

The Effects of Presentation and Recall of Material in Free-Recall Learning¹

ENDEL TULVING

University of Toronto, Canada

The results of two free-recall learning experiments, in which the number and sequential combinations of input and output phases were systematically varied, indicated that (a) recall tests facilitate the overall immediate post-input recall approximately to the same extent as do the events occurring in input phases, (b) input and output events are approximately equally responsible for forgetting of individual items from one output phase to another, (c) the earlier contention that intratrial retention remains constant over trials is untenable, and (d) recall of individual items is quite labile even when overall recall is stable. The findings were interpreted in support of the view of memory as a limited-capacity retrieval system in which the limit is set by the number, but not by the nature of the contents, of accessible memory units.

A trial in a typical verbal learning experiment consists of two operationally distinguishable phases. In the input phase material is presented to the S for study, and in the output phase its recall is tested. Input and output phases may be arranged in tandem, with all items presented in an uninterrupted input phase followed by a single output phase, as in the recall method of paired-associate learning and in free-recall learning; or, input and output phases associated with individual units of material may be interlaced within a trial, as in the anticipation method of serial or paired-associate learning. In every case, however, attempted retrieval of a given

item occurs after various numbers of presentations and successful or unsuccessful attempts to recall the same item and other items on preceding trials.

The fact that the probability of retrieval of an item increases as a function of number of prior inputs has been demonstrated repeatedly under a variety of experimental conditions and thus requires no additional confirmation as an empirical phenomenon. But little is known about the effects of prior recall tests on the subsequent recall of items. A recall test given only once obviously cannot influence the amount of learning that has taken place, or the amount and organization of the material that has been stored in memory prior to the test. If, on the other hand, recall tests occur as an integral part of the sequence of events preceding any given output phase—as is the case in typical multi-trial free-recall experiments—the amount and organization of the material in the memory storage, and hence the S's performance in a given output phase, may well be influ-

¹ This research was supported by Grant No. GB 810 from the National Science Foundation and Grant No. APT-39 from the National Research Council of Canada. The paper was prepared for publication during the author's tenure of a National Research Council of Canada Senior Research Fellowship at the Institute of Human Learning, University of California, Berkeley. The assistance of Marlene Segal and Marcia Ozier with various aspects of the research is gratefully acknowledged.

enced not only by prior inputs but also by prior outputs.

Two experiments reported in this paper were designed to separate and to compare the effects of prior presentations and prior recalls on subsequent recall in a multi-trial free-recall situation. The general approach was to vary the number and sequential combinations of input and output phases preceding a given recall test and to observe the effects of such variations on the overall recall scores as well as on the "fate" of individual items as manifested in the pattern of their recall over successive recall tests.

EXPERIMENT I

Method

Two groups of 22 Ss, female undergraduate students, learned a list of 22 nouns covering a wide range of frequency-of-occurrence values (List 2 in Tulving, 1964). Words were presented by means of a memory drum at the rate of 1 word/sec on 16 trials. The order of words was systematically changed from trial to trial so that no word occurred in the same serial position more than once, nor was it preceded by or followed any other word in the list more than once.

The Ss in Group S (Standard Group) were given 44 sec at the end of each trial to recall as many words as they could, in any order they wished. The Ss in Group R (Recall Group) were told at the end of the 44-sec output phase on the first trial to recall "all the words you remember once more" and were given another 44-sec period for this second recall. The same procedure was followed on all 16 trials. The Ss recalled the words orally. Their responses were written down by *E* and tape-recorded for later verification.

Results

The mean number of words correctly recalled over all 16 trials by Ss in Group S was 16.46. Group R recalled, on the average, 17.27 words in the first and 16.46 words in the second intratrial output phase. The learning curves for the two

groups are shown in Fig. 1. For Group R the data for both the first and the second recall period on each trial are plotted.

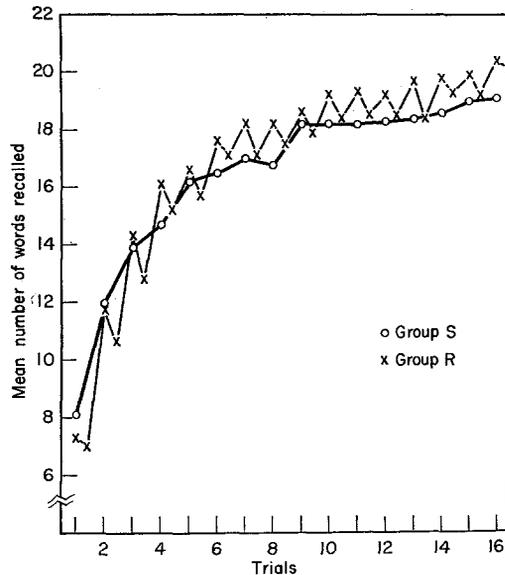


FIG. 1. Mean number of words recalled as a function of trials in Groups S and R (Exp. I).

An analysis of variance of the recall scores of Group S and of the recall scores from the first output phase of Group R showed that there was a significant interaction between groups and trials, $F(15, 630) = 1.77, p < .05$, indicating that recall increased over trials at a faster rate for Group R than Group S. The mean number of words recalled, averaged over 16 trials, was less in the second intratrial output phase than in the first one for 21 of 22 Ss in Group R, and, averaged over all Ss, less on all 16 trials, indicating reliable interphase "forgetting."

Comparison of the recall scores of the two groups with respect to trials ignores the fact that the total time spent on the task was greater for Group R than for Group S. Thus, these data do not provide unequivocal evidence for the facilitating effects of additional recall *per se*.

Apart from the very first trial on which the mean recall of Group S was slightly, but not significantly, higher than that of Group R, there were three output phases in Group S that coincided in time with the first intratrial output phases in Group R. These occurred on Trials 6, 11, and 16 in the former group, and on Trials 4, 7, and 10 in the latter. Prior to these output phases both groups had had equal amounts of total practice time, but Group S had spent a larger proportion of this time in studying the words and a smaller proportion in recalling them than had Group R.

The mean number of words recalled in these coincident output phases were 16.5, 18.2, and 19.1 in Group S, and 16.1, 18.2, and 19.2, respectively, in Group R. These data indicate that if the duration of prior practice is held constant, the level of recall is independent of temporal proportions between input and output phases. This conclusion holds, however, only if recall is measured immediately following the input of the material. The decrement in recall from the first intratrial output phase to the second in Group R implies that output events produce a certain amount of "forgetting." The invariance in the level of the immediate post-input recall in Groups S and R, therefore, is probably achieved through some facilitative mechanism that offsets such output interference in Group R.

EXPERIMENT II

The second experiment was designed to sharpen the contrast between presentation and recall, and to allow more direct comparisons between the effects of inputs and outputs on subsequent recall at comparable points of time in the overall sequence of events. In addition, longer lists were used to lessen possible "ceiling effects" that may have occurred in Exp. I.

Method

The Ss learned lists of 36 common nouns, randomly selected from among the five thousand most frequent words as tabulated by Thorndike and Lorge (1944). For all the Ss, the total duration of practice on a list was divided into six "cycles," each consisting of four 36-sec periods. Thus all groups were given 24 "practice" periods.

Three groups of Ss worked with different input and output sequences. For Group S (Standard Group) a cycle consisted of two input and two output phases, administered in an alternating sequence as in a typical FRL experiment. For Group P (Presentation Group) the first three 36-sec periods in each cycle were input phases and the fourth one an output phase. For Group R (Recall Group), the first period was an input phase and the next three output phases. Thus, every fourth period in the total sequence was an output phase for all groups. Since all periods were of equal length, 36 sec, these output phases coincided temporally in the overall sequence of events in all groups. In addition, in Groups S and R, the second period in each cycle was also an output phase.

In each group there were 18 Ss, undergraduate students of both sexes. The Ss were assigned to the three groups randomly, with the restriction that at no time did the number of Ss already tested in one group deviate from the number of Ss tested in other groups by more than one. Each S learned two lists, A and B, under the same conditions. Half the Ss in each group learned List A first and List B second, for the other half the order was reversed.

The Ss were tested individually. Words were presented by means of a memory drum, at the rate of 1 word/sec. The order of words in input lists and the order of word-orders in successive input phases were systematically varied for Ss in a given group.

Recall in output phases was oral. The intervals between successive phases, both within and between cycles, were 2 or 3 sec, just long enough for *E* to signal *S* that a new phase would begin. There were no intervals between consecutive input phases in Group P.

At the beginning of the experiment, all Ss were fully informed as to the nature of their task, the length of the list, the distribution of input and output phases relevant to their group, and the number of total cycles. All Ss were also told that the amount of time they would have for recall in each output phase was not very long and that they should, therefore, speak the words

they remembered quickly, but that they should not worry if they could not recall as many words as they thought they should, since they had a chance to improve their recall as they went on.

Prior to the learning of the first list, Ss were informed that they would learn two lists, one after the other. At the completion of the first list, Ss were given approximately 2 or 3 min of rest, and then the second list was presented under identical conditions.

Results

An initial analysis of the data showed that the mean number of words recalled over all conditions was approximately the same for Lists A and B, 16.09 and 15.73, respectively. Since this difference was relatively small and since no other systematic differences between the two lists were detected, no further distinction was made between Lists A and B in the analyses.

Overall Recall Data. The data were first examined with respect to the number of words correctly recalled in all output phases in all groups. These data are summarized in Fig. 2.

The curves of Groups S and P are rather similar to one another, while the curve of Group R lies slightly below the other two and has a serrated appearance. Figure 2 shows that the level of recall in the first intracycle output phases in Group R does not seem to be too different from that in the coincident output phases in Group S, but the level of recall does drop slightly in the second output phase in all cycles, remaining more or less constant in the third output phase. Mean number of words recalled over all six cycles was 15.90 and 17.06 in the first and second intracycle recall phases in Group S, 16.73 in Group P, and 15.12, 14.23, and 14.22 in the first, second, and third output phases, respectively, in Group R.

An analysis of variance was performed on the arcsin transformed numbers of words recalled in the output phases which

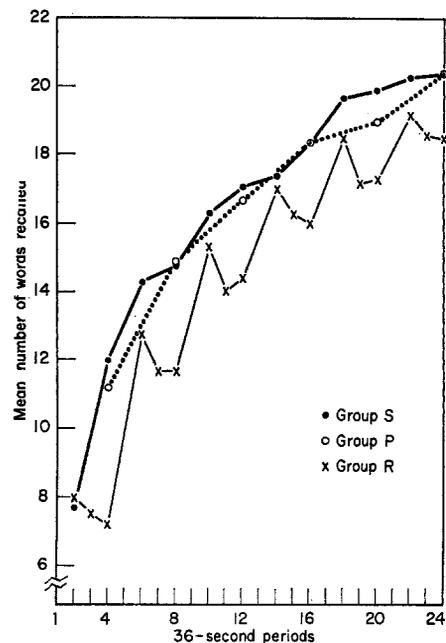


FIG. 2. Mean number of words recalled as a function of practice periods in Groups S, P, and R (Exp. II).

coincided temporally for all three groups, that is, in the *fourth period* of each cycle. The treatment variables in this analysis were Groups (S, P, and R), Order of Lists (first and second), and cycles (1 to 6). The results of the analysis showed that, in addition to cycles, significant sources of variance were Groups, $F(2, 51) = 3.91, p < .05$; Order of Lists, $F(1, 51) = 4.34, p < .05$; and the interaction between Groups and Cycles, $F(10, 255) = 2.54, p < .001$. None of the other interactions approached significance.

Thus, when we consider recall scores from the fourth period of each cycle, we can conclude that Group R recalled significantly fewer words than the other two groups and that the difference decreased over cycles. An incidental finding was the significant learning to learn or warm-up effect. The overall mean recall for the first list was 15.70, for the second list, 16.30.

A similar analysis of variance was performed on the number of words recalled in the first output phase within each cycle in Groups S and R, output phases that also coincided temporally. The only significant source of variance was associated with cycles. Thus there is no evidence that level of recall in the first intracycle output phases differs for Groups S and R.

"Fate" of Individual Items. The recall data were next analyzed with respect to patterns of recall of individual items over successive output phases. The set of items recalled by an S on each trial in a typical FRL experiment can be partitioned into two subsets: (a) those that were also recalled on the previous trial, and (b) those that were not recalled on the previous trial. With respect to performance on Trial n , the two components can be labeled $C_{n-1}C_n$ and $N_{n-1}C_n$ (or simply CC and NC). It has been argued elsewhere that the sizes of these two subsets of items provide estimates of intertrial and intratrial retention (Tulving, 1964).

In the present experiment, only Group S followed the typical pattern of alternating input and output phases. The two components of performance of this group are shown in Fig. 3. The upper panel shows mean $C_{n-1}C_n$, plotted against successive output periods (standard trials) on a logarithmic scale, where subscripts refer to the ordinal number of output periods. The lower panel shows mean $N_{n-1}C_n$, plotted against successive output periods on a linear scale. The straight lines whose equations are shown in Fig. 3 were fitted to the data by the method of least squares.

The CC and NC data depicted in Fig. 3 confirm the general picture of "fate" of individual items in a typical FRL task described previously (Tulving, 1964). Mean CC is an increasing logarithmic function of trials and mean NC is a decreasing linear function of trials. The equation of

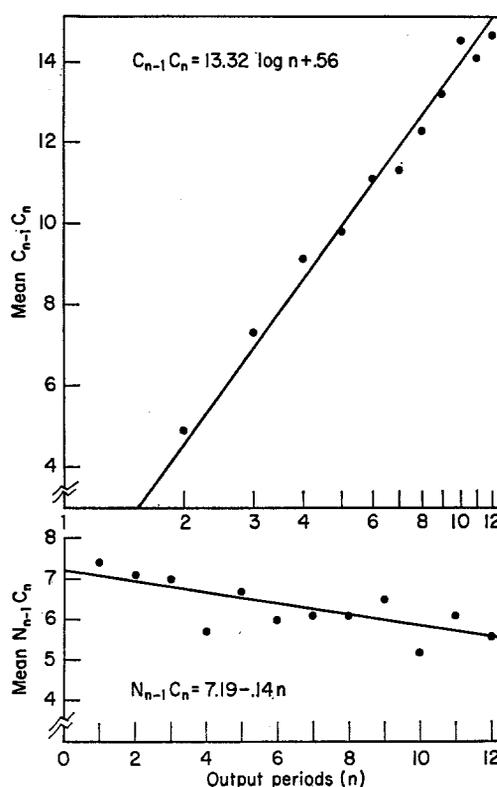


FIG. 3. Mean $C_{n-1}C_n$ (upper panel) and mean $N_{n-1}C_n$ (lower panel) over 12 output phases in Group S (Exp. II).

the standard learning curve is given by the algebraic sum of the equations of the CC and NC functions, in the present case, $R_n = 13.32 \log n - .14n + 6.63$, where R refers to the mean number of words recalled and n refers to the ordinal number of the output phase.

The $C_{n-4}C_n$ data for Group P are shown in Fig. 4, together with the corresponding data from the coincident output phases in Group S. In each case, $C_{n-4}C_n$ represents the number of words correctly recalled in the fourth period in a given cycle that were also recalled in the fourth period in the preceding cycle. Subscripts here refer to the ordinal number of successive 36-sec periods rather than just the output periods. The two functions are so similar that no

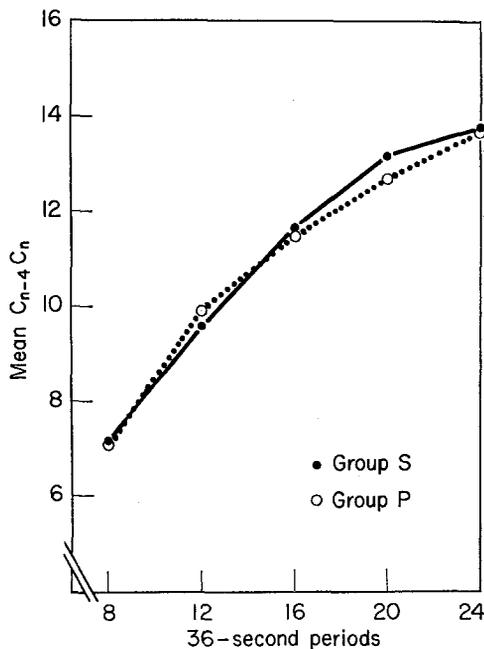


FIG. 4. Mean $C_{n-4}C_n$ as a function of practice periods in Groups S and P (Exp. II).

further analysis was undertaken. The similarity of the two functions suggests that intertrial retention (Tulving, 1964) is independent of the nature of the events, inputs or outputs, interpolated between two recall tests.

There are eight possible patterns of recall of a word over the three output phases *within* a cycle in Group R. These eight

patterns and the mean number of words associated with each of them are shown for all six cycles in Table 1. Recall pattern $N_1N_2N_3$ means that the word was not recalled in any of the three intracycle output phases, pattern $N_1N_2C_3$ means that the word was recalled in the third phase, but not in the first two, and so on.

The data in Table 1 indicate that there was a considerable amount of fluctuation in recall of individual words over successive recall phases. Of all the words recalled at least once within a given cycle, only approximately 50% were recalled in all three output phases.

Thus, although the overall level of recall in the first two intracycle output phases differed only slightly, though significantly, and in the second and third phases did not differ at all in a discernible fashion, this relative invariance in recall was to a large extent a consequence of trading words "forgotten" in a given output phase against those "recovered" from earlier phases. A simple retabulation of the data given in Table 1, for instance, shows that in Cycle 6 the mean number of words "forgotten" from the second to the third output phase (mean C_2N_3) was 3.97, while the mean number of words "recovered" in the third phase from earlier phases (mean N_2C_3) was 3.89.

TABLE 1

PATTERNS OF RECALL OF INDIVIDUAL WORDS WITHIN THREE OUTPUT PHASES OF SIX CYCLES IN GROUP R

Recall pattern	Cycles						Mean
	1	2	3	4	5	6	
$N_1N_2N_3$	26.42	20.33	16.86	14.22	12.08	11.39	16.88
$N_1N_2C_3$.33	.81	1.03	1.28	1.39	1.67	1.08
$N_1C_2N_3$.25	.47	.78	.78	1.22	1.14	.77
$N_1C_2C_3$	1.00	1.64	2.03	2.75	2.78	2.61	2.14
$C_1N_2N_3$	1.50	2.25	2.19	2.56	2.64	2.11	2.21
$C_1N_2C_3$.22	.94	1.89	1.67	2.64	2.22	1.60
$C_1C_2N_3$.67	1.19	1.72	2.42	2.75	2.83	1.93
$C_1C_2C_3$	5.61	8.36	9.50	10.33	10.50	12.03	9.39

The intracycle $C_{n-2}C_n$ and $N_{n-2}C_n$ components of performance in Group R can be compared with the same two components from coincident output periods in Group S. Subscript n within a given cycle refers to the fourth temporal period, subscript $n - 2$ to the second period. These periods were output phases in both Groups S and R. In Group S the two output phases were separated by an input phase, while in Group R they were separated by another output phase. The comparison would therefore demonstrate the effects of presentation and recall of material on what have been regarded as measures of intertrial and intratrial retention in a typical FRL experiment (Tulving, 1964). The relevant data are summarized in Fig. 5. The 36-sec periods for which numerical values are shown on the abscissa correspond to six cycles.

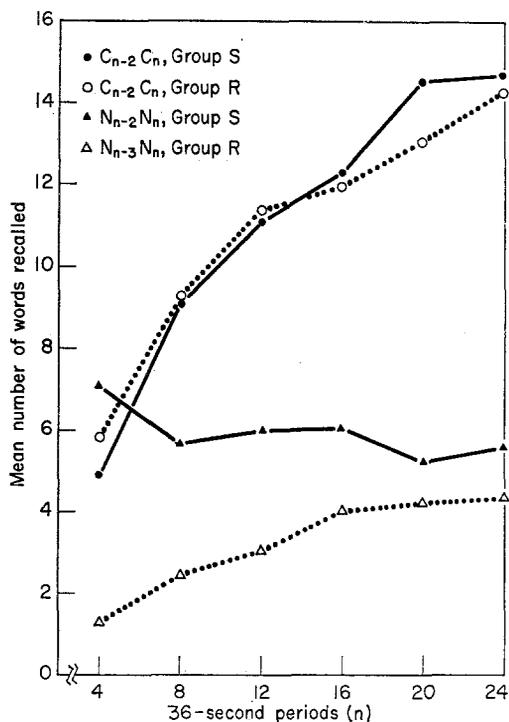


FIG. 5. Mean $C_{n-2}C_n$ and mean $N_{n-2}C_n$ as a function of practice periods in Groups S and R (Exp. II).

The two CC curves in Fig. 5 are rather similar to each other. Statistical analysis showed that neither the groups nor the Groups by Cycles interaction were significant sources of variance. The two NC functions shown in Fig. 5, however, are completely different. Of special interest is the finding that a sizable proportion of words not recalled in the first intracycle output phase were recalled in the third intracycle output phase in Group R, despite the absence of an interpolated input phase.

DISCUSSION

Evidence from both experiments suggests that the amount of immediate post-input recall, that is, the number of words recalled in an output phase immediately following an input phase, depends primarily on the total amount of time spent on the task, and that it is relatively little affected, if at all, by the distribution of this time between studying and recalling the material. This finding clearly implies that a recall test in FRL serves other functions beside that of measuring the amount or degree of learning. For one thing, it seems to facilitate subsequent immediate post-input recall approximately to the same extent as does study of the material in input phases, at least within the limits of the experimental conditions examined in the present experiment.

This conclusion may seem paradoxical in the light of observed decrements in recall from the first to the second intratrial or intracycle output phases in Group R, both in Exps. I and II. But the paradox resolves itself if we note that increments in recall from the final output phase in Cycle n to the first output phase in Cycle $n + 1$ were greater in Groups R than in Groups S. In Exp. II, for instance, the mean amount of such increase over all six cycles was 3.21 words, while the corre-

sponding increase in Group S, calculated for coincident periods, was only 1.17 words. Thus, intracycle decrements in recall in Group R were somehow compensated by larger intercycle increments.

Postman and Phillips (1965) have demonstrated sizable losses in recall over intervals of 15 and 30 sec, interpolated between a single presentation and a single recall test and filled with backward counting of numbers, these losses being limited mainly to words in the final input positions. It is possible that similar "forgetting" occurs in intervals filled with presentation or recall of words from the list being memorized. Thus, as has been argued elsewhere (Tulving, 1964), a trial in a typical FRL experiment may serve still another function beside providing the learner with an opportunity to study and retain individual items. Certain events occurring during the input and output phases of a trial must also be responsible for forgetting of some previously recalled items. Some evidence as to the amount of forgetting within and between trials is provided by the data on patterns of recall of individual items.

Before we briefly discuss some of these data, it should be pointed out that the finding of approximately equal overall facilitating effects of presentation and recall on subsequent recall is at variance with the results of a well-known experiment by Gates (1917). Gates found that time spent in active recitation of the material (recall) was more effective than time spent in reading the material. Because of the differences in Ss, materials, and procedure, possible reasons for the discrepant findings cannot be pinpointed.

Three main points are worth making with respect to the data on recall patterns of individual words. First, the breakdown of trial-by-trial recall scores into the CC and NC components in Group S in Exp.

II confirmed the shapes of the CC and NC functions reported previously for a standard FRL task in which input phases alternate with output phases (Tulving, 1964). In addition, the general similarity of the CC data based on pairs of output phases separated by different kinds and sequences of other phases (Figs. 4 and 5) lends plausibility to the notion that input and output events are approximately equally responsible for forgetting of individual items from one output phase to another. Since a simple decay theory cannot readily handle data such as those summarized in Table 1, such forgetting of individual items from one output phase to another should be attributed to interference, with the tentative conclusion that input and output interference effects in an FRL experiment are approximately equal in magnitude.

The second point concerns the identification of the NC component in a typical FRL task with intratrial retention that was suggested in an earlier paper (Tulving, 1964). Figure 5 shows that a sizable NC component was found even in absence of an interpolated input phase in Group R, thus implying that the NC component in Group S, working under the standard FRL conditions, must also include many words retained not from the immediately preceding input phase but rather from some earlier phase. A more reasonable estimate of intratrial retention in Group S might be provided by the difference between the NC functions for Groups S and R. This difference clearly decreases over trials. Thus, the earlier contention that intratrial retention remains relatively constant over trials (Tulving, 1964) seems to be untenable.

The third point of interest derives from the intracycle recall patterns of individual words in Group R in Exp. II. The findings, summarized in Table 1, gave evidence of

a truly remarkable instability of recall of individual words in the face of a rather stable overall recall performance in the three output phases within a given cycle. As noted earlier, only approximately 50% of all the words recalled at least once within a cycle were consistently recalled in all three output phases. To what extent such labile recall of individual words was a consequence of short durations of output phases, of course, is not known. With oral recall, however, Ss certainly had enough time to say more words in an output phase than they ever did. Moreover, data suggestive of similar variability in recall have been reported by Brown (1923) under conditions where Ss had ample time for recall.

The data from the present study have cast serious doubt on the usefulness of the conceptually appealing and frequently, albeit implicitly, adopted view according to which the "strength" of items in an FRL experiment increases as a consequence of "reinforcement" in input phases and that recall tests simply provide a running account, over trials, of "degree of learning" of the list. First, recall tests seem to both facilitate and interfere with subsequent recall of individual items, their overall effect being very much the same as that of the events in input phases. Secondly, a list always consists of individual items and hence the concept of degree of learning of a list should be translatable into a concept such as strength of an average item. But the analysis of recall patterns has clearly shown that the "average" item is a highly abstract and elusive entity having no readily identifiable counterparts in the empirical realm.

On the more positive side, the findings of the present study may be regarded as having added support to the notion that the limitation of the information processing mechanism of memory lies primarily in re-

trieval rather than in storage (Miller, Galanter, and Pribram, 1960). Both the present data and those from other experiments (e.g., Tulving and Pearlstone, 1966) have suggested that at any given time many more items are available in storage than are accessible for recall. The present data also suggest that retrieval is limited by the number of accessible memory units and not by their contents. In Exp. II, the number of words recalled in the third intracycle output phase in Group R that were not recalled in the second phase rather accurately matched the number of items that were recalled in the second but not in the third output phase. One might perhaps attribute the occurrence of such a "trading" relation to oscillation at the threshold of recall, but whether we talk about oscillation or about a limited-capacity retrieval system that is insensitive to the contents of accessible memory units, the major theoretical problem remains the same—how to conceptualize the events occurring in input and output phases and to account for the effects of these events on availability and accessibility of individual items. Only if these events and their effects on individual items are understood can we understand the processes involved in memorization of a collection or a list of items.

REFERENCES

- BROWN, W. To what extent is memory measured by a single recall? *J. exp. Psychol.*, 1923, 6, 377-382.
- GATES, A. I. Recitation as a factor in memorizing. *Arch. Psychol.*, 1917, 6, No. 40.
- MILLER, G. A., GALANTER, E., AND PRIBRAM, K. H. *Plans and the structure of behavior*. New York: Holt, 1960.
- POSTMAN, L., AND PHILLIPS, L. W. Short-term temporal changes in free recall. *Quart. J. exp. Psychol.*, 1965, 17, 132-138.
- THORNDIKE, E. L., AND LORGE, I. *The teacher's word book of 30,000 words*. New York: Teachers College, Columbia University, 1944.

- TULVING, E. Intratrial and intertrial retention: Notes towards a theory of free recall verbal learning. *Psychol. Rev.*, 1964, **71**, 219-237.
- TULVING, E., AND PEARLSTONE, Z. Availability versus accessibility of information in memory for words. *J. verb. Learn. verb. Behav.*, 1966, **5**, 381-391.
- (Received June 1, 1965)