DOCUMENT RESUME

ED 082 169 CS 000 762

AUTHOR	Underwood, Benton J.; And Others
TITLE	The Locus of the Retention Differences Associated with Degree of Hierarchical Conceptual Structure.
INSTITUTION	Northwestern Univ., Evanston, Ill.
SPONS AGENCY	Office of Naval Research, Washington, D.C. Personnel and Training Research Programs Office.
PUB DATE	Sep 73
NOTE	44p.; See related documents CS 000 761 and CS 000 763
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Association (Psychological); Cognitive Processes; College Students; *Concept Formation; Educational Research; Expectation; *Paired Associate Learning; Reading; *Recall (Psychological); *Retention; *Retention Studies; Structural Analysis

ABSTRACT

Constant-order paired-associate lists were used in which the numbers 1-24 were stimulus terms, and 24 nouns were response terms. The order of the nouns was varied across five lists to produce a different number of the hierarchical conceptual levels in the lists. There were two degrees of original learning and three types of retention tests after 24 hours. The study-test method was used. Learning rate was related directly to the degree of conceptual structure, but retention was uninfluenced by structure. A further experiment showed that the direct relation between recall and structure found in an earlier study is to be attributed to the anticipation method in which information at recall is an amount that is directly related to the conceptual structure. (Author)

ED 082169

THE LOCUS OF THE RETENTION DIFFERENCES ASSOCIATED WITH DEGREE OF HIERARCHICAL CONCEPTUAL STRUCTURE

Benton J. Underwood, John J. Shaughnessy and Joel Zimmerman

Northwestern University

September 1973

Project NR 154-321
Sponsored by
Personnel & Training Research Programs
Psychological Sciences Division
Office of Naval Research
Arlington, Virginia
Contract No. N00014-67-A-0356-0010

CS 000 762

THE LOCUS OF THE RETENTION DIFFERENCES ASSOCIATED WITH DEGREE OF HIERARCHICAL CONCEPTUAL STRUCTURE

Benton J. Underwood, John J. Shaughnessy and Joel Zimmerman

Abstract

Constant-order paired-associate lists were used in which the numbers 1-24 were stimulus terms, and 24 nouns were response terms. The order of the nouns was varied across five lists to produce a different number of hierarchical conceptual levels in the lists. There were two degrees of original learning and three types of retention tests after 24 hr. The study-test method was used. Learning rate was related directly to the degree of conceptual structure, but retention was uninfluenced by structure. A further experiment showed that the direct relation between recall and structure found in an earlier study is to be attributed to the anticipation method in which information at recall is in an amount that is directly related to the conceptual structure.

THE LOCUS OF THE RETENTION DIFFERENCES ASSOCIATED WITH DEGREE OF HIERARCHICAL CONCEPTUAL STRUCTURE[1]

Benton J. Underwood,[2] John J. Shaughnessy and Joel Zimmerman

Northwestern University

In a previous study (Underwood & Zimmerman, 1973), 16 words were ordered serially so that a three-level conceptual hierarchical structure resulted. If S, in learning the list, followed the rules implied by this structure, placement of each word in its appropriate position within the list was possible. Other lists were constructed from the same words in a way such as to violate the appropriateness of successive conceptual levels. The purpose of this previous study was to determine the role of conceptual structure on the learning and retention of the lists. Two findings emerged. First, learning rate was related directly to degree of conceptual structure up to a point, and second, recall after 24 hr. was related directly to degree of structure. This latter finding conformed in general to the notion that associations learned in the laboratory which are compatible with already established associations will show less rapid forgetting than will be the case for associations which are in conflict with established habits. In the previous study, this latter case was represented at the extreme by the 16 words presented in random order. It was presumed that for such a list the long-established conceptual habits would interfere with the appropriate ordering of the words at recall. In fact, however, the overt-error data gave no evidence that the poorer recall of the random list than of the structured list was due to such interference. Thus, the reasons for the differences in recall remained obscure.

The intent of the present study was to identify more precisely the characteristics of the memory for structured and unstructured lists. By so doing it was believed that the characteristic or characteristics responsible for the differences in retention might be isolated. Two different levels of learning were used as a means of varying the degree to which the conceptual structure had become a part of the memory for the lists. Since the utilization of a conceptual structure to mediate item placement at the time of retention may be critically time dependent, three different types of retention tests were used, namely, paced and unpaced recall, and unpaced associative matching.

Method

Lists. Five different lists were constructed, all from the following 24 words: robin, owl, bobolink, trout, guppy, sturgeon, apple, lemon, fig, rose, lilac, marigold, beer, rum, sherry, milk, soda, cocoa, diamond, opal, sapphire, iron, brass, tungsten. These 24 words, presented as in the order listed above, formed the most highly structured list, to be called List 5. It will be noted that at the lowest conceptual level in List 5 there are three instances of each of eight concepts, birds, fish, fruit, flowers, alcoholic beverages, nonalcoholic beverages, precious stones, and metals. At the next conceptual level there are four concepts, animals, plants, beverages, and minerals. At the third level, living and nonliving things divide the list in half.

In List 5 the three instances within each concept were ordered serially (as above) such that the first instances had high frequency in the Battig-Montague (1969) norms, the second medium frequency, and the third, low frequency. Given this ordering the S could, in a manner of speaking, run off the three instances within a concept according to a frequency rule. In the previous study (Underwood & Zimmerman, 1973), this variable, with two instances of each concept, did not influence learning. However, there were reasons to believe that the earlier lists were not entirely satisfactory for a test of the frequency rule and so this variable was included again in the present experiment.

The nature of the other four lists may now be described. For List 4 the conceptual structure remained the same as in List 5, but the order of the three instances within each of the eight concepts was randomized to neutralize any influence of a word-frequency rule. In List 3 the order of the six animal names was arranged so that the concepts birds and fishes were not appropriate for three successive items, but the concept animal was appropriate for a block of six words (robin, trout, sturgeon, owl, guppy, bobolink). The same was true for plants, beverages, and minerals. Therefore, the four intact concepts in List 3 could mediate placement only within a block of six positions. For List 2 only the living-nonliving distinction was maintained so that the implementation of this distinction by the S would restrict placement to halves of the list. Finally, in List 1, the ordering was random so that no conceptual mediation of placement of groups of words was possible.

The lists were presented as constant-order paired-associate lists with the numbers 1-24 in order as stimuli and with the words as response terms. In the previous experiment, this procedure gave results which were essentially equivalent to those found when the words were presented as a true serial list.

Conditions. One variable, of course, was defined by the five lists as explained above. A second variable was the degree of learning prior to the 24-hr. retention interval. Half of the Ss learned to a criterion of 12 correct responses on a single trial, half to a criterion of 20 correct responses on a single trial. A third variable was the nature of the retention test. Half of the Ss were given a paced-recall test, the rate being the same as used during learning. For these Ss, recall was followed by relearning to one perfect recitation but with a minimum of three relearning trials after the recall trial. Half of the Ss were given unpaced retention tests consisting of two steps. Initially, a sheet was provided the S on which the numbers 1 through 24 were listed with a blank after each. The S was given unlimited time to write down all of the appropriate response terms he could, guessing being encouraged. Following this step, a list of the 24 response words was provided the S on a second sheet and he was asked to match each word with a number, using each only once. He was required to fill each blank with a word even if it involved guessing.

The five lists, two degrees of learning, and two types of retention tests resulted in 20 different conditions. Four further conditions were added to provide controls for possible differential performance on the post-criterial trials of learning. The study-test method was used during learning, and although the use of this method is normally expected to minimize differences on post-criterial trials as a function of rate of learning, it was believed necessary to provide some minimal information about the matter. Consequently, four groups were given immediate retention tests after achieving the criterion of 20 correct responses on a single trial. Two of these groups learned List 1, and two learned List 5. For each list, one group was given paced recall and relearning, one unpaced recall, followed by matching.

There were 24 groups of college-student Ss, one corresponding to each of the 24 unique conditions, with 18 Ss in each group. Assignment to a particular group was made from a schedule containing 18 blocks of conditions, with each condition occurring once within each block. A different random order was used for the conditions within each block.

Procedure. All lists were presented for alternate study and test trials at a 1.5-sec. rate for both. When the appropriate criterion was achieved, the S either was dismissed from the laboratory (to return 24 hr. later) or was given an immediate retention test. On the immediate paced tests the experimenter stopped the memory drum, told the S that he would now have another test trial, with further study and test trials to follow, and that he should try to get as many correct as possible. For the unpaced tests (both immediate and delayed) the recall sheet was given the S and he was asked to follow the printed instructions on the sheet as the experimenter read these instructions aloud. He was given a second sheet (described earlier) for the matching test. All Ss having 24-hr tests were reminded that they had learned a list of words the previous day and that a test of their memories for this list was the purpose of the present session.

Finally, all Ss having the 24-hr retention tests were given an open-ended questionnaire concerning rehearsal activities they might have engaged in over the 24 hr. They were asked to describe any experiences they had with the words during the 24-hr period, whether they had rehearsed or thought about the words, and so on. Honesty and accuracy were emphasized since (the *S* was told) such replies were valuable in helping to understand the nature of memory.

Results

Original learning. For an examination of learning as a function of list structure, the *Ss* who would subsequently receive different retention tests were combined to provide 36 *Ss* having had each of the five lists at each criterion of learning. The data are plotted in Fig. 1 in terms of the mean number of trials required to reach the two criteria. For convenience, the five lists are equally spaced along the baseline, indicating increasing structure from List 1 to List 5. It can be seen that as list structure increased, trials to learn decreased. This is true for both criteria, and summed across the two, the effect is significant statistically, F (4,350)=8.08, $p < .01$. Although the influence of list structure appeared to be somewhat greater for the higher criterion of learning than for the lower criterion, the interaction was not reliable, F (4,350)=2.14, $p > .05$. In the previous study (Underwood & Zimmerman, 1973), learning rate increased from List 1 through List 3 with no further increase for Lists 4 and 5. It is not known whether this difference is due to methods of learning (anticipation versus study-test) or to list differences. There was some evidence in the earlier study that some of the instances defining the lowest-level concepts were not always understood by the *S*.

It will be remembered that List 5 differed from List 4 only in terms of the ordering of the three words within each of the eight concepts. For List 5 the words were ordered from high to low in terms of the frequency with which the instances were produced to the concept name in the Battig and Montague (1969) normative study. In List 4, the three words were randomized. As seen in Fig. 1, List 5 was learned more rapidly than was List 4, and this was true for both criteria. An analysis of variance with Lists 4 and 5 as one variable, and the two criteria as the other, showed the difference between the lists to be reliable, F (1,140)=5.45, $p<.05$.

The effect of word frequency for all lists was also examined. Three scores were obtained for each S, these representing the number of times the eight high-frequency instances were given correctly, the number of times the eight medium-frequency instances were given correctly, and the number of times the eight low-frequency instances were given correctly. Across all lists combined for the lower criterion of learning the three means were 13.99, 11.02, and 9.84, for the high-, medium-, and low-frequency instances, respectively (F=87.59). These differences essentially disappeared at the higher criterion of learning for Lists 4 and 5, but were still quite evident for the other three lists. It appears, therefore, that one of the characteristics of initial learning of these lists by the study-test method was that either because of ease of learning of words with high frequencies, or because of priority effects, or both, the S acquired the high-frequency instances initially regardless of the conceptual structure of the list.

Still another finding could be interpreted as a word-frequency effect. Some observations of the experimenters suggested that the learning of the concepts differed for men and women. The learning of Lists 4 and 5 were evaluated to see if these observations had validity. The 144 Ss (both criteria of learning) consisted of 77 men and 67 women. The number of correct responses within each of the eight concepts was determined for each group. These are plotted in Fig. 2 for two reasons. First, they show the nature of the serial-position curve which obtains even for the highly structured lists. Second, they indicate differential learning of certain concepts by men and women. Overall, the learning of the two groups did not differ ($F < 1$), but the interaction between concepts and sex was reliable, $F(7,994)=3.13$, $p < .01$. Three concepts produced substantial differences in learning between men and women, namely, flowers, alcoholic beverages, and precious stones. It may be that cultural experiences have resulted in women having a greater familiarity for words representing flowers and precious stones than is true for men, and that the opposite is true for alcoholic beverages. In any event, these results are similar to those reported by Bousfield and Cohen (1956) when they used free-recall learning of groups of conceptually related words.

Figure 2 suggests that the words which represented the first occurrence of a nonliving concept (alcoholic beverages) caused a distinct break in the position-performance curve. However, this is primarily a function of the particular words involved. For Lists 2 and 3, where the living-nonliving break also came between positions 12 and 13, there was no evidence of discontinuity between the two positions.

Overt errors. The first overt error analysis to be presented dealt only with the number of errors across lists. The 360 Ss from which the acquisition data were derived (as plotted in Fig. 1) were used. The error measure was the proportion of times an error was produced per opportunity (number of overt errors/number of omissions plus overt errors). For the 10 groups this ratio varied between .10 and .19. The values were greater statistically for the high-criterion groups than for the low-criterion groups, $F(1,350)=8.33$, $p<.01$, the means being .16 and .12. The errors increased with list structure (.11, .11, .15, .15, .16), $F(4,350)=4.08$, $p<.01$. The F for the interaction was less than one. Undoubtedly, some correct responses resulted from guesses within concepts, particularly for the more structured lists, but in view of the relatively low proportion of overt errors to cases of not responding, it does not seem likely that mere guessing is heavily involved in the correct-response data.

For Lists 4 and 5, overt errors can be identified as having one of four levels of appropriateness in terms of the conceptual structure for these lists. These will be described and illustrated. An error may occur within the appropriate block of three instances representing the same concept. Saying "guppy" to one of the two stimuli to which it was an incorrect response in the block of three fishes was such an error, and will be called here an error at Level 1. Six words were involved in the animal block; if "guppy" was given to one of the three stimulus terms appropriate for birds, it was defined as a Level-2 error. Twelve words were included in the living block; if "guppy" was given to any of the six stimuli paired with plants, it was called a Level-3 error. Finally, if "guppy" was given as a response to any of the 12 stimuli paired with nonliving objects, it was called a Level-4 error. All errors were classified into one of these mutually exclusive categories for Lists 4 and 5. Conceptual level and position within Lists 4 and 5 are perfectly confounded. Position is also tied to stimulus number. Therefore, any analysis of the appropriateness of overt errors to the conceptual structure must consider the number of such errors which were produced by position per se. To handle this problem, the overt errors made in learning all five lists were categorized in exactly the same way as was done for Lists 4 and 5. For example, if, for List 1 (no structure), a misplaced response occurred among the first three pairs, it was classified as a Level-1 error, just as was done for List 5. Or, if an error made to the stimulus-term 9 was an appropriate response for the stimulus-term 3, it was classed as a Level-3 error. To summarize: for all lists the errors were classified as falling within one of eight blocks of three positions (Level 1), one of four blocks of six positions (Level 2), one of two blocks of 12 positions (Level 3), and, all others (Level 4).

The errors made by each S were allocated to the appropriate level and the percentage of errors in each level was determined. The means of these values were used to construct Fig. 3. The sums of the percentages for a given list do not always equal 100 since a few Ss produced no overt errors, and the means shown in Fig. 3 were always based on N=18. It should be noted that list structure is along the abscissa, with the error levels (1, 2, 3, 4) as the plotting parameter. The left panel represents the data for the Ss who learned to a criterion of 12 correct, the right panel for those who learned to a criterion of 20 correct. Several features may be noted. For Lists 4 and 5, the greatest percentage of errors by far was that for errors falling into Level 1. Over half the errors made by the Ss learning these lists represented errors that were appropriate to the narrowest conceptual categories. The Level-1 values for List 1, values which should represent in relatively pure form the effects of position as such, are at 25% for the lower criterion, 40% for the higher criterion. In List 3 the narrowest conceptual category consisted of six words. If conceptual structure were used by the S in the placement of items in this list, he could reduce his possibilities to one of six positions. As may be seen in Fig. 3, the maximum number of Level-2 errors occurred in learning List 3, these errors reflecting placement within a block of six words. For List 1, the random list, the maximum frequency of errors for the lower criterion was in Level 4, which means giving a response in one half of the list which actually belonged in the other half. There is only slight evidence that the living-nonliving distinction influenced the errors, this being shown in the fact that fewer errors at Level 4 were made by the Ss learning List 2 than by those learning List 1. As would be expected, the major difference in the errors for the two criteria of learning was that there was an increase in Level-1 errors from the lower to the higher criterion, with the increase being the greatest for the lists with low structure. In summary, Fig. 1 showed that learning was related directly to the number of appropriate conceptual levels involved in the list; the error data of Fig. 3 indicate that these conceptual levels aided learning because they limited the number of possible numerical stimuli or positions for which a particular word was appropriate.

There is the possibility that with zero conceptual structure (List 1), the presence of conceptually related words actually interfered with learning. If such interference was present, it was not manifest in the error data. The errors made in learning List 1 were divided into two categories, those which were given to a stimulus for which another instance of the same concept was appropriate, and those which were not appropriate in the above sense. The concepts used were the eight with the three instances each. To illustrate: saying "guppy" to the stimulus appropriate for <u>trout</u> would be viewed as evidence for interference; giving "guppy" to the stimulus for <u>milk</u>, would not. Errors which would constitute evidence for interference constituted 7.6% of the total errors. Chance responding would be expected to yield 8.7% of such errors. This indicates that interference resulting from the conceptual relationships among the words was minimal in the unstructured list.

<u>Retention</u>. The retention data will be presented first for the <u>Ss</u> who learned to a criterion of 12 correct on a single trial. The results for all three retention tests are shown in Fig. 4. To replicate the

earlier study (Underwood & Zimmerman, 1973), performance should have increased as list structure increased. It is obvious from Fig. 4 that list structure had little effect on retention. The means for the paced 24-hr. recall show little variance ($F < 1$). Even for the unpaced 24-hr. test, where List 5 seems high relative to the other four lists, the F is only 1.37. The fact that matching exceeded unpaced recall indicates that the Ss were unable to recall some response words but could pair them appropriately when given the matching test. No immediate test was given to Ss learning to the lower criterion so it is difficult to estimate the amount of forgetting which took place over the 24 hr. However, the number correct on the last trial of original learning does give some basis for a rough estimate. An immediate test would probably have shown some loss due to the fall normally found after reaching a criterion. As will be shown later, it would not be unreasonable to expect this loss to be at least one to two items. If so, the amount of forgetting was between 30% and 40% for the five lists. The critical fact, however, is that list structure is unrelated to the amount lost over 24 hr.

Turning next to the retention scores for those Ss who learned to a criterion of 20 correct responses, the conclusions are much the same as for the lower criterion. The essential data are shown in Fig. 5. For these comparisons the immediate retention tests are available for Lists 1 and 5, but it may be useful to note here that the mean number correct on the last test trial of original learning varied between 21 or 22 correct responses. Although there seems to be more variance among the means for paced 24-hr recall than was true at the lower criterion, the statistical conclusion is the same, the F for the paced recall test being only 1.74. Even if Lists 1 and 5 are tested for paced recall along with the immediate paced recall for those lists, the interaction falls short of acceptable levels of significance, F (1, 68) = 3.34, p .05. Using the immediate paced recall as a base, forgetting under paced recall is estimated at 25% across the lists as a whole. It is obvious also that forgetting occurred over 24 hr. for those Ss given the unpaced test.

The number of misplaced responses at recall was examined for all paced-recall groups. These numbers did not differ as a function of the criterion of learning, but for both levels of learning the number of misplaced responses at recall increased as list structure increased, F (4, 170) = 5.81, p .01. This same relationship was found during learning and so it not a phenomenon peculiar to recall. For unpaced recall, however, list structure did not influence the number of incorrectly-paired responses given.

As noted earlier, during learning of the constant-order paired-associate lists, very clear serial-position effects were present. These position curves were in evidence for all of the types of retention tests. Furthermore, there was a strong relationship between the number of times an item was given correctly in original learning and the number of times given correctly at recall, which is to indicate that the position effects remained essentially constant from learning to recall. For example, for the lower criterion, the correlations between number of times correct in original learning and number of times correct at recall varied between .77 and .91 for the five lists. Degree of original learning was obviously a powerful determiner of the items which were recalled.

Relearning. Differences in relearning as a function of list structure appeared on the first test trial after recall, and in terms of trials to reach a criterion of one perfect recitation, were reliable, F (4,170)=4.14, $p < .01$, with, of course, rate of relearning being directly related to structure. Thus, list structure, which did not influence recall, quickly reinstated during relearning the influence it had had during original learning.

Rehearsal and recall. The Ss were given an open-ended questionnaire concerning their rehearsal activities over the 24 hr. The replies to these questionnaires were rated on a 9-point scale for the amount of rehearsal implied in the protocols. These ratings were carried out independently by three different people. Interrater reliabilities were determined for 18 subblocks of 20 Ss each. Of the 54 correlations possible, 11 were between .71 - .80, 29 between .81 - .90, and 14 were .91 or greater.

There were 20 groups of Ss of 18 each having 24-hr. recall. All correlations between rated rehearsal activity and retention were positive, varying between .14 and .81. There was no relationship between rated rehearsal activity and list structure, and the magnitude of the correlations between rehearsal and retention did not vary systematically as a function of list structure.

The positive relationship between reported rehearsal and retention allows several alternative interpretations. For example, rehearsal may have increased retention, or, as another interpretation, Ss with good retention may have rehearsed. The concern of the present study was whether or not rehearsal differed as a function of list structure. Since it didn't, it seems unlikely that the failure to find an influence of list structure on retention could be due to differential rehearsal. Also, if the present evidence on rehearsal can be generalized to the earlier study (Underwood & Zimmerman, 1973), it seems unlikely that the positive relationship between list structure and retention as reported for that study was due to differential rehearsal.

Preliminary Discussion

In the previous study, paced recall and list structure were directly related and the forgetting over 24 hr. for the list with the highest structure was only 5%. In the present study, list structure was unrelated to either paced or unpaced recall, and the forgetting for the highest criterion of learning used was estimated at 25%. The discrepancy between these two studies is the topic of this preliminary discussion.

The contradiction in the retention results for the two studies pointed immediately to a method difference as the likely source for the contradiction. In the previous study the anticipation method was used. During the recall trial, therefore, the S would be informed of the particular concept whose instances were appropriate at the moment for the structured lists. If, for example, he remembered there were six animal names in succession, the appearance of the first instance would inform him that the next several positions also contained animal names. With the study-test method, the S would not be given this information. If he did not give the first instance of a concept at the appropriate point (to the correct stimulus), his following responses would be incorrect unless he remembered a particular response term associated with a stimulus term within the series of concept instances. There were, in fact, some cases in which S did give three to six correct response words in the correct order but which were scored as wrong because the initial response in the series was not paired with the appropriate stimulus. This did not occur frequently (and the data show that scoring these as correct did not change the basic conclusion), but the uncertainty felt by the S may have prevented him from responding overtly. List structure clearly influenced learning and relearning, but some portion of the learning which allowed the S to align response terms and stimulus terms correctly during learning must have been forgotten over the 24 hr. It is possible that the numerical stimulus terms were not always used as the effective stimulus terms and that the lists were treated more as serial lists than paired-associate lists, although what this means theoretically is not known. It appears that if S had learned, for example (for Lists 4 and 5), that number 1 was paired with the first bird instance, number 4 with the first of three fish instances, and so on, that this part of the memory was lost over 24 hr. Otherwise, there is no reason why providing this information at recall (as is done in part under the anticipation method) should result in better performance.

It seemed necessary to make a test of the notion that the discrepancy in the recall results for the two experiments was due to a difference in methods. In the auxiliary experiment, to be reported now, Lists 1 and 5 were presented for anticipation learning, with recall taken after 24 hr. The expectation was that recall of these two lists would differ, with List 5 giving better recall than List 1.

Auxiliary Experiment

Method. Lists 1 and 5 were presented at a 1.5:1.5-sec. rate for anticipation learning. The criterion of original learning was 12 correct responses on a single trial. Paced recall occurred after 24 hr. with relearning carried to one perfect trial, but a minimum of three trials beyond the recall trial. Each list was learned by a separate group of 18 Ss assigned to one of the two lists by a block-randomized schedule.

Results. In presenting the results, comparisons will be made with the two groups of Ss from the major experiment having the study-test method and paced recall, and who had learned Lists 1 and 5 to a criterion of 12 correct responses before the retention interval. There were 18 Ss in each of these two groups. Since the auxiliary experiment was conducted after the major experiment, it is not known if the groups (study-test vs. anticipation) represent the same or equivalent populations. Therefore, differences in the levels of performance will not be readily interpretable, although interactions between the methods and lists should be meaningful.

The mean numbers of trials to learn to a criterion of 12 correct, and the mean numbers of trials to relearn to one perfect trial, are shown for the four groups in Fig. 6. For original learning, list structure appears to have a greater influence for anticipation learning than it does for the study-test method. List structure is significant (F=25.33), as is the interaction, $F(1,68)=4.19$, $p<.05$. The interaction between list structure and method does not occur during relearning, indicating that the methods influence on learning is confined to the early stages.

It will be remembered that for the study-test method the frequency of concept instances was directly related to learning. This was quite evident at the lower criterion of learning. For List 5, under the anticipation method of the auxiliary experiment, the reverse was found. More specifically, two findings held across all eight concepts of three instances each. First, the initial word of the three (the high-frequency instance) was never given correctly more times than the second instance (the medium-frequency instance). Second, the third word in each of the eight concept triads (the low-frequency instance) was always given correctly more times than the first instance of the succeeding concept. For List 1 under anticipation learning, the effect of word frequency was in evidence just as was true for all of the lists learned by the study-test method. The above facts would indicate that item learning under the two methods would be more reliable for List 1 than for List 5. The product-moment correlations for item learning were .83 for List 1, and .58 for List 5. The positive relationship for List 5 reflects the commonality in learning by the two methods produced by serial position of the words, differences which were apparent under both methods.

Mean overt errors per opportunity were greater in learning under the anticipation method (.22) than under the study-test method (.13), and the difference was reliable, $F(1,68)=9.24$, $p<.01$. This may represent a greater tendency to guess under the anticipation method than under the study-test method. However, of the overt errors made under the two methods, the percentage of these errors within the appropriate concept position (Level-1 errors as described earlier) was about the same for List 5, being 53% for anticipation and 54% for the study-test method.

The mean numbers of correct responses at recall are shown in Fig. 7, along with the mean number correct on the last learning trial. Under the anticipation method, recall was directly related to list structure, $F(1, 34)=4.97$, $p<.05$. The interaction between lists and methods for recall was also reliable, $F(1, 68)=4.44$, $p<.05$. Recall for List 5 under the anticipation method was higher than performance on the last learning trial, 24 hr. earlier. To some extent, these comparisons are all in error, a matter which needs discussion.

Consider first the recall of List 1 under the two methods. A conclusion from an inspection of Fig. 7 might be that recall is superior following anticipation learning to that following learning by the study-test method. However, two factors must be considered. First, the criterion fall which may occur under the study-test method, and second, the learning which occurred on the last trial under the anticipation method. This last issue can be handled directly. A multiple-entry probability analysis (Underwood, 1964) was carried out to project performance to the hypothetical next trial under the anticipation method. The mean expected value was 14.89. Since recall was 11.67 items, forgetting over 24 hr. was 3.22 items, or 22%. To incorporate the criterion fall into the calculations for List 1 for the study-test method requires a rough estimation of values. In the main experiment, groups were given a recall test immediately after reaching a criterion of 20 correct responses. On this immediate test under paced recall, performance fell from a mean correct of 22.00 correct on the last test trial to 18.50 correct on an immediately following test trial. This represents a loss of 3.50 items which is referred to here as the criterion fall. None of the groups had an immediate test following learning to a criterion of 12 correct responses. Assume, however, that the group having the study-test method and learning to a criterion of 12 correct responses actually forgot 22%, the same as forgotten by the group having the anticipation method. This would require a criterion fall of 2.75 items. In light of the criterion fall of 3.50 items shown by the group learning to 20 correct responses, a fall of 2.75 items for those learning to 12 correct would not seem to be seriously in error. Given this assumption, the conclusion is that for List 1, the amount of forgetting shown under the two methods is roughly the same.

Turning next to List 5, it must first be noted that it was not possible to project "next trial" performance for the Ss learning under the anticipation method because several of them learned in one or two trials. As an estimate, however, the percentage projection obtained for List 1 was used. This gives a mean expected next-trial value of 17.03 items. Since recall was 15.69 items on the average, the loss over 24 hr. was about 7%. Deriving an expected loss for List 5 following study-test learning (using the value of 2.75 for the criterion fall) gives a loss of 3.75 items from 12.92, or 29% forgetting. These values must be considered approximations, but they lead to the conclusion that the loss was about the same under the two methods for the unstructured lists, but that the anticipation method led to better recall than did the study-test method for the highly structured list. Both the estimate of loss for List 1 for the anticipation method (22%) and the loss for List 5 under the same method (7%) correspond to the findings of the earlier published experiment using this method (Underwood & Zimmerman, 1973).

The sources of the differences in the retention for List 5 following the learning by the two methods needs more detailed examination. The data which seem to aid in reaching decisions about the sources of differences are shown in Table 1. The data sheets were examined for the last trial of original learning and for the recall trial. A tabulation was made of the number of correct responses which resulted from producing all three instances of a concept, the number which resulted from producing only two instances of a concept correctly, and the number which resulted from giving only one correct response from among the three possible for each concept. When two correct responses were given, they were broken down by position within the three possible positions (1 & 2, 1 & 3, 2 & 3). When a single correct response was given, the number falling in each of the three positions was noted. All of this information is given in Table 1 for the last original learning trial, and for the recall trial, under both methods. For example, on the last learning trial under the anticipation method, there were 39 cases in which all three instances were given correctly, resulting in 117 correct responses, this latter value being shown in Table 1. For this same condition there were 13 cases in which a single correct response was given for a concept and which consisted of the second of the three instances in the series of three.

Several facts are to be noted in Table 1. First, under the anticipation method there was an actual *increase* over 24 hours in the number of cases in which all three instances were correctly given (117 to 162). In view of the learning which may have taken place on the last anticipation trial of learning, this increase must be viewed cautiously. Nevertheless, it is in marked contrast to the results found with the study-test method, where there was a loss of 108 (210-102) complete triads. The second fact to note is the difference between the two methods in both learning and recall when less than three instances were given correctly. Under the anticipation method, the most probable positions for correct responses when two were given are positions 2 and 3, and position 3 when a single correct response was produced. Comparable relationships do not exist for the study-test method, where there was more or less constancy among the positions. The third fact is that if the number of cases in which the first instance (by position) of a concept was given correctly was calculated (regardless of the outcome for the two following instances) for recall, the sums were found to be 198 cases for anticipation, and 146 for study-test. This seems to indicate that, (1) associative learning between the first instance of the concept and its position or stimulus number was greater or better during anticipation learning than during study-test learning, or (2), showing the S the response terms during recall allowed him to "deduce" the subsequent occurrence of a first instance of a different concept. Both factors may be involved, although the latter seems more reasonable. Finally, the data in Table 1 are clear in demonstrating that failure of S to recall the first instance of a concept provides very little penalty under the anticipation procedure, since he can proceed to give other instances of the concept for the two following positions.

The evidence seems to point to the fact that learning under the study-test method is relatively fragile or weak with regard to positions or stimulus numbers which mark the first instance of a new concept. The same is probably true for the anticipation learning. However, on the recall trial the S having learned by the study-test method receives no feedback information and he has few means to apply corrective procedures based on any knowledge he had about the conceptual structure. In effect, he was reduced to responding on an item by item basis, much as is the case for an unstructured list. On the other hand, the S having anticipation learning could apply corrective procedures based on his knowledge about the conceptual structure of the list.

General Discussion

The higher the conceptual organization of the lists the more rapid was the learning. It is presumed that this relationship was produced because the greater the number of the conceptual levels the more precise was the placement of an item if the rules indicated by the conceptual relationships were followed. Nevertheless, this should not be taken to mean that the hierarchical structure in its totality entered into the learning of the most completely structured lists. The S could learn the sequence of eight concepts of three instances each without reference to the higher-order concepts present in the list (animals, plants, beverages, minerals). So also, List 3 could be learned by reference only to the four successive concepts of six instances each and without the living-nonliving distinction per se entering into the learning. The error data were clear in showing that at least one level of conceptual responding was involved in the learning of Lists 3, 4, and 5, but these error data do not speak to the question of whether two or more conceptual levels were involved in learning Lists 4 and 5. To determine if two or more levels were involved would require the use of list in which the blocks of three instances were ordered randomly with respect to the more inclusive concepts. A list of this type was not included in the present study.

In the previous study (Underwood & Zimmerman, 1973), learning was facilitated up through the structure corresponding to List 3 in the present study. No further enhancement occurred for Lists 4 and 5. This is believed due to the fact that the present lists were made up of more obvious concepts and more obvious concept instances than were the previous lists. Unlike the present findings, frequency of the concept instance with a concept did not influence learning in the previous experiment. In view of the findings of the auxiliary experiment, it seems now that this contradiction is another by-product of the differences in the learning by the two methods.

The major purpose of the present study was to attempt to identify the component(s) in memory which had led to the increase in recall over 24 hr. as list structure increased. In an inelegant way, this study was quite successful in achieving its purpose. The use of the study-test method of learning was followed by a complete lack of any differences in the recall tests, whether paced or unpaced. Only relearning, reflecting the same relationship as was found in original learning, was related to list structure. With the study-test method the original learning was apparently based on very weak associations between stimulus number or position and words marking the conceptual changes which occurred throughout the list. The loss of these associations over 24 hr. essentially made the conceptual nature of the lists useless as a recall vehicle. The items which were recalled were apparently based upon factors specific to them.

Considering now the broader context of these studies, the question may be raised about the systematic importance of the earlier study (Underwood & Zimmerman, 1973), and the auxiliary experiment of the present report, in both of which the anticipation method was used. Recall and list structure were directly related in these studies. However, the error data from this previous study gave no support to the idea that forgetting of unstructured lists was produced by interference among the conceptual associations among the words. In the present study there was no evidence that conceptual associations played any interfering role in learning the unstructured lists. It does not seem now, therefore, that these types of lists have any special or pointed use for the study of interference as a source of forgetting.

The fact that the structure-recall relationship was found only with the anticipation method further reduces the value which one might place on the relationship as a basic and important fact of forgetting which needs theoretical appraisal. Under the anticipation method, if the S remembered something about the nature of the conceptual relationships involved (and they were almost inescapable), the information supplied on the recall trial by the anticipation method essentially served as a relearning trial. Corresponding information would not be given in the study-test method until the study trial after the recall trial. If recall following both methods of learning the structured lists was taken by the unpaced technique as used in the present experiment it is doubtful if differences for the two methods would have occurred. Or, if on the unpaced test every fourth word (the first instance of each new concept) was supplied, recall would improve markedly and equally following both methods of learning, and the improvement would be greater than for unstructured lists. These observations, if correct, indicate that the systematic problems in the study of forgetting will not be greatly illuminated by the use of conceptually structured lists designed to restrict potential placement positions in what is essentially serial learning. The more systematic issues appear to lie in studying the learning of structured lists, whether of a serial nature, free-recall, or some other type of task into which the conceptual structure may be inserted.

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, 1969, 80, (3 Pt. 2).

Bousfield, W.A., & Cohen, B.H. Masculinity-feminity in the free recall of a categorized stimulus word list. Perceptual & Motor Skills, 1956, 6, 159-166.

Underwood, B.J. Degree of learning and the measurement of retention. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 112-129.

Underwood, B.J. & Zimmerman, J. Serial retention as a function of hierarchical structure. Journal of Experimental Psychology, 1973, 99, 236-242.

Footnotes

[1]This research was supported by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract N00014-67-A-0356-0010, Contract Authority Identification No. NR 154-321.

[2]Requests for reprints should be sent to Benton J. Underwood, Department of Psychology, Northwestern University, Evanston, Illinois 60201.

Table 1

Comparison of Study-Test and Anticipation Methods with Regard to Concept

Concept Instance Learning and Recall for List 5

(See text for complete explanation)

	Anticipation		Study-Test	
	Last OL	Recall	Last OL	Recall
Three	117	162	210	102
Two				
1 & 2	22	14	18	18
1 & 3	28	20	18	10
2 & 3	62	68	18	8
One				
1	6	2	11	16
2	13	13	2	1
3	21	12	5	9

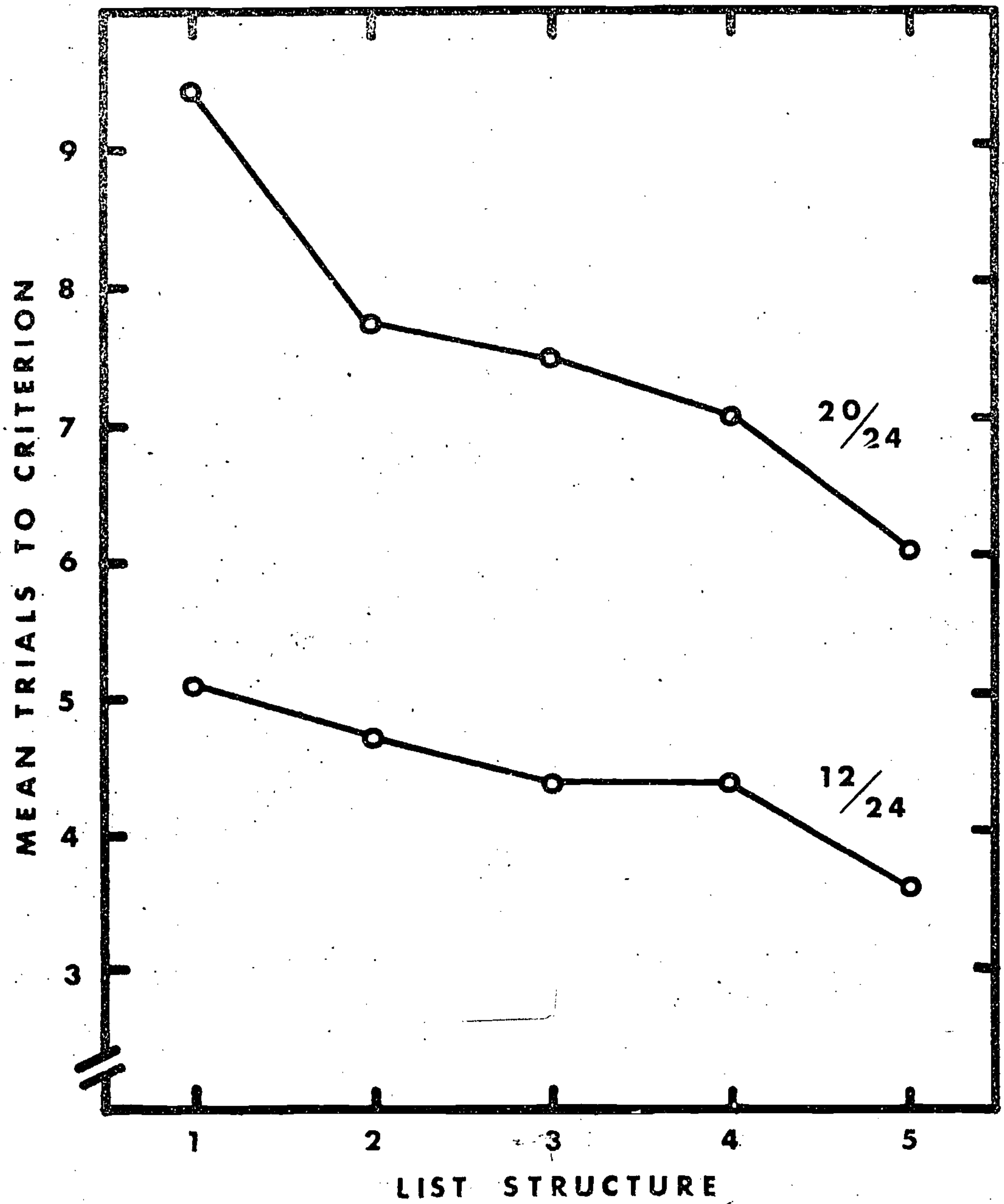

Fig. 1. Mean number of trials to reach two different criteria of learning (12 correct on a single trial, and 20 correct), for lists varying in conceptual structure from low (List 1) to high (List 5).

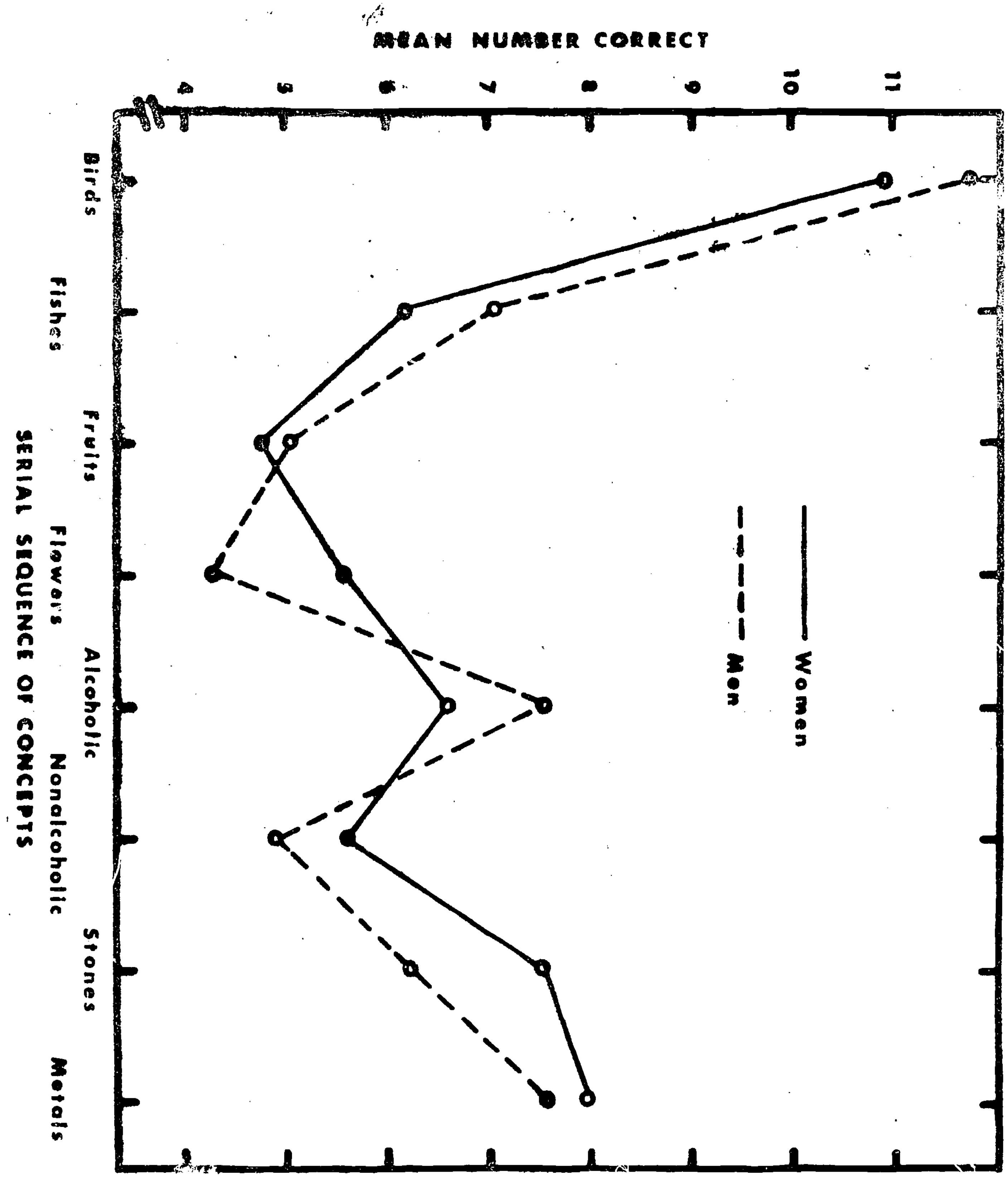

Fig. 2. Learning of the highly structured lists as a function of the particular concepts (and their positions in the list) by men and women.

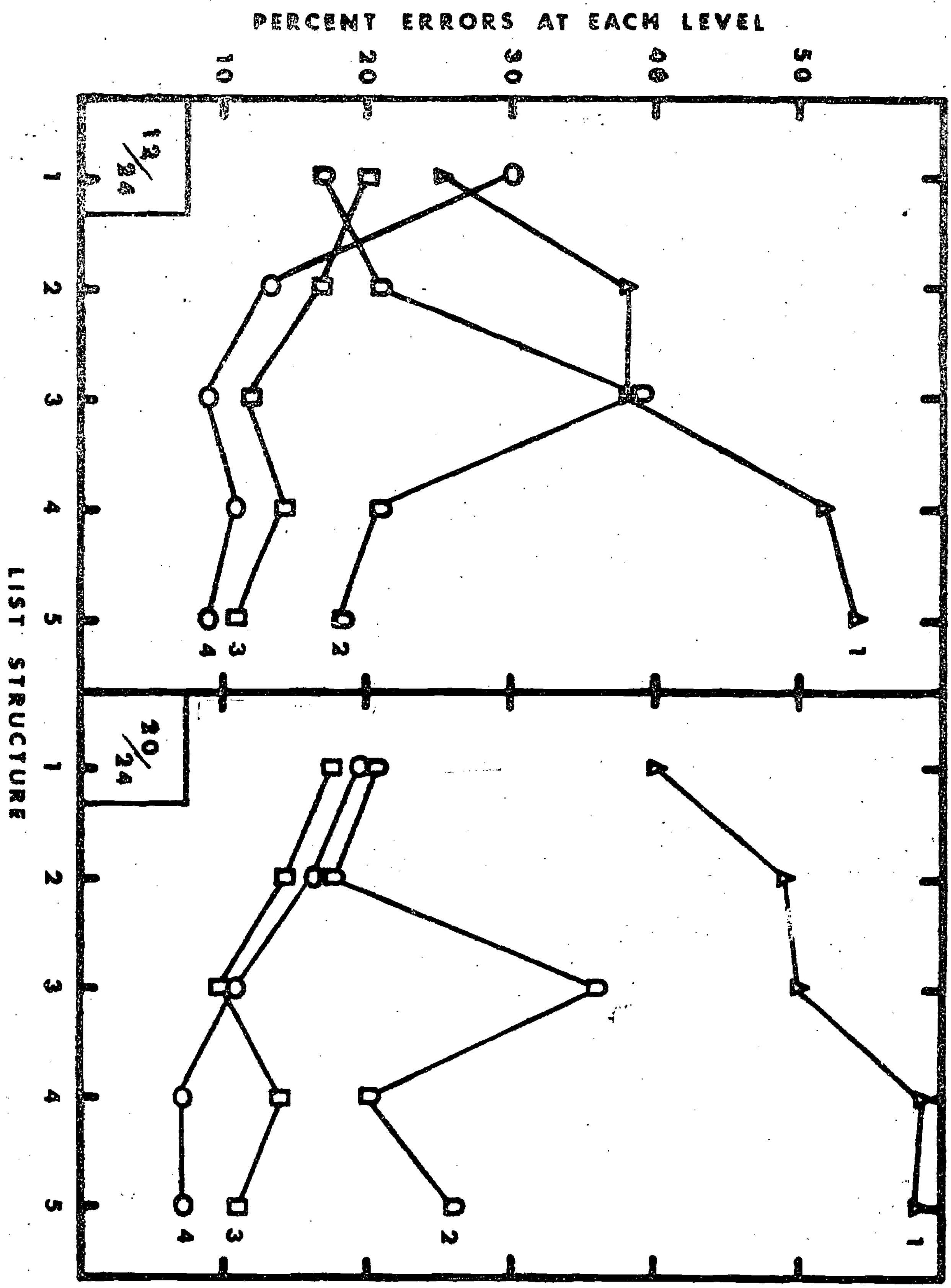

Fig. 3. Locus of overt errors as a function of list structure and degree of learning. The numbers (1, 2, 3, 4) represent an increase in the discrepancy between the locus of an error and its correct position as defined by the conceptual structure of Lists 4 and 5.

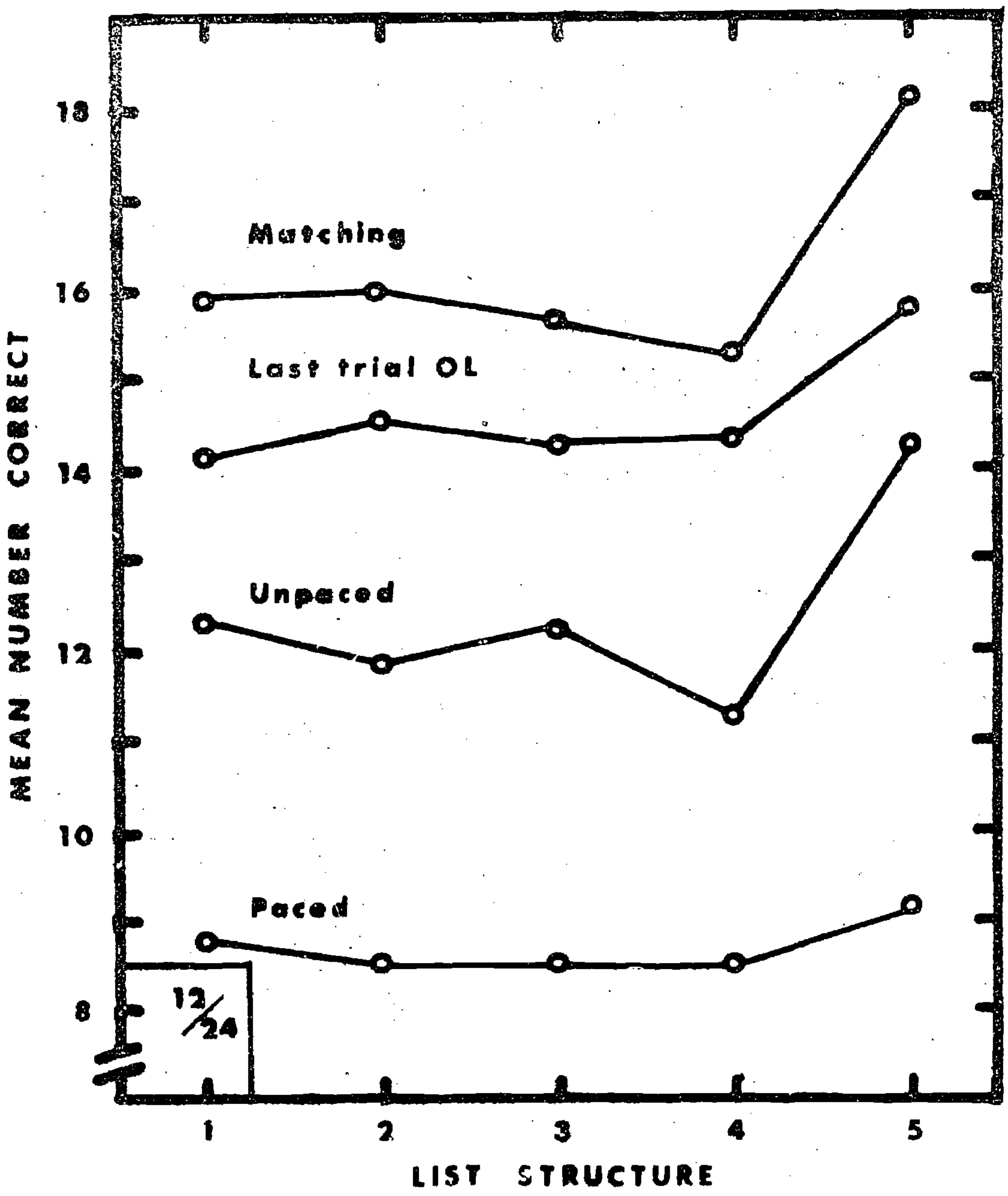

Fig. 4. Number of correct responses on the various retention tests after 24 hr. following original learning to 12 correct responses. Number correct on the last learning trial is also shown.

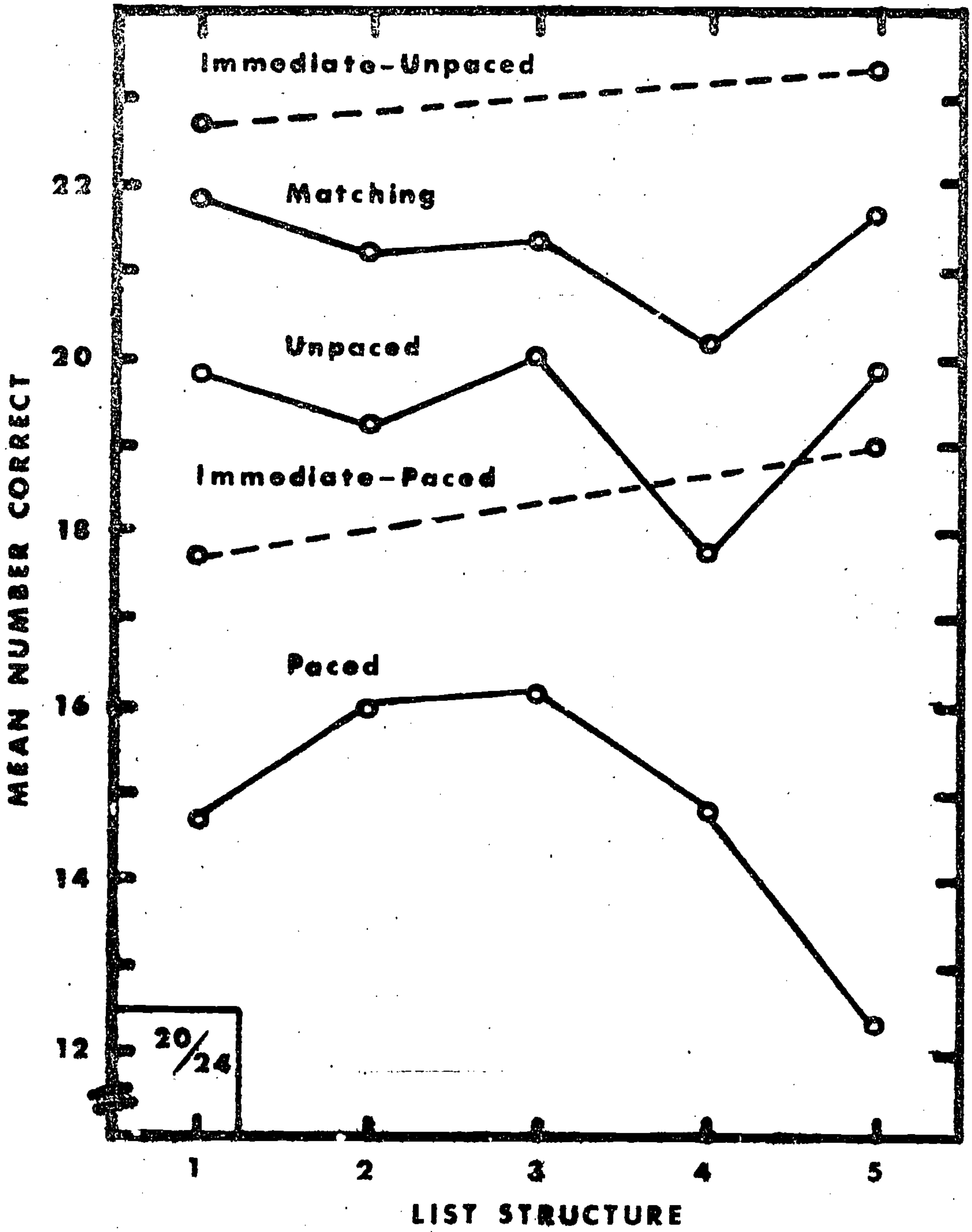

Fig. 5. Number correct on the various 24-hr. retention tests following original learning to 20 correct responses. Number correct on the immediate tests for Lists 1 and 5 is also shown.

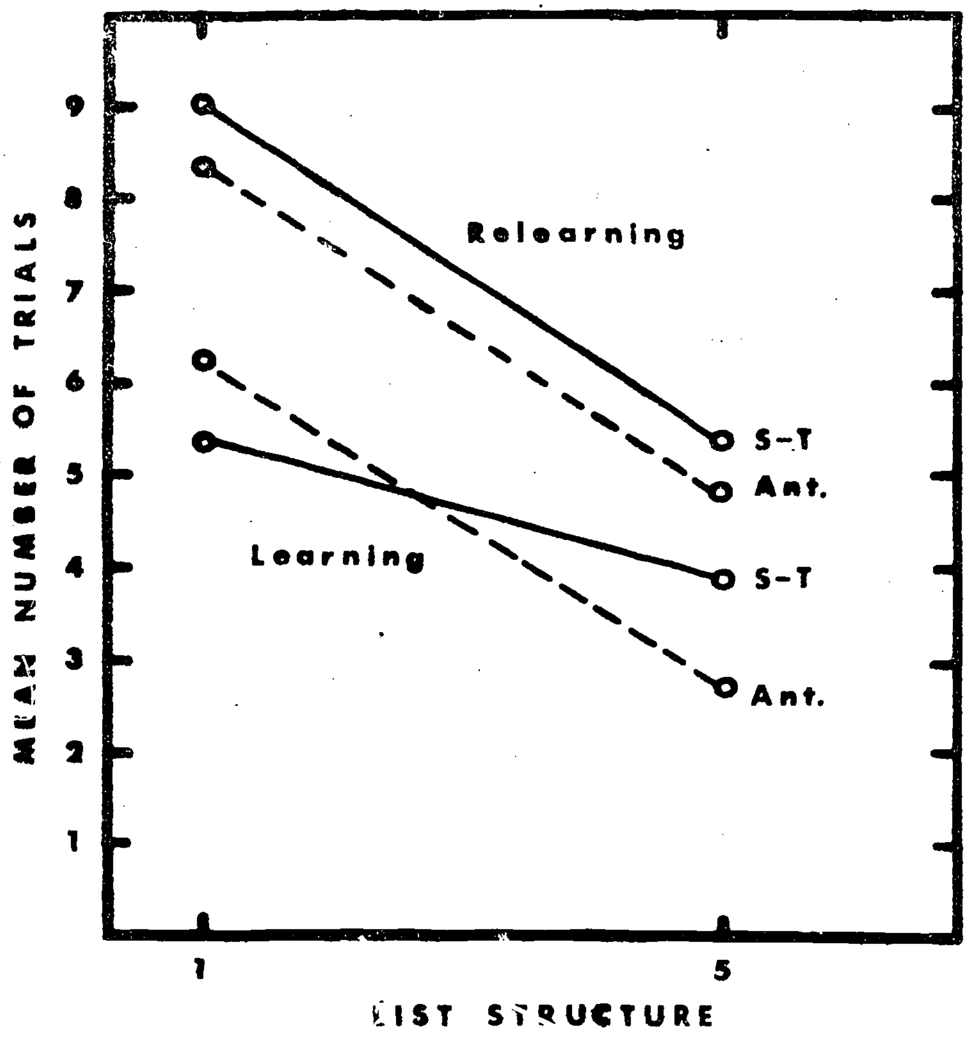

Fig. 6. Learning and relearning scores for Lists 1 and 5 as a function of the study-test (S-T) and anticipation (Ant.) methods.

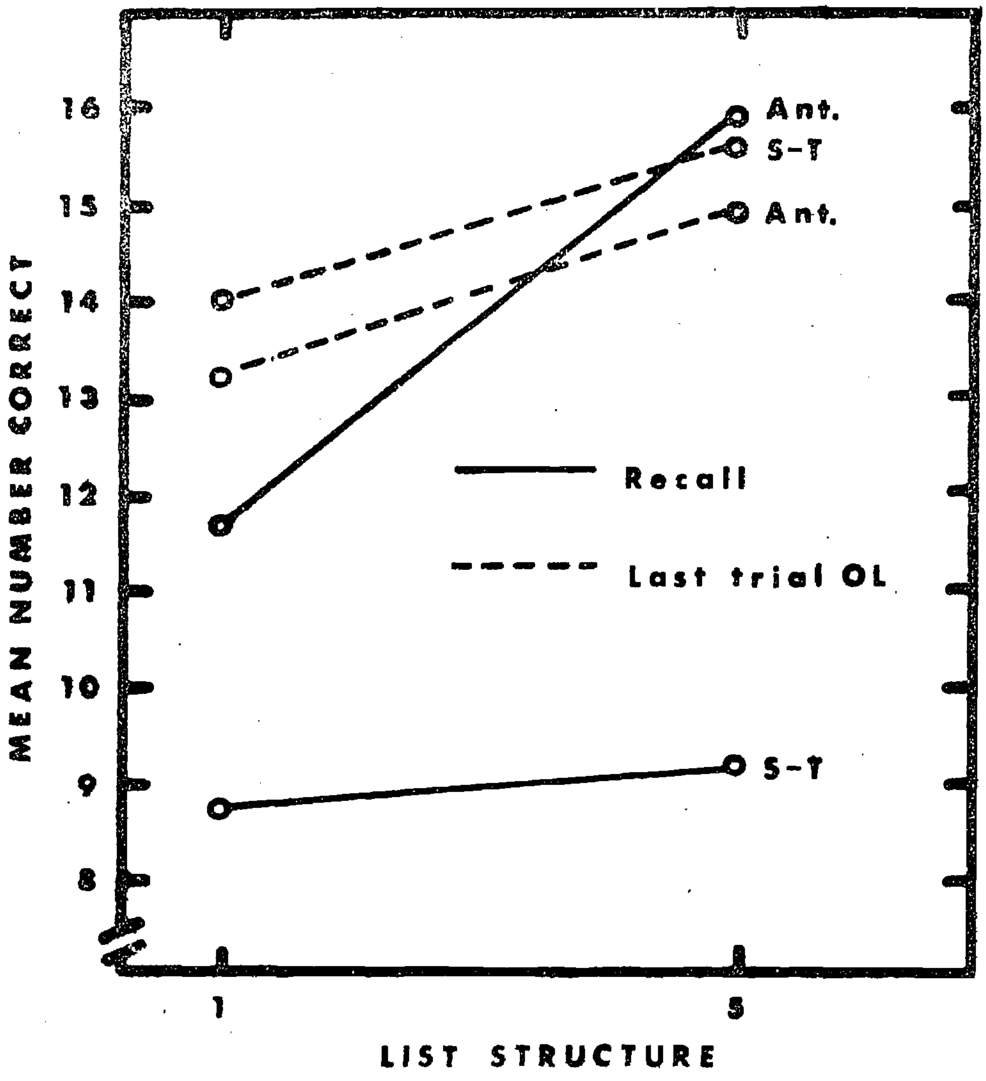

Fig. 7. Number correct on the last trial of original learning and number recalled after 24 hr. for the study-test and anticipation methods.

DISTRIBUTION LIST

Navy

4 Dr. Marshall J. Farr, Director
Personnel & Training Research Programs
Office of Naval Research
Arlington, VA 22217

1 Director
ONR Branch Office
495 Summer Street
Boston, MA 02210
ATTN: C.M. Harsh

1 Director
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101
ATTN: E.E. Gloye

1 Director
ONR Branch Office
536 South Clark St.
Chicago, IL 60605
ATTN: M.A. Bertin

6 Director
Naval Research Laboratory
Code 2627
Washington, DC 20390

12 Defense Documentation Center
Cameron Station, Bldg. 5
5010 Duke St.
Alexandria, VA 22314

1 Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054
ATTN: Dr. G.D. Mayo

1 LCDR Charles J. Theisen, Jr., MSC, USN
4024
Naval Air Development Center
Warminster, PA 18974

1 Mr. Lee Miller (AIR 413E)
Naval Air Systems Command
5600 Columbia Pike
Falls Church, VA 22042

1 Dr. Harold Booher
NAVAIR 415C
Naval Air Systems Command
5600 Columbia Pike
Falls Church, VA 22042

1 Commanding Officer
Naval Medical Neuropsychiatric Research
San Diego, CA 92152

1 Technical Reference Library
Naval Medical Research Institute
National Naval Medical Center
Bethesda, MD 20014

1 Dr. John J. Collins
Chief of Naval Operations (OP-987F)
Dept. of the Navy
Washington, DC 20350

1 Technical Library
Bureau of Naval Personnel
Dept. of the Navy
Washington, DC 20360

1 Dr. James J. Regan, Technical Director
Navy Personnel Research & Development Center
San Diego, CA 92152

1 Special Asst. to Chief of Naval Personnel
Bureau of Naval Personnel (Pers Oe)
Arlington Annex, Room 2403
Washington, DC 20370

1 Commanding Officer
Navy Personnel Research & Development Center
San Diego, CA 92152

1 Superintendent
Naval Postgraduate School
Monterey, CA 92940
ATTN: Library (Code 2124)

1 Technical Library
Naval Ship Systems Command
National Center, Bldg. 3
Room 3S08
Washington, DC 20360

1 Chief of Naval Training Support
Code N-21
Bldg. 45
Naval Air Station
Pensacola, FL 32508

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education & Training
Naval Training Command, Code 01A
Pensacola, FL 32508

1 Mr. Arnold Rubinstein
Naval Material Command (NMAT-03424)
Room 820, Crystal Plaza #6
Washington, DC 20360

1 Dr. Martin F. Wiskoff
Navy Personnel Research & Development
Center
San Diego, CA 92152

1 Dr. John Ford, Jr.
Navy Personnel Research & Development
Center
San Diego, CA 92152

1 Director of Research
U.S.Army Armor Human Research Unit
ATTN: Library
Bldg. 2422 Morade St.
Fort Knox, KY 40121

1 U.S. Army Research Institute for the
Behavioral & Social Sciences
1300 Wilson Blvd.
Arlington, VA 22209

1 Mr. Edmund F. Fuchs
U.S. Army Research Institute
1300 Wilson Blvd.
Arlington, VA 22209

1 Chief, Unit Training & Educational
Technology Systems
U.S. Army Research Institute for the
Behavioral & Social Sciences
1300 Wilson Blvd.
Arlington, VA 22209

1 Dr. Stanley L. Cohen
Work Unit Area Leader
Organizational Development Work Unit
Army Research Inst. for Behavioral &
Social Science
1300 Wilson Blvd.
Arlington, VA 22209

1 Dr. Leon H. Nawrocki
U.S. Army Research Inst.
Rosslyn Commonwealth Bldg.
1300 Wilson Blvd.
Arlington, VA 22209

Air Force

1 Dr. Martin Rockway
Technical Training Division
Lowry Air Force Base
Denver, CO 80230

1 Maj. P.J. DeLeo
Instructional Technology Branch
AF Human Resources Laboratory
Lowry AFB, CO 80230

1 Headquarters, U.S. Air Force
Chief, Personnel Research & Analysis
Division (AF/DPSY)
Washington, DC 20330

1 AFHRL/AS (Dr. G.A. Eckstrand)
Wright-Patterson AFB
Ohio 45433

1 AFHRL (AST/Dr. Ross L. Morgan)
Wright Patterson Air Force Base
Ohio 45433

1 AFHRL/MD
701 Prince St. Room 200
Alexandria, VA 22314

1 AFOSR(NL)
1400 Wilson Blvd.
Arlington, VA 22209

1 Headquarters Electronic Systems Division
ATTN: Dr. Sylvia R. Mayer/MCIT
L.G. Hanscom Field
Bedford, MA 01730

Marine Corps

1 Mr. E. A. Dover
Manpower Measurement Unit (Code A01M-2)
Arlington Annex, Room 2413
Arlington, VA 20370

Coast Guard

1 Mr. Joseph J. Cowan, Chief
Psychological Research Branch (P-1)
U.S. Coast Guard Headquarters
400 Seventh St., SW
Washington, DC 20590

Other DOD

1 Lt. Col. Austin W. Kibler, Director
Human Resources Research Office
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

1 Dr. Ralph R. Canter
Director for Manpower Research
Office of Secretary of Defense
The Pentagon, Room 3C980
Washington, DC 20301

Miscellaneous

1 Dr. Scarvia Anderson
Executive Director for Special Devpt.
Educational Testing Service
Princeton, NJ 08540

1 Dr. Richard C. Atkinson
Stanford University
Dept. of Psychology
Stanford, CA 94305

1 Dr. Ronald P. Carver
American Institutes for Research
8555 Sixteenth St.
Silver Spring, MD 20910

1 Century Research Corporation
4113 Lee Highway
Arlington, VA 22207

1 Dr. Kenneth E. Clark
University of Rochester
College of Arts and Sciences
River Campus Station
Rochester, NY 14627

1 Dr. Allan M. Collins
Bolt Beranek & Newman
50 Moulton St.
Cambridge, MA 02138

2 ERIC
Processing & Reference Facility
4833 Rugby Ave.
Bethesda, MD 20014

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Edwin A. Fleishman
American Institutes for Research
8555 Sixteenth St.
Silver Spring, MD 20910

1 Dr. Robert Glaser, Director
University of Pittsburgh
Learning Research & Development Center
Pittsburgh, PA 15213

1 Dr. Robert Tenneyson
Florida State University
Center for Computer-Assisted Instruction
Tallahassee, FL 32306

1 Dr. Henry J. Hamburger
University of California
School of Social Sciences
Irvine, CA 92664

1 Human Resources Research Organization
Division #3
P.O. Box 5787
Presidio of Monterey, CA 93940

1 Human Resources Research Organization
Division #4, Infantry
P.O. Box 2086
Fort Benning, GA 31905

1 Human Resources Research Organization
Division #5, Air Defense
P.O. Box 6057
Fort Bliss, TX 79916

1 Human Resources Research Organization
Division #6, Library
P.O. Box 428
Fort Rucker, AL 36360

1 Dr. Lawrence B. Johnson
Lawrence Johnson & Associates, Inc.
200 S Street, N.W., Suite 502
Washington, DC 20009

1 Dr. David Klahr
Carnegie-Mellon University
Graduate School of Industrial Admin.
Pittsburgh, PA 15213

1 Dr. Robert R. Mackie
Human Factors Research, Inc.
6780 Cortona Dr.
Santa Barbara Research Park
Goleta, CA 93017

1 Dr. Andrew R. Molnar
Technological Innovations in Education
National Science Foundation
Washington, DC 20550

1 Dr. Leo Munday, Vice President
American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Donald A. Norman
University of California, San Diego
Center for Human Information Processing
La Jolla, CA 92037

1 Mr. Luigi Petrullo
2431 N. Edgewood St.
Arlington, VA 22207

1 Dr. Diane M. Ramsey-Klee
R-K Research & System Design
3947 Ridgemont Dr.
Malibu, CA 90265

1 Dr. Joseph W. Rigney
Behavioral Technology Laboratories
University of Southern California
3717 South Grand
Los Angeles, CA 90007

1 Dr. George E. Rowland
Rowland & Company, Inc.
P.O. Box 61
Haddonfield, NJ 08033

1 Dr. David J. Weiss
Dept. of Psychology
University of Minnesota
Minneapolis, MN 55455

1 Dr. Anita West
Denver Research Institute
University of Denver
Denver, CO 80210

1 Dr. Kenneth Wexler
University of California
School of Social Sciences
Irvine, CA 92664

1 Dr. John Annett
The Open University
Milton Keynes
Buckinghamshire ENGLAND

1 Dr. Charles A. Ullmann
Director, Behavioral Sciences Studies
Information Concepts Incorporated
1701 No. Ft. Myer Dr.
Arlington, VA 22209

1 Mr. Dexter Fletcher
Dept. of Psychology
Univ. of Illinois at Chicago Circle
Box 4348
Chicago, IL 60680