Journal of Experimental Psychology
1972, Vol. 94, No. 3, 300–307

ORGANIZATIONAL UNITS IN FREE RECALL AS A SOURCE OF TRANSFER [1]

MARCIA K. JOHNSON [2]

State University of New York at Stony Brook

Experimental *S*s learned a paired-associate (PA) list composed of a subset of items from a previously learned free recall (FR) list. Items from adjacent positions in *S*'s FR output order were either used intact or re-paired in the PA lists; control *S*s' FR and PA lists contained no items in common. There was marked negative transfer in re-paired conditions and slight but generally nonsignificant facilitation when the pairs were from adjacent positions. This pattern of results was found for two methods of PA learning (PAL). After PAL, all experimental groups showed some evidence of a temporary disruption of recall of the original FR list; however, the decrement was not related to how the pairs were selected and there was no permanent loss. The results indicated that FR output order reflects item membership in subjective units, but evidence that recall depends on organizational stability was minimal.

There is considerable evidence that interitem relations develop during free recall learning (FRL) of "unrelated" lists. For example, the organization developed in FRL seems to be a source of interference when *S*s must develop a new free recall (FR) organization for part of the items (Novinski, 1969; Tulving & Osler, 1967) or when *S*s must learn new, prescribed associations among the items (Postman, 1971). Logically, the transfer results reported in these studies could be obtained if all items become equally associated with each other during FRL. On the other hand, several investigators (e.g., Mandler, 1967; Tulving, 1968) have pointed out that the recall order consistencies found in FRL suggest that a more segmented type of subjective organization occurs consisting of relatively independent units. If this is the case, it should be possible to manipulate the likelihood that items included in a second task come from the same organizational unit and therefore manipulate the degree and/or direction of transfer.

In the present study, an FRL to paired-associate learning (PAL) design was used to test the assumption that FR output order reflects unit membership. Pairs were composed either of items from adjacent positions or from nonadjacent positions from *S*'s FR output order. These experimental conditions were contrasted with control conditions in which the FR and PA lists contained no items in common. Negative transfer was expected to be greatest when the prescribed associations of PAL were the least consistent with the subjective organization of FRL.

The FRL to PAL design was set within a whole–part learning framework; a subset of the FR items was used in the PA list. After PAL, *S*s were required to recall the entire FR list. It was assumed that if original learning (OL) depended on the particular organizational units developed, then requiring *S*s to reorganize part of the items during PAL would produce an OL recall decrement relative to transfer control *S*s. Moreover, the more subjective units which were disrupted, the greater the recall decrement was expected to be.

METHOD

Design.—The basic design involved FRL followed by PAL, followed by a test of retention of

[1] This paper is based upon a dissertation submitted in partial fulfillment of requirements for the PhD degree at the University of California, Berkeley. The author would like to thank G. Keppel, W. D. Rohwer, Jr., and especially L. Postman for their helpful guidance. The research was supported by a Predoctoral Traineeship awarded to the Institute of Human Learning by the National Institute of General Medical Sciences, National Institutes of Health.

[2] Requests for reprints should be sent to Marcia K. Johnson, Department of Psychology, State University of New York at Stony Brook, Stony Brook, New York 11790.

FRL. For generality, two study-test methods of PAL which differed only in test trial procedure were used—free recall of pairs (R) and cued PAL (Q). In the R method, Ss were allowed to recall the pairs in any order. This method was used because it is the PA method which is most continuous with prior whole–part FRL to FRL designs. The Q method is the standard study-test procedure in which stimuli are presented on test trials and S is required to supply the correct responses. This method of PAL was included in order to obtain measures of PAL performance which were not obscured by failures to recall a unit. Within each method there were two types of PA lists—those designed to maintain (M) the FR organization and those designed to disrupt (D) the FR organization. Experimental (E) groups learned either M or D PA lists consisting of items from their FR lists. For transfer control (C) groups, FR and PA lists were made from separate pools of items; therefore in the control groups, organizational treatment (M or D) was a dummy variable required to equate for possible differences in M and D list difficulty. As an example of the notational system, an S in QDE received cued PAL of a disrupt list consisting of items from his FR list.

In addition to the eight transfer groups, there was an RI control group (Rest) which did not have PAL during the retention interval.

Materials.—Each of the three FR lists consisted of 32 common two-syllable English nouns. An attempt was made to minimize obvious relationships among the 96 words, and no more than 3 words beginning with the same letter were allowed in a given list. The lists were approximately equivalent in Thorndike-Lorge frequency and in rated concreteness (Paivio, Yuille, & Madigan, 1968). Each was used equally often in each condition as the OL list.

Paired-associate lists were constructed in the following manner: eight contiguous pairs (critical pairs) were selected from each E S's criterial trial output order. Two different selection patterns (1 and 2) were randomly assigned with the restriction that each be used equally often with each FR list. The pairs according to Selection 1 consisted of items in Output Positions 4–5, 7–8, 10–11, 13–14, 16–17, 19–20, 22–23, and 25–26; according to Selection 2, 3–4, 6–7, 9–10, 12–13, 15–16, 18–19, 21–22, and 24–25

The first item of a pair was always a stimulus and the second item a response in the resulting PA list. Pairs were used intact in the M groups; responses were reassigned to stimuli in D groups. For D groups, the pairs according to Selection 1 consisted of items in Output Positions 4–20, 7–17, 10–23, 13–5, 16–26, 19–11, 22–14, 25–8; according to Selection 2, 3–19, 6–16, 9–22, 12–4, 15–25, 18–10, 21–13, 24–7. In each case, an average of nine items had intervened between the stimulus and response in the criterial trial output order. A unique PA list was thus determined from each E S and subsequently used for a C S.

Procedure.—FR items were presented on a Carousel projector at a 2-sec. rate. Test trials consisted of 90 sec. for written recall; Ss were also required to call out the items as they wrote them down. Study and test trials alternated until S reached a criterion of 28/32. This was immediately followed by an additional, postcriterial, recall trial. An S was dropped if he failed to reach the criterion within 20 trials.

The OL–IL (interpolated learning) interval was 12 min. For 10 min., Ss practiced a motor learning task, which provided a rest for Ss and time for E to select the items and to make up the appropriate PA list. Two minutes were allowed for instructions for the IL task. The Rest group worked on arithmetic sequence problems for 10 min. Transfer groups had eight trials of PAL.

For both Q and R groups, the study-trial procedure was the same. Stimuli and responses were presented successively at a 1.5-sec. rate on the Carousel projector and each pair was separated from the next by a blank slide. On test trials, the first member of each pair was presented at a 4-sec. rate for the Q Ss, who were required to give the appropriate second member orally. R Ss had 32 sec. for oral FR of the pairs. The Ss were instructed prior to IL about the type of test to expect. Four random orders of each list were used equally often as starting orders in OL and IL. There were also four random test-trial orders of the stimuli in Q conditions.

A 3-min. retention test of OL was given immediately after the end of IL, and the entire retention interval was 22 min. The retention test was followed by three relearning trials in which the procedure was the same as during OL.

Subjects.—The Ss were 162 University of California, Berkeley, students. Sixteen Ss per group were paid for their participation; 2 Ss per group were psychology students satisfying a course requirement. The Ss were assigned randomly in blocks comprising the four E groups and the Rest group. The S arriving immediately after an E S was assigned to the C group requiring the PA list made for that E S. Eight Ss were dropped from the experiment (and immediately replaced) for failing to reach the OL criterion with 20 trials. The Ss who were dropped had been assigned to the following conditions: RME (2), RMC (2), QME (1), QDE (1), QDC (1), and Rest (1).

Results and Discussion

Original Learning

One-way analyses of variance were used to assess comparability of the nine groups. They did not differ significantly ($p > .05$) on any of the relevant measures. The measures and the overall means were: trials to criterion (9.34); total recalled on the criterial trial (28.62); total recalled on the postcriterial trial (27.92); backward and forward pair-wise intertrial repetitions

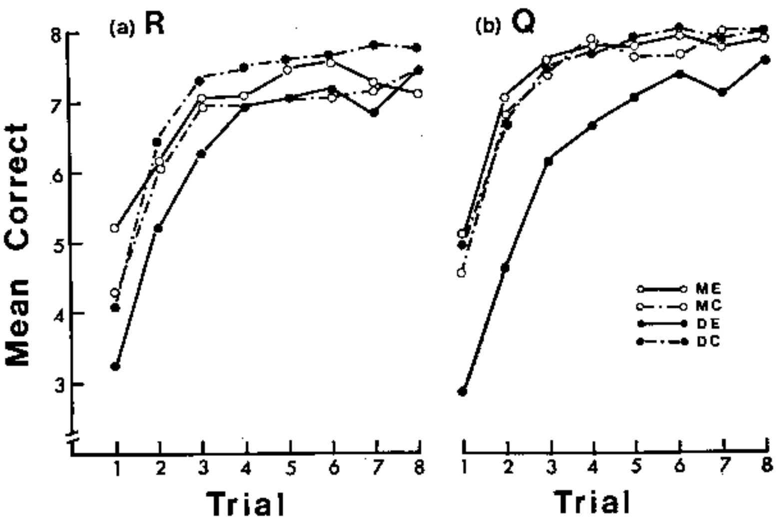

FIG. 1. Mean number correct pairs on successive transfer trials.

(ITRs) between the criterial and postcriterial trial, corrected for chance and expressed as a percentage of the maximum possible (32%); number of critical pairs which clustered on the postcriterial trial for chance (2.32).

Transfer

The evaluation of PAL performance involved first an overall comparison of the two methods of PAL (Q vs. R) and then within each method an assessment of the main effects of condition (E vs. C), the main effects of organizational treatment (M vs. D), and the interaction of condition with organizational treatment. To determine if significant transfer was obtained, planned comparisons were also made between each E group and its corresponding C group.

Learning curves for the R groups are shown in Fig. 1A, and those of the Q groups in Fig. 1B. In total correct over the eight trials, there was a general superiority of the Q over the R method of PAL: F (1, 142) = 6.32, $p < .05$.

Within the Q method, the main effect of condition reflects better performance of C compared to E groups: F (1, 68) = 7.22, $p < .01$. The main effect of organizational treatment indicated a superiority of M over D groups: F (1, 68) = 8.39, $p < .01$. More important, the Condition × Organizational Treatment interaction, F (1, 68) = 10.54, $p < .01$, indicates that the effect of prior FRL of items included in the PA list depended on whether or not the pairs were contiguous in *S*s' output order. The comparisons of each E group with its C group confirmed the general impression given by the learning curves. There was a slight but nonsignificant advantage for the QME group as compared to the QMC group. There was clear negative transfer in the QDE group as compared to the QDC group: F (1, 34) = 11.48, $p < .01$.

This general pattern of results was replicated in the R groups. Here, neither main effect was significant, but the critical interaction was obtained: F (1, 68) = 8.39, $p < .01$. Again, the positive transfer was not significant but the negative transfer was clear: RDE versus RDC, F (1, 34) = 13.00, $p < .01$.

In a separate analysis of the number correct on Trial 1, the general performance of Q and R groups was comparable: F (1, 142) < 1. Within the Q method, the Trial 1 results mirrored those for the analysis of the entire eight trials. Within the R method, a slightly different picture is given by the Trial 1 transfer analysis.

The positive transfer in RME as compared to RMC was significant: F (1, 34) = 4.24, $p < .05$. The negative transfer was not. In general, the chance of observing positive transfer over the eight trials was greatly reduced by a performance ceiling since learning was very rapid in the C groups.

Errors.—Error rates (the percentage of the total number of paired items given during the eight PAL trials which were incorrectly paired) tended to increase under D, but not under M, E conditions in both methods of PAL. Stimulus and response errors were equivalent in QDE (36 and 37, respectively). The higher error rates in the DE groups did not seem to be the result of *S*s' persistence in pairing the items according to the organization reflected in the criterial trial output order. In QDE, only six of the response errors were consistent with the contiguous critical pairs from FRL. In RDE, only 1 of 22 incorrectly matched pairs was a critical pair. The percentage of errors in second-list learning which represent correct pairs from the first list is typically much higher in an A-B, A-B_r paradigm involving successive paired-associate lists (e.g., Postman, 1964).

There was little evidence of interference from items not included in the PA lists (Other items). Other word intrusions during PAL were rare (a total of four in the combined E groups vs. two in the combined C groups in Q, and nine and two in E and C groups, respectively, in R). In addition, there was no evidence of negative transfer in either ME group, even for the most difficult items. Thus, E groups seemed to have had little difficulty restricting their responses to the appropriate set of PAL items; items not included in the PA list were rather easily suppressed (e.g., Postman, Stark, & Fraser, 1968).

Suppression of Other items when the tasks are reversed (transfer from PAL to FRL) is evidently more difficult. Wood's (1969) finding of increasing negative transfer as a function of the number of single items from pairs included in the FR list indicated greater associative interference from Other items than found in the present study. This apparent contradiction in results suggests an important difference between units formed in PAL and those formed in FRL. In PAL, units become more stable and the relationships developed between items are highly specific. In FRL, the multiple relationships an item may enter into leads to a less cohesive unit. Any particular link may be more easily "set aside" unless a potentially competing item is specifically carried into the PA task.

Competition from Other items does represent a potential source of interference, however. Therefore, the absence of positive transfer may not entirely be due to the above mentioned ceiling effect. Also, it should be noted that the whole–part design requires that *S* differentiate a subset of items from a larger set. Problems in list differentiation do not necessarily involve organizational factors. To the extent that nonorganizational cues mediate performance during FRL, any manipulation which would degrade the value of these cues would disrupt performance. Logically, eliminating some of the items from the list would require *S*s to develop a new "scale" for relative judgments among the items on the basis of such features as strength, difficulty, frequency of recall, etc. Thus, nonorganizational factors (as well as those attributable to interitem relations) may operate to offset positive transfer in M groups.

Recall

The statistical analysis of the recall scores of the original FR list was similar to that used above, with the addition of individual comparisons of each of the transfer groups with the Rest control in order to evaluate RI.

The total time allotted for recall was 3 min. The number recalled in 90 sec. was noted and separate analyses were conducted on the scores for these two intervals in order to determine whether there might be transitory as well as more persistent effects of the interpolated learning.

TABLE 1

MEAN NUMBER ITEMS RECALLED

Items	RME	RMC	RDE	RDC	Rest	QME	QMC	QDE	QDC
Pair									
90 sec.	14.28	13.56	13.56	13.50	13.83	13.39	13.56	12.67	14.33
3 min.	15.17	14.78	15.28	15.22	15.11	15.06	14.94	14.94	15.06
Other									
90 sec.	10.44	12.11	9.11	10.89	11.56	10.83	12.61	10.56	12.22
3 min.	13.22	13.61	13.00	12.89	13.56	13.72	14.22	13.44	14.00

Also, each *S*'s recall was divided into critical pair items (Pair) and the remaining (Other items). These two sets were analyzed separately. In general, retention was very good; although some of the differences reported below are reliable, their absolute size is quite small.

Pair items.—Recall of Pair items at 90 sec. (Table 1) was comparable for R and Q groups: F (1, 142) < 1. Within Q, E *S*s recalled fewer items than the C *S*s: F (1, 68) $= 5.18$, $p < .05$. Largely, this is the result of the poorer performance of the QDE group: QDE versus QDC, F (1, 34) $= 9.88$, $p < .01$. Within R, no significant comparisons were obtained. None of the eight transfer groups differed significantly from the Rest group. By the end of the 3-min. recall period, the apparent reduced availability of paired items in QDE disappeared and the performance of all nine groups was approximately equal.

Other items.—Mean recall of Other items is also shown in Table 1. At 90 sec., the R groups recalled fewer items than did the Q groups: F (1, 142) $= 4.72$, $p < .05$. Within Q, the E *S*s recalled fewer items than the C *S*s: F (1, 68) $= 13.83$, $p < .01$. Within R, the M groups were above the D groups: F (1, 68) $= 7.40$, $p < .01$. And, as in the Q analysis, the E *S*s recalled fewer items than the C *S*s: F (1, 68) $= 13.45$, $p < .01$. In both R and Q, each of the E groups was significantly below its control. The *F*s (1, 34) are: QME versus QMC, 6.25, $p < .05$; QDE versus QDC, 7.91, $p < .01$; RME versus RMC, 6.63, $p < .05$; RDE versus RDC, 6.84, $p < .05$. Although all of the E groups fell below the Rest, the difference was significant only in the case of RDE (Dunnett's test, [9, 153] = 3.77). At 3 min., these differences were diminished, but the R groups remained below the Q groups: F (1, 142) $= 7.48$, $p < .01$. The advantage of Q groups can probably be attributed to the fact that Q PAL involves a greater change in method of learning the two lists than does R; similar results have been reported previously (e.g., Wood, 1970).

When the two types of items were added together and total recall was analyzed, QME, QDE, and RDE produced fewer items at the 90-sec. recall than did the Rest (Dunnett's test [9, 153] = 2.78, 5.15 and 6.45, respectively).

In summary, the most notable result was the temporarily reduced recall of items which were not included in the PA lists found for each E versus C comparison. The Condition × Organizational Treatment interactions were not significant. There-

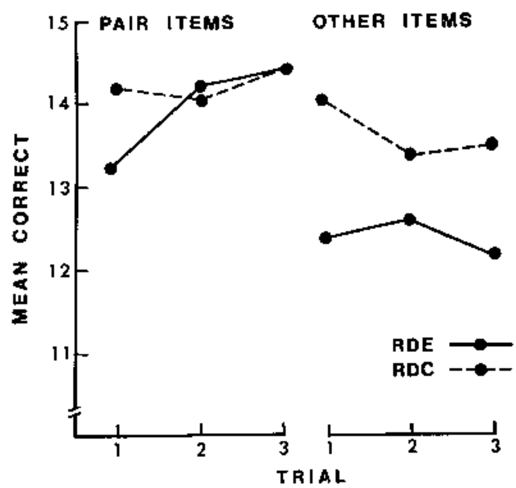

FIG. 2. Mean number correct words on successive relearning trials.

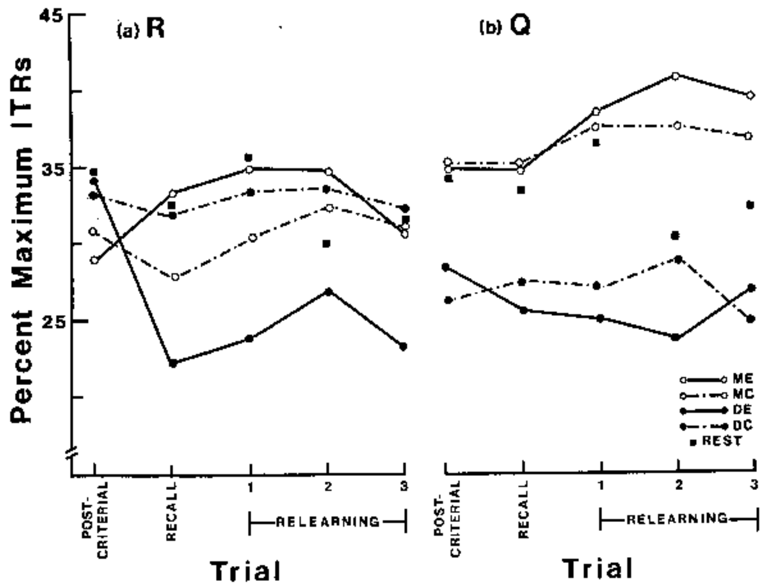

FIG. 3. Mean percent ITRs between criterial trial and each subsequent recall.

fore, there is no evidence that the D manipulation resulted in a greater loss of availability of the FR items than did the M manipulation.

Relearning

In the analysis of the total correct over the three relearning trials, the only significant comparison was RDE ($\bar{X} = 79.06$) versus RDC ($\bar{X} = 83.67$): $F(1, 34) = 4.31$, $p < .05$. From Fig. 2 it is apparent that most of this difference can be attributed to continued difficulty with Other items in the RDE group.

Organization

Figure 3 represents a comparison of the criterial trial output order with that of the recall and each of the relearning trials. The basic point of interest here was the extent to which *S*s used their original FR organization after the interpolated PAL. The percent total ITRs between the criterial and postcriterial trials is also shown in Fig. 3 so that differences in initial stability of the criterial organization might be taken into account.

Consider first the percent total ITRs between the criterial trial and the recall test for the R groups (Fig. 3A). The results indicate a degree of success in experimentally manipulating the original FR organization. The Condition × Organizational Treatment interaction is significant, $F(1, 68) = 8.64$, $p < .01$. Although the recall output order looks more like the criterial trial output order in RME than in RMC, the difference is not significant. The decrease in the RDE group of the OL organization on the recall trial is apparent: RDE versus RDC, $F(1, 68) = 8.47$, $p < .01$. For all R groups combined, the OL organization tends to recover and then decrease again during relearning; however, the RDE group remains below the other R groups.

The Q groups (Fig. 3B) present a less orderly picture. Here both QDE and QDC apparently had a less stable criterial organization than the other groups prior to the IL. Evidently this is a sampling difference. This is reflected in the recall trial in the main effect of organizational treatment, $F(1, 68) = 8.08$, $p < .01$. Neither condition nor Condition × Or-

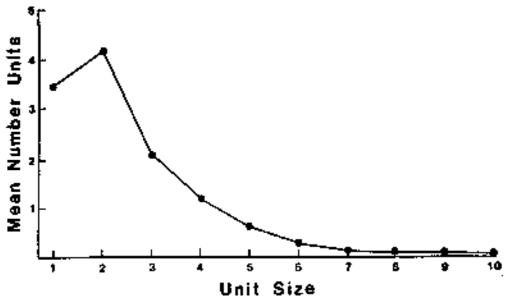

FIG. 4. Distribution of reported unit sizes.

ganizational Treatment is significant. During relearning, QME tends to show the same recovery-loss pattern found in the R groups. In contrast, for the QDE group there is a slight continued decrease in original organization in the first two relearning trials and then apparent recovery on the third trial.

Subject reports of subjective units.—At the end of the experiment, *S*s were asked to indicate the interitem relationships they had developed during FRL. In scoring *S*s' reports, the largest value was assigned whenever there was some ambiguity in the size of a unit. For example, if *S* indicated that two groups of words were connected by a single word, all of the words were assumed to belong to the same unit. Therefore, the results perhaps slightly overestimate the size of units and slightly underestimate the number. Since the forms of the distributions were very much alike in all the groups (including the Rest), all *S* reports were combined to get the best estimates of the mean number of units of various sizes. Figure 4 gives a rough picture, then, of the typical level of subjective organization. According to *S*s, the average list was organized into 12.1 discrete units. Most of the items were related to at least one other item—but there were 3.5 single-word units. The most probable unit size was two items. Large units (e.g., six or more items) were relatively rare.

Subject reports of related pairs.—The *S*s were also asked which of the paired items they had thought of as related prior to PAL from their FRL. The mean number of related pairs reported was 4.2 in ME groups and .8 in DE groups. This indicates that the selection procedure was probably more successful than indicated by the level of clustering (2.32) by critical pairs on the postcriterial trial.

In summary, the transfer analysis yielded the expected interaction of condition with organizational treatment for both R and Q methods of PAL. These interactions reflect the fact that there was marked negative transfer when contiguous items from each *S*'s FR output order were re-paired in the PA list and that there was a slight but generally nonsignificant facilitation of PAL when the pairs were composed of items which were contiguous in *S*'s FR output order. If the type of associative network developed resulted in every item being linked equally to every other item, the pattern of transfer reported here would be very unlikely. Some items, then, are more closely related than others; related items are likely to be recalled contiguously. The *S* report data confirmed this picture, but also indicated that the level of organization in the present study was less (at a fairly high degree of learning) than that implied by the usual theoretical description of organization based on a limited-retrieval-system model of memory (e.g., Bower, 1969, p. 610).

With respect to the question of the consequences of reorganization for retention, the results are more ambiguous. Recall of the original FR list was only temporarily depressed as a consequence of interpolated PAL. The small negative effect on retention did not appear to be related to the amount of associative interference during PAL or to the extent to which *S*s' recall order changed between the criterial trial of OL and the retention test. In addition, there was no permanent impairment of retention in E groups. Recently, evidence suggesting IL involving part of the OL items has little effect on OL retention in FR has also been reported by Wood (Exp. 1, 1970) and Ehrlich (1970). Within the limits of the present study, either recall does not depend on the particular organizational units established during OL, or output order stability is an inadequate index of whatever interitem relationships may be critical for retention.

The one indication that disrupting organization may have some permanent effect on performance was the relatively low score of the RDE group during the relearning trials. This may indicate that relearning is more sensitive to organizational changes than is recall and therefore is a more useful task for assessing the contribution of organizational stability to item retrieval.

On the other hand, during PAL which involves heavy associative interference, *S*s may develop a set to suppress competing responses which they then carry into FRL. Presumably, competition occurs during PAL when *S* must fix on the prescribed relationship when several, more highly practiced relationships may be available from prior FRL. These alternative links, which yield "flexibility of access" (Postman, 1971) in FRL, must be suppressed in order to meet the requirements of the PAL task. If these old associations are reactivated during relearning when the original list context is reintroduced, two consequences are logically possible. The *S* may utilize both old and new associations as alternative ways of linking the items. Or, *S* may attempt to fix on a specific relationship—in effect, setting up conditions of associative interference. In this latter case, then, *alternative* links would become *competing* links. If *S* develops a set to suppress potentially useful associations, a performance decrement would be expected.

REFERENCES

Bower, G. H. Chunks as interference units in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 610–613.

Ehrlich, S. Structuration and destructuration of responses in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 282–286.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*, Vol. 1. New York: Academic Press, 1967.

Novinski, L. Part–whole and whole–part free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 152–154.

Paivio, A., Yuille, J. C., & Madigan, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 1968, **76** (1, Pt. 2).

Postman, L. Studies of learning to learn: II. Changes in transfer as a function of practice. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 437–447.

Postman, L. Organization and interference. *Psychological Review*, 1971, **78**, 290–302.

Postman, L., Stark, K., Fraser, J. Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 672–694.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Tulving, E., & Osler, S. Transfer effects in whole–part free-recall learning. *Canadian Journal of Psychology*, 1967, **21**, 253–262.

Wood, G. Higher order memory units and free recall learning. *Journal of Experimental Psychology*, 1969, **80**, 286–288.

Wood, G. Organization, retroactive inhibition, and free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1970, **3**, 327–333.

(Received November 22, 1971)