Effects of semantic and nonsemantic cued orienting tasks on associative clustering in free recall*

ROBERT E. TILL, RANDY L. DIEHL, and JAMES J. JENKINS†

Center for Research in Human Learning, University of Minnesota, Minneapolis, Minnesota 55455

During presentation of a randomized list of nonadjacent associative pairs, Ss heard a cue after each word designating the task to be performed. All Ss performed two tasks. In the identical condition, Ss performed the same task on both members of an associative pair. In the nonidentical condition, Ss never performed the same task on pair members. Semantic tasks led to greater recall than did nonsemantic tasks. Also, percentage of clustering was greater when the second member of a recalled cluster had been used in a semantic task rather than a nonsemantic one. Identical and nonidentical conditions did not differ in the measure of recall. When the two tasks were a combination of one semantic and one nonsemantic task, the identical condition showed a greater percentage of clustering than the nonidentical condition. However, when the two tasks were both semantic, no difference in clustering was obtained.

Hyde and Jenkins (1969) found that when Ss are presented with a randomized list of high-strength associative pairs, recall and clustering depend on the nature of the orienting task performed on the words. The amount of recall and associative clustering were greater for Ss who used words in a semantic task (pleasantness rating) than for Ss who used words in a nonsemantic task (checking for "e"s or estimating the number of letters).

Till and Jenkins (1973) studied the effects of orienting tasks on a list of associated words by using a procedure in which a cue was presented after each word to indicate the task to be performed. Their results constituted a within-S confirmation of the differential effects reported by Hyde and Jenkins (1969): Recall and associative clustering were greater for words used in a semantic task than for words used in a nonsemantic task. Till and Jenkins also found that, when the *same* task was performed on both members of each associative pair, clustering was greater than when *different* tasks were performed on pair members. We refer to this result as the "sameness effect."

The present study was designed to replicate and generalize the results of Till and Jenkins (1973) and, more importantly, to determine whether the sameness effect can be obtained by using *similar* but nonidentical tasks in the "same-task" condition or whether the effect can only be obtained by using identical tasks.

Two semantic and two nonsemantic orienting tasks were used. The two semantic tasks were pleasantness

rating (P) and importance rating (I), and the two nonsemantic tasks were "e" checking (E) and letter estimation (L). Each S performed only two of these tasks, and five of the possible pairwise combinations of tasks were studied: E and P, E and I, L and P, L and I, and P and I. (A sixth possible task combination, E and L, was used in the experiment but is not discussed in the present report. Recall and clustering scores were very low and did not differ on either the task variable or the identicalness variable. No differences were expected, yet the failure to detect any could simply be due to floor effects.) Because separate analyses were to be performed on each pairwise combination, the *task* variable was defined within Ss. For each of the five task combinations, an *identicalness* variable consisting of two conditions was defined between Ss. In the identical condition, the identical task was performed on each member of any given associative pair. In the nonidentical condition, the identical task was *never* performed on both members of an associative pair. In both conditions the two tasks were performed the same number of times (i.e., a given task was either performed on both members of half the pairs or on one member of every pair). Altogether there were 10 groups to study (5 task combinations by 2 identicalness conditions)

First, consider the eight groups which performed a combination of a semantic and a nonsemantic task. In accordance with the results of Till and Jenkins (1973), it was hypothesized that recall and clustering would be greater for words used in a semantic task (P or I words) than for words used in a nonsemantic task (E or L words). (When discussing the effect of the task variable on clustering, we are actually concerned with the task performed on the *second* member of a cluster in recall.) Furthermore, we hypothesized that Ss in the identical condition would have greater clustering than those in the nonidentical condition. On the other hand, as long as the particular measure of recall was chosen to be relatively independent of clustering, there was no strong evidence

*This research was supported in part by the Center for Research in Human Learning at the University of Minnesota. The Center is supported by grants from the National Science Foundation (GB 17590), the National Institute of Child Health and Human Development (HD 01136 and HD 00098), and the Graduate School of the University of Minnesota. Further support was received from the National Science Foundation in the form of a fellowship to the second author.

†Reprint requests should be sent to James J. Jenkins, Center for Research in Human Learning, 205 Elliott Hall, University of Minnesota, Minneapolis, Minnesota 55455.

either for or against the hypothesis that recall would differ as a function of the identicalness variable.

Next, consider the two groups for which the tasks were both semantic. These groups were included to enable us to determine whether the sameness effect is obtainable using similar though nonidentical tasks or whether strictly identical tasks are required. As before, we expected marked associative clustering in the identical condition, but the crucial question was how much clustering would be found in the nonidentical condition. Rating a word for pleasantness is obviously different from rating a word for importance. Yet, at a more abstract level, it is also clear that these tasks are similar. Both require processing of the meaning of the word, and both require the S's evaluation of that meaning on some kind of subjective scale. If strict identity of tasks is necessary in order to obtain the sameness effect, then clustering should be reduced when one member of an associative pair is used in the P task and the other is used in the I task. However, if what counts is merely the abstract similarity between the tasks, then the observed clustering should be the same as in the identical condition. The latter alternative was hypothesized to be correct.

Because of this similarity between the two tasks, we also expected no difference in recall between the identical and nonidentical conditions. In addition, no difference in recall or clustering was expected as a function of the task (P or I). Thus, in sharp contrast to the predictions for the other eight groups, the null hypothesis was predicted in each case.

METHOD

Subjects and Design

All Ss were college students enrolled in an introductory psychology course at the University of Minnesota. There were 42 Ss randomly assigned to each of the five task combinations. For each task combination, 21 Ss performed in the identical condition and another 21 Ss performed in the nonidentical condition.

The data for any group of 21 Ss were obtained by pooling the data from several test sessions with small groups of Ss. Furthermore, the Ss in a single test session did not all perform the same two tasks, although they necessarily were all in the same identicalness condition (since only one taped version of the list could be played during the session).

The Ss who performed the E and P tasks made up Groups Ia and Ib. Group Ia was in the identical condition and Group Ib served in the nonidentical condition. (Similarly, in the groups which follow, a refers to the identical condition and b refers to the nonidentical condition.) Groups IIa and IIb performed the L and P tasks. Groups IIIa and IIIb performed the E and I tasks. Groups IVa and IVb performed the L and I tasks, and Groups Va and Vb did the P and I tasks.

Apparatus

A Wollensak 3000 tape recorder was used to present a list of words to Ss. All Ss were tested in the same group-experiment laboratory. The Ss marked their ratings for the different tasks on answer sheets divided into two sections. The relationship between the cues and the tasks was defined for each group of Ss at the top of the answer sheets. After the answer sheets were collected, a separate sheet of lined paper was given to each S for use during recall.

Materials

The stimulus list consisted of 27 words selected from the Palermo and Jenkins (1964) norms. The first 24 words were a randomized arrangement of 12 pairs of words having high associative strength. The range of intrapair associative strengths from the college-level norms varied from 0.31 to 0.73, with a mean forward associative strength of 0.59 and a mean backward associative strength of 0.47 (Palermo & Jenkins, 1964). Backward associative strength was determined from D.G. Doren's 1968 unpublished norms using Palermo-Jenkins associative norm responses as stimulus words. All associative strengths between words in the list, except the intrapair strengths, were less than 0.05. The last three words in the list served only as fillers to counteract any recency effects. These fillers had less than 0.01 associative strengths with each other or with any other word in the list.

Procedure

The words were presented at the rate of one word per 5 sec. Each word was followed after a .5-sec pause by a cue letter, "A" or "B." Each cue letter represented one of the four tasks: E, L, P, or I. Two versions of the list were prepared. In one version, both members of an associative pair were followed by the *same* cue letter. In the other version, the members of each associative pair were followed by *different* cue letters. A single randomized presentation order was then determined for the words.

During each test session, the tasks were described. On each answer sheet, the cue letters were defined as designating two particular tasks, and each S was instructed to perform on cue the appropriate task of the two assigned to him. Since the relation between the cues and the designated tasks could be permuted, it was possible to counterbalance across Ss any effect due to the interaction of a task with a specific word.

Each task required the selection of one of three categories of response. For the P task Ss chose "pleasant," "unpleasant," or "indifferent;" for the I task they chose "important," "unimportant," or "neutral." For the E task Ss selected "0," "1," or "2 or more" according to the number of "e"s in the spelling of the word. For the L task Ss chose "3," "6," or "9," whichever was closest to the actual number of letters in the word.

Before the list was presented, all Ss were warned about the recall test. Afterward, they were given 4 min to recall as many words as possible, in any order.

RESULTS

For all groups, the dependent variables consisted of a measure of recall and a measure of associative clustering. The recall measure in several past studies has been the *total* number of words recalled; however, this measure obviously varies with the amount of clustering. In order that our measure of recall would be relatively independent of clustering, we decided that only the *first* member of any given associative pair to occur in recall would be counted in the measure. The subsequent occurrence of the other member of the pair was *not* counted as an additional word recalled. This measure is referred to as "adjusted recall." (This measure is equivalent to the number of "opportunities to cluster," terminology used by Till & Jenkins, 1973.) Finally, the measure of clustering was the ratio of the number of clusters to adjusted recall. That is, the clustering percentage was conditional on the appearance of the first pair member in recall, just as in previous research.

The results for the eight groups which performed a combination of a semantic and a nonsemantic task are

reported in Table 1. The means and standard deviations are presented for both adjusted recall and percentage of clustering.

For example, consider Groups Ia and Ib in Table 1. The Ss in Group Ia had a mean adjusted recall of 2.81 E words and 4.28 P words. Similarly, Ss in Group Ib had a mean adjusted recall of 2.48 E words and 5.09 P words. In addition, Group Ia obtained a mean percentage of clustering of 56% when both members of a cluster had been used in the E task (EE clusters) and a mean percentage of clustering of 82% when both members had been used in the P task (PP clusters). Group Ib obtained a mean percentage of clustering of 20% when the first member of a cluster had been used in the P task and the second member in the E task (PE clusters). The Ss in Group Ib also obtained a mean percentage of clustering of 52% when the first member of a cluster had been used in the E task and the second member in the P task (EP clusters).

For each of the pairs of groups, Ia and Ib, IIa and IIb, IIIa and IIIb, and IVa and IVb, the data for the two dependent variables were analyzed independently using a two-way analysis of variance (with repeated measures on one factor). By means of the Hartley test, the assumption of homogeneity of variance was found tenable for Groups Ia, Ib, IIa, IIb, and IVa, but not for the within-S dimension of Groups IIIa, IIIb, and IVb. Accordingly, an approximate F test for heterogeneous variances (Winer, 1962) was used in the latter cases. The between-Ss factor was the identicalness variable (identical vs nonidentical) and the within-S factor was

the nature of the orienting task (semantic vs nonsemantic).

For most of the groups, there were one or two Ss who had no adjusted recall of words which had been used in a particular task (for example, a S might not recall any E words as first members of clusters). The result would be a zero in his clustering ratio denominator. In such cases, the procedure of filling in missing data with column means was adopted. An alternative analysis was performed for the group with the greatest number of instances (five) of missing data. All the data of Ss not having clustering opportunities (i.e., nonzero adjusted recall) under one or both tasks were deleted. This analysis yielded the same pattern of results, with no changes in significance levels.

For each of the four pairs of groups, the analysis of variance indicated significantly greater adjusted recall of words used in a semantic task than words used in a nonsemantic task: for Groups Ia and Ib, $F(1,40) = 41.93$, $p < .001$; for Groups IIa and IIb, $F(1,40) = 38.30$, $p < .001$; for Groups IIIa and IIIb, $F(1,20) = 61.09$, $p < .001$; and for Groups IVa and IVb, $F(1,20) = 89.42$, $p < .001$. No difference in adjusted recall was obtained on the identicalness dimension for any of the pairs of groups. The interaction was nonsignificant for three of the four analyses, but for Groups IVa and IVb the interaction achieved marginal significance [$F(1,20) = 6.11$, $p < .05$].

For each of the four pairs of groups, the percentage of clustering was greater when the second member of a cluster had been used in a semantic task rather than a

Table 1
Means and Standard Deviations for the Semantic-Nonsemantic Combinations

Word Type*	Group Ia Identical		Group Ib Nonidentical		Group IIa Identical		Group IIb Nonidentical	
	E	P	E	P	E	P	E	P
Adjusted Recall								
Mean	2.81	4.28	2.48	5.09	2.33	4.38	2.62	4.48
SD	1.03	1.35	1.40	1.37	1.11	0.86	1.46	1.43
Cluster Type**	EE	PP	PE	EP	LL	PP	PL	LP
Percentage of Clustering								
Mean	56	82	20	52	35	68	28	66
SD	36	30	28	33	35	35	35	37
Word Type	Group IIIa Identical		Group IIIb Nonidentical		Group IVa Identical		Group IVb Nonidentical	
	E	I	E	I	L	I	L	I
Adjusted Recall								
Mean	2.28	4.48	2.14	4.67	2.24	4.52	1.48	5.38
SD	1.10	0.81	0.96	1.74	0.94	1.08	1.21	1.83
Cluster Type	EE	II	IE	EI	LL	II	IL	LI
Percentage of Clustering								
Mean	55	80	22	58	36	69	20	48
SD	39	18	24	44	41	26	20	41

*This refers to the task performed on the recalled word: "e" checking (E), letter estimation (L), pleasantness rating (P), or importance rating (I).

**This refers to the tasks performed on each word of a recalled cluster.

nonsemantic task: for Groups Ia and Ib, $F(1,40) = 25.57$, $p < .001$; for Groups IIa and IIb, $F(1,40) = 33.62$, $p < .001$; for Groups IIIa and IIIb, $F(1,20) = 18.18$, $p < .001$; and for Groups IVa and IVb, $F(1,20) = 16.33$, $p < .001$. For three of the four pairs of groups, a significantly greater amount of clustering was found in the identical condition than in the nonidentical condition: for Groups Ia and Ib, $F(1,40) = 17.61$, $p < .001$; for Groups IIIa and IIIb, $F(1,40) = 14.64$, $p < .001$; and for Groups IVa and IVb, $F(1,40) = 7.31$, $p < .01$. For Groups IIa and IIb, no difference in percentage of clustering was obtained. The interaction was nonsignificant for each of the pairs of groups.

The results for the two groups which performed a combination of two semantic tasks are reported in Table 2. The means and standard deviations are presented for adjusted recall and percentage of clustering. The table is constructed in the same way as Table 1.

As before, the data were independently analyzed for the two dependent variables. The between-Ss factor was the identicalness variable, and the within-Ss factor was the task variable (P vs I). No significant differences were found in adjusted recall as a function of the identicalness variable or the task variable. Similarly, no differences in percentage of clustering were indicated for either variable. Interactions were nonsignificant.

After computing the proportion of Ss who recalled a given word used in a given task, it was possible to investigate the correlations, over items, between tasks. These correlations, which range from 0.74 to 0.85, are attributable to the fact that the same individual words in the same serial order were being compared. The fact that the correlations remained so high across all the comparisons suggests that the tasks did not interact strongly with specific words.

DISCUSSION

The present study replicates and extends the results of Till and Jenkins (1973). For the groups which performed a combination of a semantic and a

nonsemantic task, words used in the semantic task were better recalled than words used in the nonsemantic task. Tasks which required the processing of the meaning of words facilitated recall, while tasks that focused on the orthographic aspects of words led to poor recall. Also, associative clustering was greater when the second member of a cluster had been used in a semantic task rather than a nonsemantic one. That is, associative clustering is not "automatic," but depends on the nature of the task activity. Given the occurrence of one member of an associative pair in recall, the recall of the other member is more likely if it was used in a semantic task.

For each combination of a semantic and a nonsemantic task, the identical and nonidentical conditions did not differ in amount of adjusted recall. However, the two conditions did differ in amount of associative clustering: It was always greater for the identical condition and usually significantly so. Since the adjusted recall measure is equivalent to the number of opportunities to cluster, it can be seen that the clustering difference between the identical and nonidentical conditions is not biased by a difference in clustering opportunities. Rather, this difference in clustering is attributable to the degree of similarity between the tasks used on associative pair members.

The above findings indicate that the nature of the orienting tasks and the manner in which they are assigned to associative pair members are crucially important variables for storage and retrieval. Still, the notion of task similarity needed clarification.

The groups which performed a combination of two semantic tasks were essential for the investigation of task similarity. No difference in adjusted recall was found between the two tasks (P and I). Similarly, no difference in associative clustering occurred between clusters whose second member was an I word and clusters whose second member was a P word. Also, for these groups, no difference in adjusted recall was found between the identical and nonidentical conditions. Most importantly, associative clustering did *not* differ between the identical and nonidentical conditions.

This last result suggests that the "sameness effect" should be redescribed. Strict task identity appears to be unnecessary to obtain the greater clustering which Till and Jenkins (1973) observed when the *same* task was performed on both members of an associative pair. The high degree of associative clustering in the nonidentical condition (with IP and PI clusters) probably reflects the fact that these two tasks are abstractly similar. Both tasks require the elaboration and subjective evaluation of word meaning. In short, the sameness effect should be obtainable when this kind of similarity exists between the two tasks.

REFERENCES

- Hyde, T. S., & Jenkins, J. J. Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, 1969, 82, 472-481.

Table 2
Means and Standard Deviations for the
Semantic-Semantic Combinations

Word Type*	Group Va Identical		Group Vb Nonidentical	
	I	P	I	P
Adjusted Recall				
Mean	3.90	4.00	3.05	3.90
SD	0.99	1.00	1.74	1.64
Cluster Type*	II	PP	II	PP
Percentage of Clustering				
Mean	76	75	67	75
SD	20	27	34	30

*See footnotes to Table 1.

- Palermo, D. S., & Jenkins, J. J. *Word association norms: Grade school through college*. Minneapolis: University of Minnesota Press, 1964.
- Till, R. E., & Jenkins, J. J. The effects of cued orienting tasks on the free recall of words. *Journal of Verbal Learning & Verbal Behavior*, 1973, 12, 489-498.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received for publication December 27, 1973;
revision received April 19, 1974.)