

Retrieval processes for serial order information

JOHN G. SEAMON

Wesleyan University, Middletown, Connecticut 06457

and

JAMES I. CHUMBLEY

University of Massachusetts, Amherst, Massachusetts 01002

Three experiments sought to specify how list structure and rehearsal pattern influence the retrieval of well-learned serial order information. Subjects learned a serial list of 12 words followed by a probed recall task measuring response time. Adjacent list items served as retrieval cues to permit probing of an item by cues which maintained or crossed semantic or rehearsal boundaries. Evidence for structure in serial recall was inferred from the large cue format effects on response time. Such effects were found to be consistent with the semantic relationships in categorized lists and the acquisition rehearsal pattern in unrelated lists. When rehearsal grouping and semantic relatedness were in conflict, the cue format effects conformed mainly to the rehearsal pattern. Extended practice over five sessions did not eliminate these effects for many of the serial items. These results suggest that the structure of the serial list, whether based on previous associations or present rehearsal patterns, can provide a basis for retrieval. A hierarchical search model based on item and order information provided good fits of the data. The model suggested that response time varies with cue formats because cues differ in their efficiency at directing search to the correct response in the list structure. The structure, which is acquired at the time of learning, determines cue efficiency and, hence, the subsequent effects upon response time.

The problem of serial order is how to represent the encoding, storage, and retrieval of temporally ordered information (Murdock, 1974). Since the time of Lashley (1951), it has been realized that a simple associative chaining model is not sufficient to account for performance on serial order tasks. Over the past decade, theoretical interest has turned to concepts such as organization, coding, and pattern detection (e.g., Bower & Winzenz, 1969; Estes, 1972; Johnson, 1970; Restle & Brown, 1970) to either replace or supplement associationism.

Common to many of the current serial order models is the assumption that information may be organized in terms of a hierarchical structure. Consistent with this view are results which indicate that retrieval of category units in organized lists is independent of retrieval of category members (Martin, 1974; Seamon, 1973). Structural effects upon performance have been observed in studies measuring transitional error probabilities (Bower & Winzenz, 1969; Johnson, 1970; Restle & Brown, 1970) or response time (McLean & Gregg, 1967; Seamon, 1973; Wilkes & Kennedy, 1970).

Appreciation is expressed to William K. Estes, Charles E. Clifton, and Jerome L. Myers, who read drafts of this paper. The final version was written while the first author was on sabbatical in the Mathematical Psychology Laboratory of Rockefeller University. The kind support of William K. Estes over this period is gratefully acknowledged. Requests for reprints should be sent to John G. Seamon, Department of Psychology, Wesleyan University, Middletown, Connecticut 06457.

The present research studied retrieval processes for well-learned serial order information. Response time for probed recall was examined in lists composed of subsets of categorically related words. These studies sought to determine if structural effects can be observed in long-term serial recall and to specify how such effects might influence response time.

EXPERIMENT 1

In the first experiment subjects learned a serial list composed of blocked subsets of categorically related words. Adjacent items served as retrieval cues to permit probing of an item by cues which maintained or transgressed category boundaries. Based on previous research which used only the preceding item as a retrieval cue (Wilkes & Kennedy, 1970), it was expected that item response time would be an inverse function of the degree to which the retrieval cues preserved the semantic context of each probed item.

Method

Subjects. Ten introductory psychology students from the University of Massachusetts served individually to fulfill a course requirement.

Apparatus. Stimuli and test trials were presented on slides with a Kodak Carousel projector. Response latencies were obtained in milliseconds from a Hunter KlockCounter timer in concert with a Gerbrands voice key, Lafayette tachistoscope, and the Kodak projector. The closure of a manual switch simul-

taneously opened the projector shutter, started the timer, and activated the voice key. A verbal response from the subject into the voice key terminated a trial.

Procedure. Using the serial recall method, subjects were presented with a list of 12 common words consisting of four groups of three semantically related words (TREE-BUSH-FERN-PONY-GOAT-LAMB-SHOE-BELT-VEST-BONE-SKIN-HAIR). Each word was shown individually for 5 sec, followed by a recall interval of 20 sec. Subjects were instructed to recall the words during the interval by writing them in their correct order on an answer sheet on which the appropriate number of blank spaces was provided. A new answer sheet was used on each trial with no feedback given. A criterion of two successive perfect recall trials was employed.

Following completion of the serial recall task, subjects were presented with test trials on each word in the serial list with response latencies recorded. A test trial consisted of a cue slide showing two words from the list which (1) immediately preceded, (2) bracketed, or (3) immediately followed the test word. Subjects were required to fill in the correct third word as quickly and as accurately as possible. Test trials examined all possible formats in which the original serial order was maintained, thereby crossing category boundaries in the process. For example, (TREE BUSH -----), (BUSH ----- PONY), and (----- PONY GOAT) were the three retrieval cues requiring the response FERN. From the 12 original serial words, 30 different word format combinations were constructed. Four random orders of the 30 combinations, or 120 test trials in all, were presented to each subject with an average interval of 4 sec between trials.

Results and Discussion

The data of primary interest are the correct response latencies to each word format combination. Error trials (5.67%) and response times greater than 8 sec (.25%) were not included. Median latencies were determined for each subject to each combination, with a mean, based on the individual subject medians, then obtained for each combination. Figure 1 presents the mean response

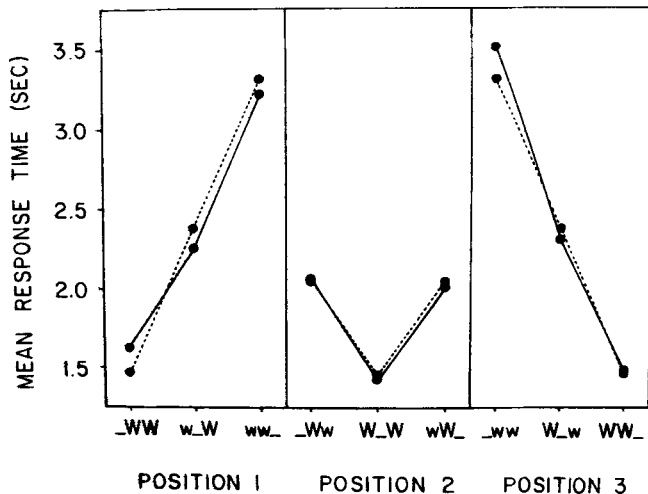


Figure 1. Mean response time as a function of cue format and word position within a semantic category for observed (solid lines) and theoretically derived (dotted lines) data in Experiment 1. Uppercase letters in a cue format designate cue words from the same category as the probed word; lowercase letters designate words from an adjacent category.

time for each item and cue format collapsed over items occupying the same position within each category. It may readily be seen that response time varied with cue format. Cues which presented two members of the same semantic category as the probed item had the shortest latencies, while cues which presented two members of a different semantic category had the longest. These results do not reflect a speed-accuracy tradeoff, as a parsing of intrusion errors by cue formats found errors to vary over cue formats in the same manner as response time.

Separate analyses of variance were conducted on words occupying the first, second, and third positions of the categories and tested on all formats. There was a significant effect of cue format on response time for first [$F(2,18) = 53.72$, $MSe = .31$], second [$F(2,18) = 12.54$, $MSe = .20$], and third position words [$F(2,18) = 36.24$, $MSe = .89$]. With the exception of a significant difference between response times for PONY and SHOE [$F(2,18) = 4.50$, $MSe = .54$], words did not differ reliably, and there were no significant Word by Cue Format interactions. A $p < .05$ rejection region was used in all analyses.

These data show that adjacent serial items are not equally effective retrieval cues in a probed recall task. The more removed a probed item from the context of its category, the longer the latency of response. Latencies were significantly faster to cues which preserved semantic boundaries than to those which crossed category lines. Different interpretations of these results are possible. It may be that response time varied inversely with preexperimental associations; strong associations yielded faster response times than weak associations. Alternatively, subjects may have structured the list during encoding on the basis of the semantic relationships; the best retrieval cues were those which were consistent with the original encoding pattern (Bower & Winzenz, 1969; Tulving & Thomson, 1973).

EXPERIMENT 2

If the previous results are based on encoding patterns, it should be possible to reproduce the response time effects with lists of words that are not categorically related. This experiment presented categorized and uncategorized lists and induced different list structures by means of perceptually grouping the stimuli during learning. The groupings either coincided or conflicted with the semantic relationships in the categorized lists, and induced a structure in the uncategorized lists. The finding of response time effects consistent with the induced grouping in each list would indicate that retrieval cue effectiveness was based on the encoding conditions during acquisition, rather than on preexperimental associative strength. The list structure imposed during learning determines response time during retrieval.

Method

Subjects. The subjects were 60 University of Massachusetts undergraduates, 10 in each condition, who served for course credit. Fourteen of the subjects were replaced for failing to meet a performance criterion.

Apparatus. Display of serial lists and retrieval cues was accomplished by using an Applied Digital Data Systems television monitor controlled by a PDP-8/I computer. Words were presented on the video monitor at a rate of 1 letter/msec. Vocal response times were obtained with a voice key.

Procedure. Three lists of words were used, two categorized lists (PONY-GOAT-LAMB-TREE-BUSH-FERN-SHOE-BELT-VEST-BONE-SKIN-HAIR and MILK-SODA-BEER-JEEP-BOAT-SLED-HAIL-RAIN-SNOW-BELL-DRUM-HORN) and one uncategorized list (BANK-DESK-DAWN-WALL-SLED-RAIN-HAIR-MILK-DRUM-BELT-TREE-PONY). The words were drawn from the Battig and Montague (1969) norms with the constraint that all words have four letters and be common responses to a category name. Categorized and uncategorized lists were used equally often. Each subject had either a different permutation of categories and words within categories or a different permutation of the unrelated words.

The serial recall method was employed again for list learning. Subjects were required to rehearse the words during acquisition in a manner determined by the experimental condition. Type of list, categorized (C) or uncategorized (U), was factorially combined with size (one, two, or three words) of rehearsal group (RG) to yield six experimental groups.

A word list was displayed on the television monitor in groups of one, two, or three words. Subjects rehearsed the word(s) aloud in time with an arrow which appeared beneath the word to be rehearsed. Each word was rehearsed four times. If more than one word was on the screen, the arrow started the rehearsal with the leftmost word and proceeded to indicate rehearsal of the words in a left-to-right order. The first rehearsal was indicated 300 msec following the appearance of the word(s) on the screen, with the arrow moving to the next word to the right 600 msec later. The interval between rehearsing the rightmost word in a group and rehearsing the leftmost word again was 500 msec, and there was a 1,200-msec interval between the final rehearsal in a group and the initial rehearsal in the next group. A criterion of three successive perfect recalls was used.

Following learning, subjects were given eight blocks of test trials. Each block contained the 30 different word format cues, for 240 recall trials in all. The 30 different cues were randomly permuted for each of the eight blocks of test trials for each subject. There was a 3-sec interval between trials and a short break between each block. If subjects failed to make a correct response before a 6-sec deadline, the cue for that trial was repeated at the end of that block of trials. Repeated trials were treated in the same manner as other trials.

Results and Discussion

Subjects who were unable to average 95% correct responses overall or who failed to contribute at least six correct responses for any of the 30 different word format cues were replaced. Use of this criterion resulted in replacing six subjects in Condition C-RG1, and two subjects in each of the other conditions except U-RG3, where it was not necessary to replace anyone. For subjects who met the performance criteria, intrusion error and long time response percentages were: C-RG1, 2.73 and 2.17; U-RG1, 3.23 and 1.87; C-RG2, 4.63 and 2.77; U-RG2, 5.47 and 1.93; C-RG3, 4.14 and 2.76; and U-RG3, 3.02 and 2.78.

The primary results are displayed in Figures 2 and 3.

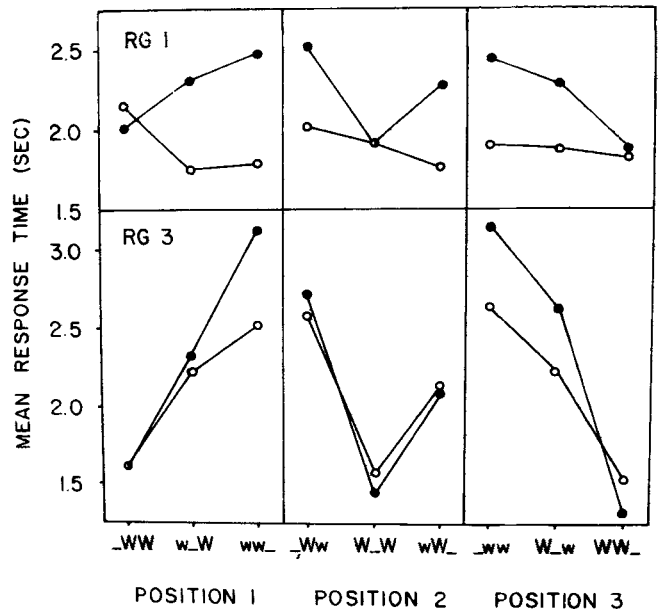


Figure 2. Mean response time as a function of cue format, rehearsal group size, and word position within a semantic category-sized unit for categorized (filled circles) and uncategorized (open circles) lists in Experiment 2.

Mean cue format times for correct responses were determined from the individual subject medians. These means were then collapsed across items occupying the same position in a category. Figure 2 contains the RG1 and RG3 data, while Figure 3 contains the data from RG2. The data in each figure are based on Serial Positions 4 through 9 only. This was done to permit analysis of cue format (F), category position (CP), and position within a category (P) as within-groups variables and

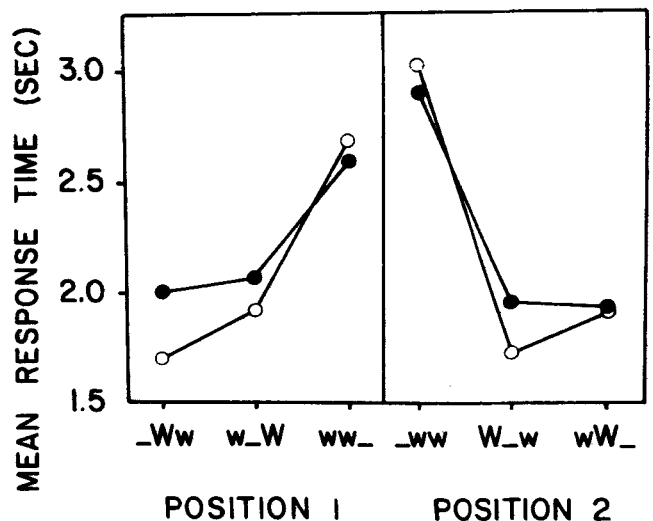


Figure 3. Mean response time for RG2 as a function of cue format and word position within the acquisition grouping for categorized (filled circles) and uncategorized (open circles) lists in Experiment 2.

presence or absence of semantic grouping (S) as a between-groups variable. The data from early and late serial positions were excluded from analysis since these items were not tested on all cue formats. Their exclusion, however, did not change the pattern of results in any manner. In addition, the analysis of the data in Figure 2 included size of rehearsal grouping (R) as a between-groups variable. Serial positions were arranged into groups of three for Conditions C-RG1, U-RG1, C-RG3, and U-RG3 and groups of two for Conditions C-RG2 and U-RG2. The former were based on semantic category size, while the latter were based on rehearsal group size. The reason for this change was that a visual inspection of the data for each serial position indicated that, for Conditions C-RG2 and U-RG2, Serial Positions 4, 6, and 8 (Position 2 in Figure 3) produced very similar results, as did Serial Positions 5, 7, and 9 (Position 1 in Figure 3). For Conditions C-RG1, C-RG3, and U-RG3, Serial Position 4 produced data similar to that of Serial Position 7, Position 5 data similar to that of Position 8, and Position 6 data similar to that of Position 9.

The primary findings shown in Figures 2 and 3 may be summarized as follows: The cue format effect on response time with categorized lists can be observed with uncategorized lists if subjects initially rehearse the serial items in groups of three (Conditions C-RG3 and U-RG3); when semantic categories and rehearsal groups are in conflict, rehearsal grouping is the more powerful variable (Conditions C-RG2 and U-RG2); when serial items are rehearsed individually, a cue format effect exists for categorized lists (Conditions C-RG1 and U-RG1). The latter result may reflect implicit grouping of semantically related words in Condition C-RG1. Oral rehearsal does not preclude noticing stimulus similarities.

The results of the analysis of variance (pooled over both categorized lists) for the data in Figure 2 indicated that only two variables produced significant main effects: semantic grouping [$F(1,36) = 7.27$, $MSe = 1.56$] and cue format [$F(2,72) = 21.13$, $MSe = .21$]. Three interactions detectable in Figure 2 all produced reliable effects: F by P [$F(4,144) = 50.82$, $MSe = .31$]; R by F by P [$F(4,144) = 26.64$, $MSe = .31$], and S by F by P [$F(4,144) = 8.81$, $MSe = .31$]. The source of the semantic grouping effect and the R by F by P and S by F by P interactions can be traced to the recall latencies for Condition U-RG1. Response times for this group were similar to those of Conditions U-RG3 and C-RG3 when the retrieval cue contained one item from the acquisition rehearsal grouping. The lack of a cue format effect indicates that all adjacent items were equally effective as retrieval cues. It may be that items in this list condition were multiply coded with other items both preceding and succeeding them on the list.

The results of the analysis of variance on the data displayed in Figure 3 indicated a significant effect of cue format [$F(2,36) = 34.75$, $MSe = .21$] and a reliable F by

P interaction [$F(2,36) = 90.93$, $MSe = .30$]. If the data grouping procedure used for the analysis of variance had been inappropriate, the F by CP by P interaction ($F < 1$) should have been significant.

In the main, these data provide strong support for the hypothesis that rehearsal patterns employed during learning determine response time to different cues during retrieval. Response time reflects list structure, not preexperimental associative strength. Associative relationships, when present, may serve to suggest a particular encoding scheme at the time of list acquisition.

EXPERIMENT 3

The previous experiments have indicated that response time varies with list structure, but they have not suggested why. One possibility is that a serial list is hierarchically acquired from the bottom up (Martin, 1974). The lower level order of items within categories may be learned before the higher level order of categories. If retrieval is hierarchically ordered from top to bottom (Estes, 1972; Martin, 1974) and response time varies with degree of learning (Seamon, 1972), the previous effects on response time may reflect unequal learning of upper and lower levels in a hierarchical structure. FERN may be produced faster in response to (TREE BUSH -----) than (----- PONY GOAT) because the latter requires knowledge of the higher level order of categories in addition to the lower level order of items within a category. To test this hypothesis, Experiment 1 was replicated over five sessions on the same list of categorized words. Cue format effects on response time should be found in Session 1 and either eliminated or greatly diminished in Session 5.

Method

Subjects. The subjects were six University of Massachusetts undergraduates who served as paid volunteers.

Apparatus. The apparatus was the same as that used in Experiment 1.

Procedure. The procedures and materials of Experiment 1 were replicated over five consecutive days. Each day subjects were required to produce two perfect serial recalls of the same categorized list before the 120 test trials were begun. Prior to the test trials in Session 5, all subjects had accurately recalled the serial list at least eight times and responded to 480 test trials, or 16 trials on each of the 30 different cue formats. By the last session all subjects had learned the list very well.

Results and Discussion

Intrusion errors and response times greater than 8 sec were not included in the data of Session 1 (7.50 and .28%) or 5 (1.39 and 0%). Intrusions in the last session reflected mainly the inappropriate naming of one of the cue items rather than an erroneous recall.

The data of principal importance are the correct response times to the different cue formats in Sessions 1 and 5. These results are shown in Figure 4 together with the data of Experiment 1 and List Condition C-RG3 of

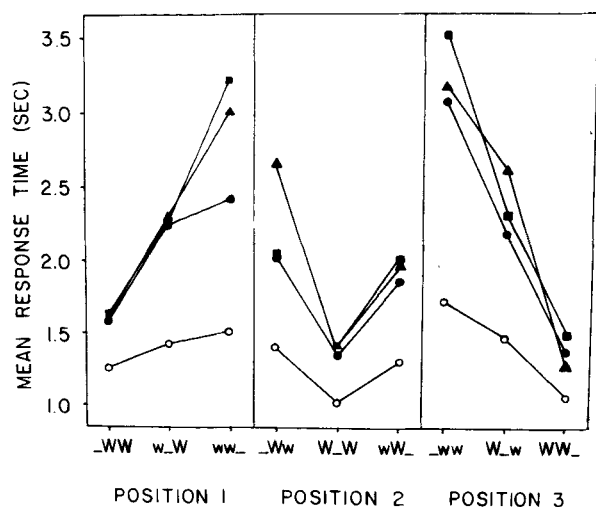


Figure 4. Mean response time as a function of cue format and word position within a semantic category for Session 1 (filled circles) and Session 5 (open circles) of Experiment 3. Also shown are the mean times for Experiment 1 (filled squares) and Experiment 2, Condition C-RG3 (filled triangles).

Experiment 2. Mean response times were obtained by averaging across items occupying the same position in a category over an entire list for each function. Most readily apparent in Figure 4 are the following observations: The cue format effects on response time in the first two experiments and in Session 1 of the present experiment are strikingly similar; response time to each cue format condition is reduced over five sessions; and cue format effects are still present after extended practice.

Separate analyses of variance were conducted on response times for first-, second-, and third-position words tested by all formats in Session 5. The analyses revealed reliable cue format effects for second- [$F(2,10) = 20.73$, $MSe = .05$] and third- [$F(2,10) = 10.96$, $MSe = .19$] position words only. The cue format effect on first-position words was significant in the first session but not in the last ($F < 1.0$). In postexperimental discussions some subjects reported that they were aware of their long response times to cues with two members of an adjacent category. They spontaneously attempted to reduce those times by recoding items across semantic boundaries in a manner reminiscent of Condition C-RG2 in Experiment 2.

These results suggest that incomplete learning of the order of categories in a serial list is not likely to account for the effects on response time. Cue format effects were still present for many of the list items after considerable practice. Such effects may reflect retrieval processes through an episodic list structure.

GENERAL DISCUSSION

This research has sought to determine if structural effects can be observed in long-term serial recall and to

specify how such effects might influence response time. Evidence for list structure was inferred from the large cue format effects on response time. Such effects were found to be consistent with the semantic relationships in categorized lists and the acquisition rehearsal pattern in unrelated lists. When rehearsal grouping and semantic relatedness were in conflict, the cue format effects conformed mainly to the rehearsal pattern. Extended practice did not eliminate these effects for many of the serial items. These results suggest that encoding patterns are used to structure a list and to provide a basis for retrieval; the best retrieval cues are those that are consistent with that structure.

One structure capable of representing a subject's serial knowledge is a hierarchy (see Crowder, 1976, pp. 434-442 for a review of Hebb's, Sheffield's, and Estes' use of hierarchical structures for serial order information). Individual items could be associated by coding elements representing the different word categories. The category elements would be associated in turn by a higher order element representing the entire serial list. Item and order information for the different hierarchical levels could be contained within each coding element (cf. Johnson, 1970). A list structured in this fashion would represent item information by vertical links between items and their category element and category elements and the higher order list element. Order information could be conveyed by position tags on the category elements and their items.

If retrieval is viewed as a hierarchical search, the cue format effects on response time can be interpreted as differences in search efficiency. Response time depends upon where a cue directs the search within the list structure. Consider the retrieval of FERN when probed by (TREE BUSH ----), (BUSH ---- PONY), or (---- PONY GOAT). A subject has two words, a blank, and their spatial arrangement to guide retrieval. In the first instance, either word and its distance from the blank can direct the search from its own representation to the probed item's representation in just two steps: The first step accesses a category element from a cue word (e.g., VEGETATION is accessed by TREE or BUSH) based on the lower level item-to-category element links; the second step unpacks the category element to reveal the item that is two positions after TREE or one position after BUSH. Since both cue words are equally efficient, a subject need process only one word and its position from the blank.

Differences in cue efficiency are observed when FERN is probed by (---- PONY GOAT). Processing either PONY or GOAT accesses the ANIMAL category element. Unpacking this element reveals that there is no item one position before PONY or two positions before GOAT. After two steps, the search must be redirected. The only direction the search can take is up. The coding element representing the entire list is accessed from the ANIMAL category element. In this case, order information alone now guides retrieval. The

list element is unpacked and the category element one position before ANIMAL is accessed. A total of six steps was needed to produce the correct response.

The retrieval of FERN by the cue (BUSH ----- PONY) can be considered as a combination of the preceding two instances. Retrieval based on BUSH will provide the correct response in just two steps, while a search based on PONY will take six steps. Since the cue provides no information on the relative importance of either word, it may be assumed that either is chosen with equal probability and a correct response is retrieved on the average in four steps. With other cues, such as (FERN PONY -----), the cue provides information to suggest that one word is more important than the other. In this case, the two words could be differentially weighted for the probability of directing search. Since complete efficiency may not be possible in a speeded task, the item closest to the blank would direct search with a probability of between .5 and 1. Depending upon how each item is weighted, the average number of steps to retrieve the correct response could be reduced from four to two. Differential weighting, however, has no effect on those cues which contain two words from either a different category (----- PONY GOAT) or the same category (TREE BUSH -----) as the probed item.

To examine the relationship between response time and the number of retrieval steps needed to produce a correct response, correlations were obtained between the mean response time for each of the 30 different cue format combinations and the hypothetical number of retrieval steps for each cue format. The 30 different formats consisted of four general types, each with its associated number of retrieval steps: (TREE BUSH -----), 2 steps; (FERN PONY -----), 3.33 steps; (BUSH ----- PONY), 4 steps; (----- PONY GOAT), 6 steps. The 3.33-steps example was obtained by making the word closest to the blank two times more likely to direct search than the item farther away. The correlations for the data from Experiments 1, 2 (C-RG3), and 3 (Sessions 1 and 5) were +.93, +.92, +.84, and +.71, respectively. Consistent with the assumption that the 3.33- and 4-step format types represent a mixture of the 2- and 6-step conditions are the findings that both the 3.33- and 4-step reaction time frequency distributions cross the fixed point of the 2- and 6-step distributions within a very small interval; the 3.33-step function, in fact, passes directly through the fixed point. These findings represent a necessary condition for the assumption that an intermediate distribution is a mixture of two extremes (Falmagne, 1968; Lupker & Theios, 1975). Bimodal distributions for the 3.33- and 4-step functions is not a necessary condition for mixture.

Collapsing the mean response times across cue formats with the same number of retrieval steps yielded an overall mean response time for the four main cuing conditions. These data, together with the best-fitting linear

equations, are shown in Figure 5. The results show response time to increase linearly over the number of retrieval steps in each instance. The fact that mean times for cue formats with 3.33 retrieval steps are less than those for 4 steps indicates that the data may not be described by a model which simply assumes that response time varies inversely with the number of cue words from the same category as the probed item. Interestingly, while no attempt was made to maximize data fitting, the 2:1 differential weighting for items in cue formats with 3.33 retrieval steps produced a very close fit in three of the four functions.

The results in Figure 5 indicate that the data for the present experiments may be well described by a linear equation which holds that response time is a function of cue encoding, search, and response emission. Cue encoding and response emission times can be inferred from the intercept of the function, while the search rate from the cue representation may be inferred from the slope. The reduced slope in the function for Session 5 of Experiment 3 might imply that search rate varies with practice. Alternatively, it may be that by Session 5 the categorized list was restructured through multiple coding. This additional coding could have interacted with cue formats to reduce times most for six-step retrievals and least for two-step retrievals.

Quantitative support for the retrieval model was obtained by comparing the best-fitting response time estimates with the observed data. The model provided reasonable fits in each instance. A mean absolute differ-

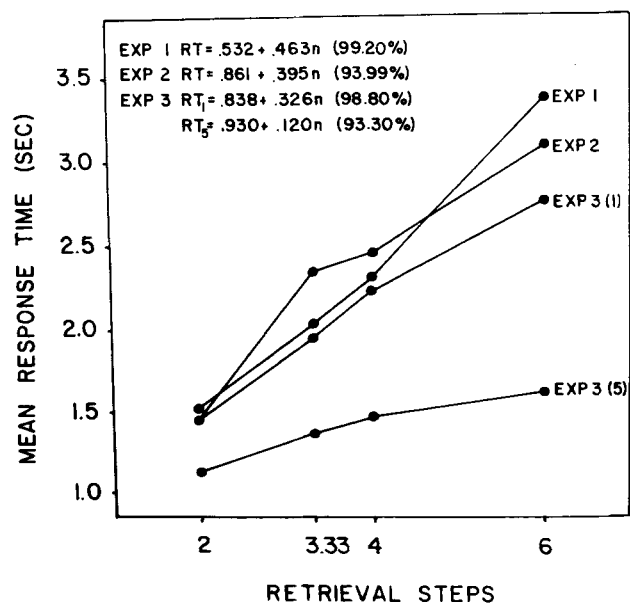


Figure 5. Mean response time as a function of the number of retrieval steps necessary to produce a correct response for Experiments 1-3. Best-fitting linear equations were obtained by the method of least squares for each condition and are shown with the percent of variance accounted for by linear regression in parentheses.

ence between theoretically derived and observed times was obtained for each experiment by averaging over the nine data pairs. The mean absolute difference was .09 sec for Experiment 1, .19 sec for Experiment 2 (C-RG3), .14 sec for Experiment 3, Session 1, and .08 sec for Session 5. Figure 1 shows the close correspondence between the derived and observed data in Experiment 1.

In summary, the present experiments suggest that retrieval processes are based on a linear search of a hierarchical structure containing item and order information. Response time varies with cue formats because cues differ in the degree to which they permit access to the correct category in the list structure. The structure, which is acquired at the time of learning, determines cue efficiency and, hence, the subsequent effects upon response time. These data are not inconsistent with expectations from multilevel associative models of serial order. Further research will have to determine if this paradigm can be applied to specific models such as Estes' perturbation model (Estes, 1972) for short-term recall. One possibility is that the present task involves long-term and short-term components: Higher level category elements may be retrieved from long-term memory and unpacked in short-term memory. Variations in list structure and probe techniques may provide additional information.

REFERENCES

- BATTIG, W. F., & MONTAGUE, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph Supplement*, 1969, **80**(No. 3, Part 2).
- BOWER, G. H., & WINZENZ, D. Group structure, coding, and memory for digit series. *Journal of Experimental Psychology Monograph Supplement*, 1969, **80**(No. 2, Part 2).
- CROWDER, R. G. *Principles of learning and memory*. Hillsdale, N.J.: Erlbaum, 1976.
- ESTES, W. K. An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. New York: Wiley, 1972.
- FALMAGNE, J. C. Note on a simple property of binary mixtures. *British Journal of Statistical and Mathematical Psychology*, 1968, **21**, 131-132.
- JOHNSON, N. F. The role of chunking and organization in the process of recall. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 4). New York: Academic Press, 1970.
- LASHLEY, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley, 1951.
- LUPKER, S. J., & THEIOS, J. Tests of two classes of models for choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **1**, 137-146.
- MARTIN, E. Serial learning: A multilevel access analysis. *Memory & Cognition*, 1974, **2**, 322-328.
- MCLEAN, R. S., & GREGG, L. W. Effects of induced chunking on temporal aspects of serial recitation. *Journal of Experimental Psychology*, 1967, **74**, 455-459.
- MURDOCK, B. B., JR. *Human memory: Theory and data*. Hillsdale, N.J.: Erlbaum, 1974.
- RESTLE, F., & BROWN, E. R. Serial pattern learning. *Journal of Experimental Psychology*, 1970, **83**, 120-125.
- SEAMON, J. G. Serial position effects in probe recall: Effect of rehearsal on reaction time. *Journal of Experimental Psychology*, 1972, **96**, 460-462.
- SEAMON, J. G. Retrieval processes for organized long-term storage. *Journal of Experimental Psychology*, 1973, **97**, 170-176.
- TULVING, E., & THOMSON, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, **80**, 352-373.
- WILKES, A. L., & KENNEDY, R. A. The relative accessibility of list items within different pause-defined groups. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 197-201.

(Received for publication February 15, 1977;
accepted September 6, 1977.)