Journal of Experimental Psychology
1974, Vol. 103, No. 5, 1014-1018

TO ORGANIZE IS TO REMEMBER:
THE EFFECTS OF INSTRUCTIONS TO ORGANIZE AND TO RECALL[1]

PETER A. ORNSTEIN[2]
University of North Carolina at Chapel Hill

TOM TRABASSO
Princeton University

PHILIP N. JOHNSON-LAIRD
University of Sussex, Brighton, Sussex, England

Mandler's 1967 finding that organization and recall instruction led to the same degree of recall was examined in two experiments, one employing taxonomically related materials and the other, unrelated words. To minimize anticipation of a recall test unless *S*s were specifically instructed, degree of organization was varied within *S*s, while recall instruction was a between-*S*s variable. A concept learning task was employed such that all *S*s classified some, but not all, of the words on a list that was subsequently tested for recall. Only one half of the *S*s were given prior instructions about the recall task. In both experiments organized material was recalled better than nonorganized material, and prior recall instructions did not lead to further increments in the recall of the organized items. However, these instructions did not invariably lead to enhanced recall of the items that were not classified; in fact, recall instructions seemed to be effective only when it was possible for *S*s to organize implicitly the noncategorized items. In this sense, organization and recall instructions are not invariably equivalent.

Mandler (1967) has shown that instructions to organize and instructions to recall have equivalent effects. Each of these types of instruction improved the recall of a list of words, but no further improvement was evident when both instructions were given. The nonadditive nature of these instructional manipulations motivated the present study, in which Mandler's finding is reexamined with a sensitive within-*S*s procedure.

In Mandler's experiment, four independent groups of *S*s were created by the combination of the presence or absence of instructions to organize or to recall a set of 52 unrelated words. Following five trials of presentation, all *S*s were instructed to recall the items. The results indicated that both types of instruction produced equivalent levels of recall, superior to that when neither recall nor organization instructions were given. However, an interaction was clearly implied by the data (Table III, p. 356), such that instructions to recall facilitated those *S*s not given categorization instructions (32.8 vs. 23.5 words recalled), but did not facilitate those *S*s instructed to organize (31.4 vs. 32.9).

The design of the present experiments was suggested by an analysis of the task presented to Mandler's *S*s, particularly those given neither set of instructions. It seems likely that these *S*s, whose task was simply to write the presented words on successive columns of a booklet, would have expected to be asked for recall later, regardless of their instructions. Indeed, Mandler reports that half of the *S*s in the no-category/no-recall condition did expect a recall test. Could this also have been true for the *S*s given categorization instructions but no recall instructions? If this were the case, clearly, the conclusion of the equivalence of categorization and recall instructions could be questioned. In addition, Mandler reports that very few of the *S*s in the categorization groups were able to

[1] This research was supported in part by Grant MH 19223 from the National Institute of Mental Health, U.S. Public Health Service, to T. Trabasso. Appreciation is expressed to Norma Johnson for assistance in conducting these experiments.

[2] Requests for reprints should be sent to Peter A. Ornstein, Department of Psychology, Davie Hall, University of North Carolina, Chapel Hill, North Carolina 27514.

reach a criterion of organizational consistency. It is possible that instructions to categorize might be facilitative above the level of instructions to recall if *S*s were taken to a strong criterion of organizational consistency.

In the present study an attempt was made to replicate Mandler's findings under conditions that appeared to be plausible for all *S*s. A concept learning task was utilized such that each *S* overtly organized some, but not all, of the words in a list. The effects of organization were then assessed by comparing the subsequent recall of the organized and unorganized words. Since the comparison was made within *S*s, individual differences in recall ability were controlled, and *S*s were easily taken to a strong organizational criterion. The effects of instructions to recall were assessed by instructing one group of *S*s (the "recall" group) to expect to recall the words, and by giving no such instructions to another independent group of *S*s (the "no-recall" group). The concept learning task was assumed to considerably reduce the chance of the no-recall group anticipating the recall test. Stimulus materials that differed in terms of the ease with which organization could be imposed were utilized in two experiments. In Experiment I, related items drawn from taxonomic categories were employed as stimuli, while unrelated materials were presented in Experiment II.

Experiment I

Method

Design and materials. All *S*s were presented with a list of 50 words, composed of 10 categories of 5 items each, and were told that the experiment dealt with learning to classify the words. Of the 50 items presented, 25 (5 items in each of 5 categories) were to be actively categorized by *S*, while the remaining 25 items were to be grouped together, thus ignoring the category structure. Active categorization was achieved by learning to associate an arbitrary category identification number with each of the 25 words that were to be categorized. One group of 10 *S*s (the recall group) was told that the recall of all of the items would be tested following the categorization task, while another group of 10 *S*s (the no-recall group) was not so instructed.

The words were 50 nouns taken from the category norms of Battig and Montague (1969). Five instances were selected from each of 10 common but nonconfusable categories such that the distribution of rank within category was approximately the same across the categories. The mean ranks ranged from 12.2 to 15.0, and all words had a Thorndike–Lorge (1944) frequency of occurrence of at least 14 per million. To validate the selection of the words, pilot *S*s were asked to sort the 50 words into 10 categories and then to supply a label for each grouping. All *S*s used exactly the same categories as those from which the words had been selected, and provided labels that were identical to or synonymous with the original labels.

The words to be actively classified varied randomly across *S*s. There were 10 separate partitions of the 50 items into the 25 words that were to be categorized and those which were not to be categorized, and one *S* in each of the groups was assigned to each partition. These partitions were such that the members of each taxonomic category were to be categorized by one half of the *S*s in each group, and not by the remaining *S*s.

Procedure. All *S*s were tested individually and sat across from *E*, separated by a low wooden partition. On each trial *E* first read a word and then said either "categorize" or "do not categorize." If instructed to categorize, *S* responded with a number from one to five corresponding to the category of the word. The *S*s were not pretrained with these category identification numbers, and they were told that they should guess on the initial trials. If the correct category number was given, *E* said "right", and when an error was made, *E* said "wrong" and indicated the correct number. While the category structure of the list was not explicitly indicated, *S*s had no difficulty in detecting relations among the items and learning to assign the category numbers. Numbers were not assigned to the 25 words that were not to be categorized; however, in an attempt to equate the amount of time spent per item, *S*s were asked to repeat aloud these words. While this was a self-paced procedure, the duration of word exposure was approximately 4–5 sec. for both categorized and noncategorized items. For each trial the words were presented in a different random order for each pair of *S*s (one in each of the two groups), with the constraint that no 2 words from the same category occurred in succession.

After reaching a criterion of two successive presentations of the list with no classification errors, *S*s wrote down in any order as many of the words as they could remember, taking as much time as they required. They were then given a cued-recall task, (Tulving & Pearlstone, 1966) in which they were presented with a sheet containing the 10 Battig and Montague category labels in a random order, and were asked to write down any additional words that they could recall. All *S*s were then asked several general questions about the experiment and the strategies that they employed.

Subjects. A total of 20 male undergraduates at Princeton University were randomly assigned to the two experimental groups. All *S*s were paid for their participation.

TABLE 1

MEAN NUMBER OF WORDS (TAXONOMICALLY RELATED) CORRECTLY RECALLED IN EXPERIMENT I

Groups	Conditions		Totals
	Classified words	Unclassified words	
Recall	17.80	12.30	15.05
No-Recall	19.50	8.80	14.15
Totals	18.65	10.55	

Results and Discussion

The experimental task proved to be a plausible one. Only two *S*s in the no-recall group indicated slight suspicion about the possibility of a retention test, and the recall performance of these *S*s did not differ from the others in the group. Further, there were no differences between the two instructional groups in the number of trials necessary to learn to classify the items. The recall group averaged 4.1 trials to criterion, while the no-recall group took 4.3 trials.

The data of major importance, the number of categorized and noncategorized words recalled under the two instructional conditions, are indicated in Table 1. The groups did not differ in total amount recalled ($F < 1$), but there was a substantial increase in overall recall for words that were classified, F (1, 18) = 126.58, $p < .01$; and, this effect was significant for both groups, ts (18) $\geq$ 5.39, $p < .01$. However, there was a significant interaction between the effects of instructions to recall and to classify, F (1, 18) = 13.04, $p < .01$. Further analysis indicated that the tendency for the recall group to recall more unclassified words than the no-recall group was significant, $t(36) = 3.33$, $p < .01$, whereas their tendency to recall fewer classified words than the no-recall group was not significant.

These results affirm the considerable importance of categorization in this type of memory task and support the view of the nonadditive nature of recall and organizational instructions. When items are classified, here to a criterion of organizational consistency, further instructions to recall them are of no additional help. However, the fact that the recall *S*s in this experiment remembered more categorized than uncategorized words is not consistent with Mandler's (1967) finding and might question his hypothesis of the equivalence of categorization and recall instructions. On the other hand, this difference does not necessarily refute the position that the processes underlying organization and recall are equivalent. It is possible that the recall group may have tried to classify the complete list, including those words which they were not asked to categorize. Since task conditions would not permit complete classification of the noncategorized items, the differences in the recall of the two classes of words by the recall *S*s may represent the effects of different amounts of organization. In addition, the attempts at classification of the noncategorized items by the recall *S*s might account for their superior recall of these items, when compared with the no-recall *S*s.

This strategy of attempted organization of the noncategorized words would be a feasible one since it is relatively easy to discover the taxonomic categories of the unclassified words. Moreover, when questioned after the experiment, the recall *S*s showed somewhat greater awareness of these categories than did the no-recall group. The cued-recall results also support this interpretation. When presented with the category names, the recall group recalled 22% of the remaining unclassified words, while the no-recall group recalled only 11%, t (18) = 2.23, $p < .05$. In contrast, the increase in the recall of the categorized words was minimal and did not differ between the groups.

These data suggest that instructions to recall are an effective incentive to *S*s to try to organize material, even words that are not explicitly required to be classified. Given the taxonomic materials employed in the present experiment, organization of the noncategorized items was possible to some extent. However, if unrelated words were presented, it would be more difficult for the recall *S*s to impose an organization on those words which were not to be categorized. Under these conditions, the facilitative effects of recall instructions might disappear, and the interac-

tion between instructions to recall and to classify might not be observed. This conjecture was examined in a replication of the first experiment using words which were taxonomically unrelated to each other.

Experiment II

Method

The design, procedure, instructions, etc., of Experiment II were identical to those of Experiment I except for the unrelated words employed as stimuli and the deletion of the cued-recall task. Fifty words, each with a minimum Thorndike–Lorge frequency of 14, were chosen, 1 from each of 50 categories of the Battig and Montague (1969) norms. The 50 words were assigned randomly to 10 arbitrary classes of 5 words each; as before, the classes to be categorized were randomly determined for each pair of *S*s. The *S*s were 20 males from the same population.

Results and Discussion

As was the case in Experiment I, *S*s seemed to accept the experimental situation as a plausible task. Only one no-recall *S* indicated that he thought a retention test might be possible, and his recall did not differ from others in his group. Although a greater number of trials were required to reach the classification criterion than in Experiment I, there was, once again, no significant difference between the two groups; the recall group averaged 9.8 trials to criterion, and the no-recall group averaged 8.3 trials.

Table 2 summarizes the recall results for Experiment II. These results corroborate Experiment I in that there was no significant difference in total recall for the two groups ($F < 1$) and that classified words were better recalled than were unclassified words, F (1, 18) = 79.27, $p < .01$. Where the two experiments differ, however, is that in the present case, as predicted, there was no interaction between instructions to recall and to classify. Thus, the recall group did not show a significant increase in memory for unclassified words, presumably because these words did not allow any known or obvious classification and thus were difficult to organize spontaneously.

These findings fail to replicate those of Mandler (1967) in yet another way. Here an instruction to recall had neither an overall effect, nor a specific effect on words that had not been categorized. To the extent that recall instructions lead *S* to try to organize words, their effect would appear to be limited by task demands. The unrelated words employed here are difficult to organize and their active classification seems to be required for effective recall. Previous results (e.g., Mandler, 1967; Nelson, McRae & Sturges, 1971) probably depend upon the extent to which *S*s in noncategorized conditions were able to impose organization upon the material when instructed to recall it.

TABLE 2

Mean Number of Words (Unrelated) Correctly Recalled in Experiment II

Groups	Conditions		Totals
	Classified words	Unclassified words	
Recall	21.40	12.70	17.05
No-Recall	21.00	10.20	15.60
Totals	21.20	11.45	

The present experiments clearly confirm the importance of organization for successful recall performance. To organize is, to a considerable extent, to remember. Active and consistent categorization is sufficient to yield a relatively high level of recall, and additional instructions for recall do not facilitate performance further. It should be added that it seems unlikely that the differences in the recall of categorized and noncategorized words that were observed in these experiments could be due to differential processing time, in the sense of *S*s devoting a greater proportion of their time to the words being classified. The repetition aloud of the words that were not to be classified probably inhibited rehearsal of the classified words during this time. But the strongest argument against any crucial effect of this kind is the difference in the recall of the nonclassified words between the two experiments. These findings, then, indicate that Mandler was correct concerning categorization. However, organization and recall instructions are not invariably equivalent. Although it seems likely that recall instructions result in attempts at organizational strategies, these instructions fail

if the basis of an organization is difficult to find. It thus appears that if unrelated information is to be remembered, it would be better to force a *S* to organize that information overtly using spatial or class categories, than merely to instruct him to recall.

It is important to indicate that these experiments were not designed to specify how the usage of an organizational strategy facilitates recall performance. However, it should be noted that the retention of classification information is required if an organizational criterion is to be achieved. Indeed, it is possible that organizational strategies are effective because they impose on *S* an implicit and highly structured recall task, namely that of remembering which items are associated with each of a relatively small number of cues or labels. Thus, although instructions to recall a list of words may encourage *S*s to organize the to-be-remembered items, at another level of analysis it seems clear that recall, albeit not of the items per se, is an integral part of the classification process.

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

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

Nelson, A. K., McRae, B. C., & Sturges, P. T. Recall as a function of instructions and trials. *Journal of Experimental Psychology,* 1971, **90**, 152–153.

Thorndike, E. L., & Lorge, I. *The teacher's word-book of 30,000 words.* New York: Columbia University, Teachers College, Bureau of Publications, 1944.

Tulving, E., & Pearlstone, Z. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5**, 381–391.

(Received May 20, 1974)