). The net effect is that similarities exist in terms of the salient list properties that are shared among the linear and nonlinear lists, with the major difference being the potential for representation of the linearly arranged list. Therefore, the purpose of the present experiment was to determine whether the monkeys are able to accurately respond during nonadjacent testing of the linearly trained list, and to determine whether the monkeys were able to respond in accordance with VTT following training on the nonlinear list.

Method

Subjects: Same as experiment 1.

Apparatus: Same as experiment 1.

Procedure: During the third experiment, each monkey was randomly assigned to a condition. In one condition (adjacent-pair training), five computer generated icons (A,B,C,D,E) were used as stimuli. All of the icons were the same in terms of shape, but each icon differed with
regard to color (blue, green, red, light blue, and yellow). For each trial, the monkey in the adjacent-pair condition (Abe) was presented with adjacent-pair subsets of the five-term series (AB, BC, CD, DE). In order to receive reinforcement, Abe was required to choose the icon that had the greatest value as defined by A>B>C>D>E, and a correct choice was always be reinforced and an incorrect choice was never reinforced. However, Bob did not receive a linear ordering, but instead received a series of conditional discriminations such that A>B, C>D, E>F, G’>H’. Additionally, for Bob reinforcement frequencies differed depending on the subset type. When responding to subset AB, a correct choice (A) was always reinforced and an incorrect choice (B) was never reinforced. However, when responding to subset CD, a correct choice (C) was reinforced 50% of the time and an incorrect choice (D) was never reinforced. The same contingencies applied with icons EF, and G’H’, with E reinforced 100% of the time and G’ reinforced 50% of the time. The primed letters represent achromatic geometric forms. As shown by Terrace (1991; with pigeons), this type of dimensional shift across (and within) lists will minimize interference while providing Bob with the same number of novel stimulus pairings as Abe. It should be noted that the four conditional pairs presented to Bob are referred to as “lists” only for the sake of simplicity. Criterion in each case was reached when the monkeys responded at 80% or better for each pair over the period of a single 200 trial session. During the initial training, pairs were presented in 10 trial blocks followed (after an 80% criterion for each block) by 2 trial blocks to facilitate learning. Following an 80% criterion for each pair, the 2 trial block requirement was removed allowing for random presentation of the stimulus items. Testing
consisted of a selection of all possible two-item pairings from the list, and dependent measures included measures of accuracy in terms of order, and in terms of latency to respond.

**Results and Discussion**

Given that Bob and Abe did not perform the same task for this third experiment, their data will be discussed separately, with Bob’s data being addressed first. During the nonlinear (or conditional discrimination) training phase, Bob was able to reach the 90% criterion level in only 6 sessions. Also, during the testing phase of the experiment, Bob performed at above chance levels in the direction consistent with the stimulus reinforcement contingencies. That is, when the S+(100) was paired with the S-(0), Bob accurately chose the rewarded stimulus for 100% of all selections (the numbers in parentheses indicate the percentage at which a selection of the specified icon yielded reinforcement). Similarly, for the other training condition, Bob showed preference for the S+(50) during 74.5% of all selections. Of course, the above response patterns are to be expected given that in all conditions, one stimulus is reinforced, and the other stimulus is not reinforced. However, in order to determine whether the separate reinforcement schedules conferred differential positive value on their nonrewarded counterparts, a requisite test pairing of the nonrewarded stimuli from each of the contingencies is needed. Examination of Bob’s preferences during the S-(0) pairings from each of the S+ contingencies reveal a strong preference for the S-(0) that was paired with the S+(100) during the training phase. Indeed, Bob preferred the nonreinforced stimulus paired with the S+(100) relative to the...
nonreinforced stimulus paired with the S+(50) during 82.4% of all selections.

Bob's data are consistent with the pigeon data presented by Zentall and Sherburne (1994) in that Bob showed consistent preference for the S-(0) paired with the S+(100) over the S-(0) paired with the S+(50). Indeed, Zentall and Sherburne's pigeons demonstrated the expected preferences for both rewarded stimuli relative to the nonrewarded stimuli, and they also showed a similar preference for the S-(0) paired with S+(100) relative to the S-(0) paired with the S+(50), with 64.6% of all selections favoring the S-(0) that was previously paired with the S+(100). Zentall and Sherburne explain this preference in terms of a differential transfer of value from the rewarded stimulus to the nonrewarded stimulus, thereby resulting in a greater amount of transferred value to the nonrewarded stimulus from the more frequently rewarded stimulus. Because the task does not easily lend itself to a linear representation (given that there is no basis to link the various conditional pairs with one another), it does not seem reasonable that the mechanisms responsible for such performance are similar to the mechanisms responsible for inferential transitivity. That is, a linear representation cannot be employed in the present situation, so therefore it is postulated here that the results demonstrated by Bob during the S- pairings are the result of differential transfer of value from the rewarded stimulus to the nonrewarded stimulus, and thus they are consistent with the expected outcome as predicted by value transfer theory.

Following 12 sessions of adjacent-pair training, Abe was able to meet the expected criterion level. Acquisition data for Abe were
examined over the course of the last 7 sessions of training (which includes the criterion session), during the period when subset pairs were presented randomly (as opposed to the earlier blocked phase). During the training phase of the current experiment, Abe showed acquisition patterns that were similar to the acquisition patterns demonstrated by Fersen’s et al (1991) pigeons. As can be seen in Figure 9, adjacent subset pairs containing an anchor (AB, DE) were generally more readily acquired than subset pairs that did not contain an end-item anchor (i.e. items from the interior of the list). Also consistent with Fersen’s et al pigeon data, Abe was able to acquire the internal pair BC more readily than the other internal pair CD.

During the testing phase, Abe was able to perform at above chance levels on all subset items, which included all of the previously nonpaired items (see Figure 7). Of course, any true test of transitivity must examine nonadjacent items that are also nonanchor items (in order to avoid reinforcement confounds given that the first-item anchor is always reinforced, and the last-item anchor is never reinforced). So, the interior nonadjacent pairs (B versus D) were examined and showed an above chance level of performance at 93.8% for the correct selection of item B. Analysis of Abe’s latency data (Figure 8) showed a general decrease in latency as the distance between subset items increased. Median response latencies for distances of one, two, three, and four yielded latencies of 1.92, 1.97, 1.78, and 1.66 seconds respectively. Thus, Abe showed a symbolic distance effect in terms of latency across the various subset distances. Although Fersen et al did not provide any latency data for their pigeons during nonadjacent subset testing, they did provide data regarding differences in accuracy at
various distances during the testing of the subsets. Interestingly, Fersen's et al pigeons demonstrated a symbolic distance effect with regard to accuracy level, in that accuracy tended to increase consistently as the distances between subset testing pairs increased. Similarly, Abe also showed greater levels of accuracy for nonadjacent items than for the adjacent subset items (during the testing phase); with a demonstrated accuracy level of 93.8% for adjacent pairs, and 97.9% for nonadjacent pairs.

One of the most salient features with regard to Abe's data is the differences in terms of latency between Abe's performance on nonadjacent subset items from the simultaneous chain training, and the performance on nonadjacent subset items from the adjacent pair training. Because increases in subset distances invariably result in shorter distances between the first-item anchor and the first subset item, it is difficult to compare response latencies to the first subset item across the two conditions (simultaneous chain vs. adjacent-pair training). One of the questions that might be asked (and thus revealing the confound) is whether first-item response latency is due to shorter distances between subset items, or due to the shorter distances between the first-item anchor and the first subset item in the test. One method to get around such problems is to hold distance between subset pairs constant while varying the distance from the first subset item to the first-item anchor, which therefore, allows one to identify the major source of variation with regard to first-item responses in the pair. That is, if decreasing latencies to the first item in the pair were due to a) decreases in the time necessary to travel through the list (associatively) due to decreases in the distance between the first-item
anchor and the first subset item (which is confounded with inter-item distance), or b) decreases in latency due to the symbolic distance effect associated with increased distance between subset-pair items. As would be expected, both lists (simultaneous chain and adjacent-pair) should show a general decrease in first-item response latencies, but there should be differences between lists with regard to first-item response latencies when inter-item distance is held constant. That is, with a distance of "one" between subset pair items, there should be no predictable pattern for first-item response latencies during the adjacent-pair training (with the exception of slight increases in reaction time toward the end of the list due to decreases in perceived distance [to be explained later]), but as distance between the first-item anchor and the first item in the test pair increases for the list trained via a simultaneous chain, there should be predictable increases in latency. As can be seen in Figure 12, this notion is supported. Although no regular pattern occurred with the adjacent-pair training, there were regular increases in latency as the distance between the first-item anchor and the first item in the subset increased. This finding, coupled with the general decreases in latency shown during testing (following adjacent-pair training) suggest that subset distance increases create the symbolic distance effect observed during testing (note the different pattern for first item response latencies when distance is held constant relative to when distance is not held constant in Figure 8).

Based on this evidence, it is clear that performance on test items derived from the two respective training conditions does not result from similar underlying mechanisms. However, at this point it seems
difficult to elucidate the associated mechanisms responsible for the differences. That is, does the underlying mechanism responsible for test performance following adjacent pair training represent a more complex process (or representation) than does test performance following simultaneous chain training? Or does it represent a simpler process? At first it is tempting to argue that test performance following adjacent-pair training would represent a more complex process. Consider the differences between the two training procedures. During a simultaneous chaining preparation, performance on an intact list does not require a high level of representation in order to successfully complete the list; and during the testing phase, above chance performance does not require anything more than a simple recall of the delineated order that was directly provided by the experimenter. However, during adjacent-pair training, the sequential architecture of the list is not readily available from the information provided during training. That is, the experimenter does not directly provide the animal with the correct sequence, but rather, the sequential order has to be inferred from the available information. Thus, it seems reasonable to argue for a more complex process to account for test performance during adjacent-pair training.

However, one of the problems associated with the above argument is related to differences in test performance across the two training conditions. Given the above interpretation, it seems reasonable to contend that performance on subset tests should be higher when the animal is trained with a simultaneous chaining procedure as opposed to an adjacent-pair training procedure. This does not seem to be the case however, especially with regard to the available data for pigeons in the
two training conditions. Fersen et al (1991) show that pigeons are capable of showing transitivity during testing conditions. That is, when nonadjacent items are presented, pigeons are able to respond at above chance levels. On the other hand, as shown by Terrace (1991), pigeons trained with a simultaneous chaining paradigm are only able to respond accurately to the nonadjacent subsets that contain at least one of the end-item anchors in the list, and furthermore, responding to non-anchored items (or interior items) does not differ from the level of responding that would be expected by chance alone.

The differences during testing shown by pigeons across the two training conditions (simultaneous chain versus adjacent pair) suggest that perhaps such an interpretation regarding the requisite representational processes necessary for transitivity may in fact be more simplistic than is described above. If a pigeon is unable to determine the order of nonadjacent test items when the order of the list is clearly delineated during training, then how is it possible that the pigeon is able to determine the order of the list when the order must be inferred from the relationships between the various test pairs? For pigeons, the underlying mechanism may be a much simpler one, and perhaps a mechanism involving inferential processes may not be appropriate; but would such a mechanism be appropriate for monkeys? Given the ease at which monkeys accurately respond to non-adjacent test pairs following training with a simultaneous chain, it certainly seems reasonable to argue for different mechanisms for pigeons and monkeys in this regard. However, although it can be convincingly argued that pigeons and monkeys represent serial order differently (given a simultaneous chain preparation), can it also be argued that different underlying mechanisms
between the species are responsible for test performance following adjacent-pair training? Abe’s data are highly similar to the pigeon data provided by Fersen et al (1991) both in terms of error patterns during acquisition, and in terms of the error patterns shown during testing. At this point however, it is impossible to determine whether Abe’s performance was the result of a higher-order representation of the list itself, or whether his performance resulted from the differential transfer of positive value to the “greater” stimulus during the earlier training phase.

**EXPERIMENT 4**

Given that the monkeys responded as expected during the third experiment, the fourth experiment was designed to determine whether representational differences (between Bob and Abe) were established during the third experiment. Thus, a wildcard task was implemented during the fourth experiment in order to determine whether transfer from the linear task in Experiment 3 would occur with regard to the present task.

**Method**

**Subjects:** Same as experiment 1.

**Apparatus:** Same as experiment 1.

**Procedure:** The procedure was identical to the procedure employed in Experiment 2 with the following exceptions: first, the stimulus icons from the third experiment were used during the current experiment, whereas novel icons were used during Experiment 2. The icons were arranged according to the direction of the linear training (for Abe) and for Bob, the icons were arranged according to descending stimulus value (AECBD, renamed ABCDE in Experiment 4 for simplicity). The other
difference between the methodology here and in Experiment 2, was that the monkeys were only retrained on the baseline sequence ABC (to familiarize them with the task). After 80% criterion for sequence ABC, the four-item sequence with a wildcard was introduced (Further training on baseline items beyond the first three was not performed because the dependent measure was the amount of ordinal transfer from Experiment 3 to Experiment 4), and following criterion performance on the four-item wildcard phase, the five-item wildcard phase was introduced.

Results and Discussion

The results obtained for this experiment are consistent with the hypothesis in that given the existing representational differences conferred from Experiment 3, there should be marked differences between Bob and Abe with regard to the amount of transfer. However, the direction of the difference was not in the expected direction. As expected, Bob's performance did not differ much from his demonstrated performance in experiment 2. For wildcard sequences ABC, ABCD, and ABCDE, Bob required 8, 16, and 13 trials (respectively) in order satisfy the predetermined criterion levels (as compared to 5, 12, and 11 sessions necessary during experiment 2). The slight differences observed in the number of sessions needed to reach criterion between experiments 2 and 4 are most likely due to the fact that Bob did not receive extensive training on baseline items prior to the implementation of the wildcard item in experiment 4 (Bob did receive extensive training on baseline items prior to the implementation of the wildcard in experiment 2). Abe on the other hand showed marked degradation in performance relative to earlier performance in the second experiment. During experiment 4, Abe required 9, 21, and 41 sessions to satisfy the
predetermined criterion levels for wildcard training phases ABC, ABCD, and ABCDE respectively (as opposed to 6, 8, and 6 sessions for experiment 2).

Although facilitation of wildcard performance did not occur for Abe, his performance during experiment 4 is inconsistent with the expected performance given the assumption that value transfer theory (VTT) is the correct underlying mechanism responsible for performance in both cases. That is, both sequences were arranged in a series of descending values as suggested by value transfer theory, and thus, given that VTT is the appropriate underlying mechanism, there should be no differential transfer for Bob and Abe. As can be clearly seen however, Abe shows heavy degradation of performance during experiment 4 (relative to experiment 2), whereas Bob does not show any performance degradation in the fourth experiment.

So, given the substantial difference in the performance of Bob and Abe in the fourth experiment, the question regarding the potential mechanism responsible for the difference remains. One of the seemingly obvious explanations is that the representational differences incurred during the third experiment may have contributed to the overall differences in the fourth experiment. One of the more salient differences between the second and the fourth experiment prior to wildcard implementation was the level of experience that each monkey had with the associated baseline list. During the second experiment, monkeys were well trained on baseline list items prior to being exposed to wildcard trials. In contrast, during the fourth experiment, monkeys were only exposed to the three-item phase (ABC) of the initial training procedure prior to being presented with wildcard trials. Therefore, a
couple of different possibilities exist. First, it is possible that training during the third experiment assisted Bob’s performance by somehow facilitating the representation of the serial list in the fourth experiment; and to account for the differences, the parallel assumption must be that training during the third experiment did not provide Abe with the representational media necessary to facilitate performance on the new task. Examination of the separate tasks in the third experiment suggest that the above scenario is unlikely. In order to delineate the potential differences that each animal carries into the fourth experiment, it is important to look at the meaningful experiential differences in the third experiment; or in other words, to examine the specific characteristics of each condition (from Experiment 3) that may enhance or degrade performance during the fourth experiment. Consider the task with which Bob was presented. Each trial consisted of a series of conditional discriminations that were completely independent of one another. Essentially, a pair of icons appear on the screen, and a selection must be made in which one icon is reinforced and the other icon is not reinforced (with some reinforced stimuli being reinforced at 100% and other reinforced stimuli being reinforced during 50% of all selections). These icons can then be arranged in series of descending values consistent with VTT. Of course, adjacent-pair training (as Abe received in experiment 3) also results in a similar serial arrangement of descending values. Because the pairs are independent of one another during the conditional discrimination phase, it is unlikely that any form of representational process is facilitating performance.

However, in Abe’s case, the potential mechanism for transfer is not limited to the descending arrangement (in terms of value determined
by VTT), but there is also the potential for transfer of the linear representation from the third experiment. Although it is not assumed here that representation is unidimensional in either condition; given that the two mechanisms described (VTT and linear representation) are the two most salient of the contributions to the performance in each condition, the only difference between the adjacent-pair training and the conditional discrimination training is the potential for linear representation of the list trained via adjacent-pair discriminations (in theory, both lists have the potential for differential value transfer). So, this suggests that the other potential mechanism for the difference in performance during the fourth experiment may be of a representational nature. That is, Abe may have been able to form a linear representation of the list in Experiment 3, and the associated representation may have disrupted performance during the fourth experiment.

This may occur when one sort of representation interferes with the development of another. Remember that the conclusion derived from Experiment 2 was that the list was being represented as an associative chain, and that accurate positioning of the wildcard was simply the result of a given response to the wildcard when a match was unable to be found between the stimuli in the array and the representation of the missing item. Similarly, during subset testing conducted in Experiment 1, latencies generally tended to increase with increasing distance between the various subset test items (Figures 2, 3, and 4), which as mentioned earlier, supports the idea that the monkeys were representing the lists in terms of an internal associative analogue of the list itself.
The testing phase during Experiment 3 however, yielded opposite latency patterns relative to those in Experiment 1. Remember that during the nonadjacent-pair testing phase, latencies generally tended to decrease as the distance between subset test-pairs increased. This suggests that although overall accuracy patterns during testing were similar across the two conditions (adjacent-pair training and training via simultaneous chain), the underlying mechanism responsible for the performance in each condition may not be the same across the two conditions. Additionally, it is certainly possible that the primary mechanisms involved in each case may not necessarily translate into enhanced performance when transferred from one condition to the other. As mentioned earlier, it is not assumed here that any of the representational components for a given task are unidimensional, but it seems reasonable that success on a given task may require a primary representational type, and that any of the other subordinate representational components must be weak enough to prevent interference with the primary mode of representation. In terms of the current experiment, the argument presented here is that the type of representation formed in experiment 3 did not provide Abe with the necessary information needed to accurately solve wildcard trials (the nature of representation formed in Experiment 3 will be discussed later). Furthermore, it is also postulated here that the type of representation formed in Experiment 3 hindered performance in the fourth experiment in that it interfered with (by competing with) the development of an associative representation.

In order to support such a claim, there needs to be evidence suggesting a shift in representation from one task to the next. So, in
order to determine whether there was a shift in representation from Experiment 3 to Experiment 4 (in Abe's case), nonadjacent subset pairs from the baseline component of the wildcard list were selected for testing following the criterion performance on the five-item wildcard list. Under the working hypothesis that a representational shift should manifest itself in terms of a concomitant latency shift across tasks; latencies at all of the possible subset distances were analyzed, as were latencies to respond to the first item in the pair. As indicated in Figure 10, Abe showed a general increase in reaction time as distance between the subset pairs increased with distances of 1, 2, 3, and 4 producing corresponding mean response latencies of .771, .734, .773, .903 seconds, respectively. Abe also showed monotonic increases in latency for first-item responses as the distance from the beginning anchor-point in the list increased, with latencies of 1.61, 2.01, 2.4, and 2.53 seconds for corresponding distances of 1, 2, 3, and 4 respectively. Furthermore, as can be seen in Figure 12, latency to respond to the first item increased steadily when distance was held constant (distance=1), suggesting that first-item response latency was a function of first-item position rather than a function of distance between the subset items.

Clearly, as demonstrated, the patterns shown during the subset tests following wildcard training are in opposition to the patterns shown during subset tests following adjacent-pair training (Figure 8 and Figure 12). Based on this evidence, it appears that in order to solve wildcard trials, Abe required a representational shift from the representation of the stimuli obtained in Experiment 3, to an associative chaining representation for experiment 4. Thus, during the
fourth experiment, Abe showed marked degradation in performance due to competing (and incompatible for the given task) representational structures, while Bob on the other hand, did not form a linear representation prior to being presented with wildcard tasks in Experiment 4, and as a result did not suffer any degradation in performance.

General Discussion

The question remains then as to the nature of the representation incurred during Experiment 3 for Abe. The reaction times certainly suggest that the representation is not an associative chain, but what exactly is the nature of the representation? Similar procedures conducted with humans (Woocher, Glass, & Holyoak 1978) yield similar patterns with regard to reaction time. When trained with adjacent subset-pairs (providing verbal descriptions of relationships such as "the baker is taller than the milkman etc."), it is found that as distance between the tested subset pairs increase, reaction time tends to decrease. One possible explanation that has been offered is that these types of relationships may be represented in terms of a representational analogue of a perceptual process. As part of an experiment examining perceptual discriminations of size, Moyer and Bayer (1976) presented undergraduate college students with pairs of circles of varying sizes, and instructed them to choose the larger of the two circles as quickly and as accurately as possible. It was found that as the difference in size increased (independent of the circles' ordinal relationship with respect to size), reaction time tended to decrease. Interestingly, Moyer and Bayer were able show that an analogous representational process could demonstrate similar reaction time
differences. Specifically, circles of various sizes were linked to consonant-vowel-consonant (CVC) nonsense syllables via a paired associate task. Subjects were then presented with pairs of CVC syllables and were asked to select the CVC that corresponded to the larger circle. Similar to the perceptual judgment task, the subjects demonstrated decreasing reaction times for judgments of CVC pairs during which the relative size differences for the corresponding circles increased (similar to the perceptual judgment); which is interesting in that it suggests there may be analogous processes occurring for perceptual judgments of the physical stimuli themselves, and for similar judgments involving their representational counterparts.

More direct approaches to examining representational analogues of perceptual judgments were offered by Moyer (1973) in an experiment in which subjects were presented with the names of various animals, and were then required to determine which animal was larger. For example, subjects might be asked "Which is larger? An ant, or a flea?" Interestingly, it was found that as the size difference between the animals increased (compare ant and elk versus ant and flea), the reaction time tended to decrease. Again, this suggests that perhaps there are analogous processes occurring in perceptual versus representational judgments. In the present case, rather than size, icons may be judged along some other dimension (perhaps magnitude), arranged linearly and examined in much the same way as a perceptual array.

More precisely, the internal representation of the list may be similar to a perceived spatial array, and that each member of the array occupies its own distinct spatial position. As a result, each member of
the array contains relative positional information, and thus, the animal (or human for that matter) can extract positional information directly from the stimulus itself; and this information then, can be used to compare subsets within the list in order to determine which stimulus occupies a position earlier or later in the list. The symbolic distance effect observed in the present case (the adjacent-pair training), may be the result of positional overlap among the stimuli. The closer (spatially) a pair of stimuli are to one another (either perceptually or representationally), the more difficult they will be to distinguish due to positional overlap, and the farther apart a pair of stimuli are to one another, the more easily distinguishable they are due to decreased positional overlap between the stimuli. As can be readily ascertained, the positional information inherent in each of the stimuli is relatively imprecise, and the precision of positional information may be dependent on where in the list a given stimulus is positioned. For example, Moyer and Landaur (1967) found a symbolic distance effect during numerical judgments similar to the one described in the present paper with regard to adjacent-pair training. That is, when subjects were presented with a pair of digits (the pair drawn from a set of integers ranging from 1-9), and asked to judge the larger of the two stimuli. As subset distances increased between subset pairs, the reaction time tended to decrease.

It is also interesting to note that with regard to numerical discrimination, quantitative differences between the numbers themselves do not necessarily translate linearly into a psychological representation of those differences. For example, the difference between 8 and 9 say, is perceptually smaller than the difference between 1 and 2. This seems to be the case as well with direct perceptual
comparisons. Buckley and Gillman (1974) demonstrate that perceptual
distances between numerosity judgments involving patterns of dots tend
to decrease when the two stimuli increase in magnitude, yet the physical
distance between them is held constant (this relationship can be
captured mathematically by logging the ratio of the larger value
relative to the smaller value subtracted from the larger). In fact, the
numerosity judgments described in the paper by Buckley and Gillman,
follow the properties of the postulated model: a) the psychological
representation of stimulus magnitude is logarithmic (as suggested above)
b) the internal representation of the physical stimuli function as a
random variable with the center of the distribution about the stimulus
magnitude, c) the subject samples (and resamples) stimulus information,
and the difference between the sampled distribution added to a
cumulative counter, and d) once the counter exceeds some threshold (in
either direction), a decision is made.

Therefore, as can be determined from the above description, as a
concurrent increase in magnitude occurs (holding the absolute magnitude
difference constant), the amount of overlap between the stimuli tends to
increase, and thus making the discrimination more difficult. In terms
of the present task, this model can also be applied to positional
overlap. Due to the nature of the question being asked, (for example,
which is larger? A or B?) Perhaps the beginning of the series (in this
case "A") represents a conceptual anchor-point in much the same way that
the number "1" represents a conceptual anchor point in magnitude
discriminations of arabic numerals; and thus, as one travels further
down the list (to selections B, C, and D etc.) the positions of the
icons tend to overlap, and thus increasing the level of difficulty in
terms of selecting the appropriate icon. Of course, pairs AB and DE should be excluded from this analysis due to their salient reinforcement contingencies, but an examination of the remaining pairs shows a similar accuracy pattern as predicted by the perceptual model, with the pair BC>CD in terms of accuracy.

It might also be predicted, that given the variability in perceptual distance involved, wildcard trials might pose a significant problem given a spatial representation. That is, when sampling stimulus information, the imprecision associated with generating a position code may inhibit precise identification of distances of a given size (between test items) that allow for accurate wildcard placement. First, it seems reasonable that during baseline trials, performance should be relatively high, but when the wildcard is introduced, there can be no consistent distance rule utilized to accurately position the wildcard within the sequence because the position codes for each item vary considerably (both with and without missing items). Therefore, it seems reasonable, that given a spatial representation, that judgments of relative order should be fairly simple, yet judgments requiring the utilization of perceptual distance should be extremely difficult (in the case of the wildcard trials). Specifically, it can be predicted that during the four-item wildcard phase (the four-item phase eliminates simple rule-based responding more effectively than the three-item phase) following adjacent pair training (as occurred with Abe), the ratio of baseline correct responses relative to wildcard correct responses should be higher than all of the ratios exhibited during Experiment 2, and should also be higher than the ratios shown by Bob in Experiment 4. As can be clearly seen in Figure 11, this seems to be the case. Although
this trend appears less stable during the five-item wildcard phase, this is to be expected given that the process of re-representation of the list has probably started to occur once criterion performance had been met during criterion performance during the four-item wildcard phase. It should also be noted, that given prior experience with wildcard trials, Abe's previous method of proceeding through the list from the beginning (in an associative chaining fashion) and selecting a wildcard when a given icon in the series could not be found may have interfered with performance in the fourth experiment in that Abe may have attempted to solve the trials in Experiment 4 in the same fashion, and in the process, failing to use the available information (inter-item distance) to solve the wildcard trials. If this is the case, then given the differences shown between Bob and Abe during the fourth experiment, it seems probable that the representation established in Experiment 3 may have interfered with the performance in Experiment 4.

Although above the possibility for a spatial representation seems an attractive explanation, a problem is that much of the evidence for a spatial representation can be accounted for by an ends-inward scanning model (for example, the symbolic distance effect). One of the explanations that has been offered for the observed symbolic distance effect that occurs as subset distance increases is an ends-inward scanning model. Generally, when judging the relative position of two stimuli, a search is initiated at both ends of the list until one of the items is located, at which point (depending on which end the icon was located), a decision will be made. An alternative view, proposed by Holyoak and Patterson (1981), is that each item within the sequence provides direct positional information, and judgments as to relative
position will be influenced by the amount of positional overlap between the items in the sequence. One of the difficulties associated with such hypotheses is the inherent problems associated with comparing these two hypotheses against one another. Given the hypothesis of less positional overlap between nonadjacent items (especially with increasing subset distance), it becomes readily apparent that the two hypotheses are perfectly confounded. How does one examine increased distances between items (and hence decrease positional overlap) without also decreasing scanning time by placing at least one of the subset-items fairly close to one of the anchors? As part of Holyoak and Patterson’s experiment (to examine their model), subjects were asked to judge the relative position of specific items within a visual array. For example, a visual array that consists of several multicolored vertical lines, and the subject may be asked “which line is farther to the right, the red line or the green line?” Similar to the current experiment with the monkeys, subjects typically show similar patterns in terms of decreasing latencies with increasing distances between the compared items, and in terms of the demonstrated serial position curves. According to Holyoak and Patterson, discrimination of position occurs as a two-stage process: a) subjects locate each member of the pair, and then generate a position code for each item. The amount of time required to generate a position code is defined in terms of a “Confusability index” which is essentially the amount of difficulty distinguishing items in the array as a function of target-to-item distance (or subset distance), and the degree to which items are similar to one another (determined by psychophysical properties of the stimulus, and by list position, and calculated via logarithmic transformations of the absolute positional values, and then
the differences between one stimulus and all the others summed (see Murdock, 1960)). The second stage b) is essentially a comparison process similar to the Buckley and Gillman (1974) model described earlier. During this stage, the positional information is sampled (from each specific positional distribution), and values are subtracted and added to a counter. Once the value of the counter exceeds a given threshold, a decision will be made. So essentially, reaction time will be determined by the number of iterations (or re-samplings of the stimuli) required to make a decision multiplied by the summed confusability indexes of the items to be compared (assuming a serial process, with the highest confusability index of the pair replacing the sum given a parallel process).

The utility of this model is that it allows one to directly compare positional discriminability versus an ends-inward scanning model. In their second experiment, Holyoak and Patterson (1981) presented two colored vertical lines placed within a homogeneous array of black lines, and once again subjects were required to determine the relative positions of the icons (in the same way described earlier). This is where predictions of the positional discriminability and an ends-inward scanning model differ. Given that distance is usually confounded with the serial positions of the stimuli, it seems reasonable that a vast reduction in scanning time (as would occur when distinct items are placed within a homogeneous set of black vertical lines), would therefore eliminate the distance effect. On the contrary, the positional discriminability model predicts that although time to locate and generate positional codes would decrease (step a), the comparison process would still affect reaction time due to decreasing positional
overlap with increasing distance. So in essence, the ends-inward model would predict an elimination of the symbolic distance effect, while the positional discriminability model would simply predict an attenuation of the symbolic distance effect. Holyoak and Patterson report that as separation distances increase from one to five, there is a steady decrease in reaction time, thereby supporting the hypothesis and the positional discriminability model. Thus, given the evidence suggesting possible links between representational and perceptual processes, and given the similarity between Holyoak and Patterson's data on perceptual judgments of visual arrays and the present data, it seems reasonable that an effective method of examining mechanisms of representational processes may be to also examine their perceptual counterparts.

In sum, there is strong evidence to suggest that perhaps monkeys are capable of representing linear order in a multi-dimensional fashion, with the characteristics of the underlying representational structure contingent upon task demands. When monkeys are presented with stimuli in a fashion that allow for a simpler level of representation, they will utilize the simplest strategy necessary to solve the particular task (as demonstrated in Experiment 2), and therefore show lower levels of representation. This of course suggests that monkeys are capable of representing linear order multi-dimensionally, and potentially utilize multiple strategies simultaneously when solving a problem involving linear order. Of particular interest, is the possibility that linear order can be represented as a spatial analogue of a perceptual array as some evidence seems to suggest. In any case, it is clear that noncognitive mechanisms such as differential value transfer are inadequate in terms of explaining representation of linear order in
monkeys; and it seems that perhaps a productive avenue of exploration may be to further examine the relationship between representation and perception, particularly with regard to linear arrangements.
References


Footnotes

1. Two of the monkeys (Peeper and Vern) had to be dropped from the study because they were unable to satisfy criterion during the first experiment.

2. Lowercase letters are used to distinguish the physical stimuli (uppercase letters) from their represented counterparts (lowercase letters).

3. Analysis of latencies during correct and incorrect trials reveal that incorrect responses were not the result of precipitative responding in that latencies for incorrect trials were on average .47 seconds longer for Bob and .95 seconds longer for Abe.

4. Because Bob was unable to consistently respond above 40% (and therefore unequivocally above chance levels) on the fifth wildcard position, a fifth performance requirement was not implemented.

5. Because there were so many possible pairings to choose from, only a subset of the testing pairs were chosen for Bob. The chosen pairs were chosen based on their relationship in terms of reinforcement contingencies. Only pairs with differential reinforcement contingencies were paired (except in the case where differential value transfer to the nonreinforced stimulus was being examined.

6. Actually, the stimulus values are arranged according to the following protocol: A=B>C>D>E, (renamed from AECBD) because A and B are reinforced 100% of the time. However, given the ease at which a two-item chain is learned, it wasn’t expected that this minor violation would represent a serious difference between the two lists in terms of their value patterns (defined by VTT).

7. It is interesting to note that the single violation of the pattern (when distance=1) is the same violation (in terms of pattern) exhibited by Abe during the first experiment, which bolsters the claim that Abe was representing the wildcard list in Experiment 3 in the same way he was representing the serial list in experiment 1.
### Table 1

**Error Patterns During the First Three Sessions and the Last Three Sessions of Training**

<table>
<thead>
<tr>
<th></th>
<th>Other backward errors</th>
<th>One-step backward errors</th>
<th>One-step forward errors</th>
<th>Other forward errors</th>
<th>Position in list</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe</td>
<td>NA</td>
<td>NA</td>
<td>70%(100%)</td>
<td>30%(0%)</td>
<td>1</td>
</tr>
<tr>
<td>Bob</td>
<td>NA</td>
<td>NA</td>
<td>78%(100%)</td>
<td>22%(0%)</td>
<td>1</td>
</tr>
<tr>
<td>Abe</td>
<td>NA</td>
<td>0%(0%)</td>
<td>82.5%(90.3%)</td>
<td>17.5%(9.7%)</td>
<td>2</td>
</tr>
<tr>
<td>Bob</td>
<td>NA</td>
<td>0%(0%)</td>
<td>90.3%(97.8%)</td>
<td>9.7%(2.2%)</td>
<td>2</td>
</tr>
<tr>
<td>Abe</td>
<td>0%(0%)</td>
<td>13.6%(2.7%)</td>
<td>67.5%(94.6%)</td>
<td>18.6%(2.7%)</td>
<td>3</td>
</tr>
<tr>
<td>Bob</td>
<td>0%(0%)</td>
<td>2.1%(4.5%)</td>
<td>71.6%(94.0%)</td>
<td>26.3%(1.5%)</td>
<td>3</td>
</tr>
<tr>
<td>Abe</td>
<td>8.8%(23.1%)</td>
<td>11.3%(7.7%)</td>
<td>80.0%(69.2%)</td>
<td>NA</td>
<td>4</td>
</tr>
<tr>
<td>Bob</td>
<td>4.7%(2.5%)</td>
<td>11.7%(5.0%)</td>
<td>83.6%(92.5%)</td>
<td>NA</td>
<td>4</td>
</tr>
<tr>
<td>Abe</td>
<td>43.2%(22.2%)</td>
<td>56.8%(73.7%)</td>
<td>NA</td>
<td>NA</td>
<td>5</td>
</tr>
<tr>
<td>Bob</td>
<td>63.8%(22.2%)</td>
<td>36.2%(77.8%)</td>
<td>NA</td>
<td>NA</td>
<td>5</td>
</tr>
</tbody>
</table>

**Note.** Percentages on the left represent the proportion of errors that fall into the specified category during the first three sessions of training. The percentages in parentheses represent the proportion of errors that fall into the specified category during the last three sessions of training.
Table 2

Percentage Correct Responses at Each Wildcard Position

<table>
<thead>
<tr>
<th>Position</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Overall Wildcard</th>
<th>Overall Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe</td>
<td>96%</td>
<td>91%</td>
<td>96%</td>
<td>71%</td>
<td>89%</td>
<td>91%</td>
</tr>
<tr>
<td>Bob</td>
<td>79%</td>
<td>92%</td>
<td>84%</td>
<td>46%</td>
<td>75%</td>
<td>94%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Position</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Overall Wildcard</th>
<th>Overall Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe</td>
<td>90%</td>
<td>86%</td>
<td>60%</td>
<td>70%</td>
<td>84%</td>
<td>78%</td>
<td>95%</td>
</tr>
<tr>
<td>Bob</td>
<td>55%</td>
<td>86%</td>
<td>60%</td>
<td>65%</td>
<td>32%</td>
<td>60%</td>
<td>93%</td>
</tr>
</tbody>
</table>
Table 3

Transition Probabilities During the Criterion Session of the Five-Item Wildcard Phase

Transition Probabilities for Abe

<table>
<thead>
<tr>
<th>Correct Selection Item A</th>
<th>A-B Transition</th>
<th>B-C Transition</th>
<th>C-D Transition</th>
<th>D-E Transition</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1000 (W)</td>
<td>.950 (W)</td>
<td>.950</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>.950</td>
<td>.950 (W)</td>
<td>1.000 (W)</td>
<td>1.000</td>
<td>.950</td>
</tr>
<tr>
<td>.950</td>
<td>1.000</td>
<td>.740 (W)</td>
<td>.860 (W)</td>
<td>1.000</td>
</tr>
<tr>
<td>.950</td>
<td>.950</td>
<td>.940</td>
<td>.880 (W)</td>
<td>.930</td>
</tr>
<tr>
<td>.950</td>
<td>.890</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000 (W)</td>
</tr>
</tbody>
</table>

Transition Probabilities for Bob

<table>
<thead>
<tr>
<th>Correct Selection Item A</th>
<th>A-B Transition</th>
<th>B-C Transition</th>
<th>C-D Transition</th>
<th>D-E Transition</th>
</tr>
</thead>
<tbody>
<tr>
<td>.950 (W)</td>
<td>.950 (W)</td>
<td>.670</td>
<td>1.000</td>
<td>.920</td>
</tr>
<tr>
<td>.950</td>
<td>1.000 (W)</td>
<td>.900 (W)</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>.850</td>
<td>.760</td>
<td>.920 (W)</td>
<td>1.000 (W)</td>
<td>1.000</td>
</tr>
<tr>
<td>.950</td>
<td>.890</td>
<td>.760</td>
<td>1.000 (W)</td>
<td>1.000 (W)</td>
</tr>
<tr>
<td>.950</td>
<td>.890</td>
<td>.750</td>
<td>.500</td>
<td>1.000 (W)</td>
</tr>
</tbody>
</table>

Note: Transition probabilities during the criterion session of the five-item wildcard phase for Bob and Abe. The "W" in parentheses indicates a wildcard transition.
Table 4

Average Transition Probabilities During the Middle Two Sessions of the Five-Item Wildcard Phase

### Transition Probabilities for Abe

<table>
<thead>
<tr>
<th>Correct Selection Item A</th>
<th>A-B Transition</th>
<th>B-C Transition</th>
<th>C-D Transition</th>
<th>D-E Transition</th>
</tr>
</thead>
<tbody>
<tr>
<td>.920 (W)</td>
<td>.835 (W)</td>
<td>.875</td>
<td>.935</td>
<td>.960</td>
</tr>
<tr>
<td>.975</td>
<td>.975 (W)</td>
<td>.945 (W)</td>
<td>.970</td>
<td>1.000</td>
</tr>
<tr>
<td>.950</td>
<td>.920</td>
<td>.970 (W)</td>
<td>.880 (W)</td>
<td>1.000</td>
</tr>
<tr>
<td>.925</td>
<td>.945</td>
<td>.885</td>
<td>.805 (W)</td>
<td>1.000</td>
</tr>
<tr>
<td>.945</td>
<td>.945</td>
<td>.850</td>
<td>.755</td>
<td>.915 (W)</td>
</tr>
</tbody>
</table>

### Transition Probabilities for Bob

<table>
<thead>
<tr>
<th>Correct Selection Item A</th>
<th>A-B Transition</th>
<th>B-C Transition</th>
<th>C-D Transition</th>
<th>D-E Transition</th>
</tr>
</thead>
<tbody>
<tr>
<td>.845 (W)</td>
<td>.848 (W)</td>
<td>.785</td>
<td>.965</td>
<td>.940</td>
</tr>
<tr>
<td>.930</td>
<td>.840 (W)</td>
<td>.885 (W)</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>.880</td>
<td>.800</td>
<td>.865 (W)</td>
<td>.895 (W)</td>
<td>1.000</td>
</tr>
<tr>
<td>.975</td>
<td>.845</td>
<td>.785</td>
<td>1.000 (W)</td>
<td>.960 (W)</td>
</tr>
<tr>
<td>.890</td>
<td>.785</td>
<td>.710</td>
<td>.275</td>
<td>1.000 (W)</td>
</tr>
</tbody>
</table>

**Note:** Transition probabilities during middle two sessions (sessions 3 and 4 of the five-item wildcard phase for Abe, and sessions 5 and 6 for Bob). The “W” in parentheses indicates a wildcard transition.
Figure Captions

**Figure 1.** The percentage of correct responses to the subset triplets derived from the five-term serial list.

**Figure 2.** Abe’s latencies to respond to the first item of a subset pair as a function of the distance between the first-item anchor and the subset icon occupying the position closest (ordinally) to the first-item anchor.

**Figure 2.** Bob’s latencies to respond to the first item of a subset pair as a function of the distance between the first-item anchor and the subset icon occupying the position closest (ordinally) to the first-item anchor.

**Figure 3.** Bob and Abe’s latencies to respond to the second subset-item following a response to the first subset-item as a function of subset-item separation distances.

**Figure 4.** (A) Bob and Abe’s latencies to respond to the first item in the triplet subset as a function of the distance between the first-item anchor and the subset icon occupying the position closest (ordinally) to the first-item anchor. (B) Bob and Abe’s latencies to respond to the second subset-item (as a function of distance) following a response to the first subset-item. (C) Bob and Abe’s latencies to respond to the third subset-item (as a function of distance) following a response to the second subset-item.

**Figure 5.** Abe(A) and Bob’s(B) actual and expected performance levels for wildcard trials during the four-item wildcard phase during criterion performance. Expected values represent the values to be expected given that the positioning of the wildcard constitutes the major source of error for wildcard trials.

**Figure 5.** Abe(C) and Bob’s(D) actual and expected performance levels for wildcard trials during the five-item wildcard phase during criterion
performance. Expected values represent the values to be expected given that the positioning of the wildcard constitutes the major source of error for wildcard trials.

Figure 5 Abe(E) and Bob’s(F) actual and expected performance levels for wildcard trials during the five-item wildcard phase during the middle two sessions. Expected values represent the values to be expected given that the positioning of the wildcard constitutes the major source of error for wildcard trials.

Figure 6 Abe’s transition probabilities from Baseline to Baseline, Baseline to Wildcard (BL-WC) and Wildcard to Baseline (WC-BL) during the criterion session (B) and during the two middle sessions (A) of the five-item wildcard phase. Bob’s transition probabilities from Baseline to Baseline, Baseline to Wildcard (BL-WC) and Wildcard to Baseline (WC-BL) during the criterion session (D) and during the two middle sessions (C) of the five-item wildcard phase.

Figure 7 Abe’s accuracy level for the 10 two-item subsets derived from the five-item series following adjacent-pair training.

Figure 8 Abe’s latency to respond to the first item of the subset pair as a function of inter-item distance.

Figure 9 Accuracy among the various training pairs during the last 7 sessions of training (during randomized presentation of stimulus pairs).

Figure 10 Abe’s latencies to respond to the second subset-item following a response to the first subset-item as a function of subset-item separation distances (after criterion performance during the five-item wildcard phase).

Figure 10 Abe’s latencies to respond to the first item of a subset pair as a function of the distance between the first-item anchor and the
subset icon occupying the position closest (ordinally) to the first-item anchor (after criterion performance during the five-item wildcard phase).

**Figure 11** Ratio of Bob and Abe's baseline to wildcard performances of the first 10 sessions of Experiment 2 and Experiment 4 during the four-item wildcard phase.

**Figure 11** Ratio of Bob and Abe's baseline to wildcard performances of the first 10 sessions of Experiment 2 and Experiment 4 during the five-item wildcard phase.

**Figure 12 (A)** Latency to respond to the first item in the subset pair while holding subset distance constant. The presented data represent performance during the testing phase of Experiment 1.

**Figure 12 (B)** Latency to respond to the first item in the subset pair while holding subset distance constant. The presented data represent performance during the testing phase of Experiment 3.

**Figure 12 (C)** Latency to respond to the first item in the subset pair while holding subset distance constant. The presented data represent performance during the testing phase of Experiment 4.
One Two Three Four

Separation Distance

<table>
<thead>
<tr>
<th>Abe</th>
<th>Bob</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>2.5</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Seconds