Department of Psychology

College of Liberal Arts

The Pennsylvania State University

University Park

Charles N. Cofer and Richard A. Olsen

Anticipation and Recall Methods in Paired Associate Learning: Effects of Rate of Item Presentation, Item Type, and Intra- and Inter-Item Similarity

Technical Report No. 6 for Contract Nonr 656(30)
(Learning, Retention and Recovery of Meaningful Materials)

between

The Pennsylvania State University and the Office of Naval Research

July, 1965

Anticipation and Recall Methods in Paired Associate Learning: Effects of Rate of Item Presentation, Item Type, and Intra- and Inter-Item Similarity

Charles N. Cofer and Richard A. Olsen
The Pennsylvania State University

Recent experimentation in verbal learning has made extensive use of the paired-associate learning procedure. Much of this work has employed the anticipation method, in which, on presentation of the stimulus member of the pair, the S is to attempt to spell or pronounce the response member before it comes into view. For any given pair, the sequence of events in time in this procedure is stimulus and then the pair. This anticipation procedure has often been said to have the advantage of the presentation of immediate knowledge of results and of reinforcement (when the anticipation is a correct one) or of immediate correction of erroneous anticipations.

For reasons arising from concern with other problems, Battig & Brackett (1961) compared the efficacy of the anticipation procedure with another procedure which they called the recall method (other investigators refer to it as the study-test method). In this technique, each trial is divided into two parts. In part 1, the pairs are presented one after another and, then, in part 2, the stimuli are presented (usually in an order different from that of the pairs) and S is asked to recall or reproduce the response which is appropriate to each stimulus. (The sequence here is P, P, P...S, S, S...). No knowledge of results, reinforcement or correction is involved in this procedure. The differences just summarized would make one expect more rapid learning for the anticipation than for the recall procedure, but Battig & Brackett reported the opposite result. They used pairs composed of nonsense-shapes as stimuli and of 2-digit numbers as responses. Their Ss learned the 12-pair lists to criterion (one errorless trial) in 6.85 trials under the recall method and 9.75 trials under anticipation; under recall a mean of 31.83 errors and under anticipation 49.58 errors were made.

Other results reported in the literature have not always been confirmatory of the superiority of recall to anticipation. Lockhead (1962) used nonsense syllables (CVC's) of intermediate association value and found learning under anticipation to be equal in difficulty to that under the recall procedure. Battig, Brown & Nelson (1963) also found no differences between the methods for pairs composed of CVC's (Exp. I, p. 699) but did find an over-all superiority for recall (Exp. IV, p. 708) over 5 lists. These lists were composed of trigrams, some of which were actual words (W), some non-words (N). The pairs used over four of the lists may be symbolized as N-N, N-W, W-N, and W-W, and there was a mixed list composed of 5 N-N and 5 W-W pairs. From the report, it is not possible to discern whether the advantage of the recall method occurred across all lists or was specific to one or two of them (the total advantage for recall was slight: 31.4 to 36.3 errors).

Battig & Brackett (1963) have compared the two methods with CVC pairs varying in inter-item similarity and have also studied performance under the methods as a function of per cent occurrence of the response member (% ORM) during acquisition.

Over-all, they found no consistent or significant superiority for the recall method or any consistent advantage for either method as a function of inter-item similarity. They did find a significant interaction between the methods and % ORM which showed the recall method to be superior under 100% ORM and the anticipation method to be superior under 50% ORM. However, this interaction arose primarily from the depressed performance of one recall group (at 50% ORM) and may not be a replicable finding. Battig & Brackett (1963, p. 513) suggest that kind of material (digits vs. trigrams) may be an important factor determining whether recall will be superior to anticipation or whether there will be no differences. Schild & Battig (in press), using CVC pairs, however, in a two-list design involving directionality of associations formed during acquisition of list 1, did find that fewer trials were required for list 2 learning under recall than under anticipation. Some of these differences may have arisen because of a pretest procedure, which resembled recall, but it is unlikely that all of the methods differences can be accounted for on this basis.

The evidence summarized suggests that when 2-digit numbers are used as responses (and perhaps 3-letter words) the recall method results in more rapid learning than the anticipation method. Except for the Schild & Battig experiment,[3] with CVC's, however, the methods do not produce differences, even with variations in inter-item similarity. Two presentation rates have been used with CVC's (4" - 4" and 2" - 2"), and in the Battig experiments the responses were always pronounced (it is not clear from Lockhead's report whether the responses were pronounced or spelled). We can concluded from the literature that anticipation is not superior to recall, in contrast to expectation as indicated above, and is sometimes inferior to recall.

The experiments reported here were undertaken in an effort to replicate the foregoing findings and to extend somewhat the conditions under which the methods have been compared. In the first experiment, 3-letter word - 2-digit number pairs were used to avoid the problem of stimulus discriminability which may have been involved with the nonsense shapes used by Battig and Brackett (1961). In addition, 3 exposure intervals were employed. In the second and third experiments, CVC trigram pairs were used and variation was introduced in terms of intra-pair (or S-R) similarity in Experiment 2 and in terms of inter-stimulus and inter-response similarity in Experiment 3.

Experiment 1[1]

Battig & Brackett (1961) used a 5-5 rate in their study of the two methods, i.e., in anticipation the stimulus (nonsense shape) was presented for 5" and then the pair (shape-2 digit number) for 5" and in recall the pair was presented for 5" in the first part of the trial and the stimulus for 5" in the second part of the trial. In addition to using words rather than shapes as stimuli, we used in this experiment 1-1 and 3-3 rates as well as the 5-5 rate.

Materials. Twelve-item lists were constructed, the stimuli of which were 3-letter words with frequencies of occurrence (Thorndike & Lorge, 1944) of AA and the responses of which were 2-digit numbers selected so as to include none ending with "5" or with "0". The list is shown in Table 1. Three random orderings of the list were prepared, to avoid serial learning.

Subjects. There were 48 male and 48 female students (naive with respect to paired-associate learning) enrolled in the class in introductory psychology at the

Pennsylvania State University for which extra credits can be earned by service in experiments. Assignment to the six groups was unsystematic, except that the sex ratio was kept constant for each group.

Procedure: The materials were presented by means of a Stowe Memory Drum. For both anticipation and recall groups there was first a study trial, in which the pairs were presented and S pronounced the response. After a 30" inter-trial interval, in one set of groups anticipation trials were begun, with the stimulus and the pair presented, in different groups, for 1", 3" or 5". Thirty seconds after the study trial with the recall groups, the stimulus items were presented and S attempted to recall the responses; the order in which the stimulus items were presented always differed from that in which the pairs were presented. There was another thirty second interval between the presentation of the stimulus terms and the next presentation of the pairs. This arrangement, of a thirty second inter-trial and a thirty second intra-trial interval for the recall method, replicates the procedure of Battig and Brackett (1961, p. 61; see also, p. 64).

Results. Total trials to criterion are reported in Table 2 separately for the eight male and eight female Ss in each group for the two methods and three exposure intervals. The analysis of variance shows significance for methods [$F(1,84) = 8.30$, $p < .01$] and for intervals [$F(2,84) = 28.85$, $p < .01$] but none of the interactions was significant. It is quite clear from Table 2 that fewer trials were required to reach criterion under the recall than under the anticipation method at every interval and that, as would be expected, fewer trials are required under either method as exposure intervals increase. Similar findings emerged for other data analyses based on errors and correct responses and will not be reported here, as they contribute no additional information.

We may conclude from Experiment 1 that the findings of Battig & Brackett have been replicated, that they are maintained at shorter exposure intervals than the one used by Battig & Brackett, and that the change from nonsense shapes to meaningful work stimuli has not affected the differences.

Because sex differences will appear later in this report, we point to one aspect of Table 2 not reflected significantly in any of the interactions. This is that females require fewer trials to criterion than males at the two short exposure intervals under recall, require many more trials than males at the shortest interval under anticipation, and fewer than males at the longest interval under anticipation.

Experiment 2

In this experiment, pairs of CVC's were employed and intra-pair similarity was manipulated under the two methods.

Materials. Four lists of 10 CVC's each were taken from Hull's list (cf. Stevens, 1951, p. 545) as having less than 20% association value. One list was used for the stimulus CVC's in this experiment, and the others were used to make pairs at one of three levels of intra-pair similarity. The CVC's are listed in Table 1, where the stimulus list is referred to as "common stimuli" and the three response lists are referred to as "high similarity", "medium similarity" or "low similarity" (i.e., zero). A given S learned a list with stimuli from the common stimulus list paired with responses from one of the response lists (high, medium, or low similarity). Inspection of Table 1 reveals that for high S-R

(intra-pair) similarity two letters were shared by each stimulus and its response; for medium similarity, one letter and for low similarity no letters were shared by each stimulus and its response. All the vowels were used in the stimulus list (2 A's, 2 E's, 4 I's, 1 O and 1 U) and in the response lists (2 occurrences of each vowel).

In the stimulus list, eight consonants were used for the first letter (V and Y were each repeated once), and six consonants were used for the last letter (F and K were each repeated once, and J twice). Nine different consonants were used as the first letter in the high similarity response list (P repeated once), eight in the medium similarity list (B and K each repeated once), and eight in the low similarity list (J and Z each repeated once). For the third letter in these lists, six consonants were used for high similarity (F and B repeated once, J two times), six for medium similarity (F and J each repeated twice), and eight for low similarity (D and T each repeated once). The response lists differ somewhat in inter-item similarity (as well as in intra-pair similarity), chiefly in terms of the third letter, a fact made almost inevitable by the other constraints operating.

Subjects. Sixty male and six female Ss, similar to those used in Experiment 1, were employed. They were assigned to conditions as they appeared at the laboratory, with the restriction that the sex ratio in each condition was maintained at the same level. Four additional Ss were discarded, three because of experimenter error and one for failure to learn at all.

Procedure. Instructions for P-A learning (either method) were given by playing a tape recorder, and S then learned three PA pairs (ZAB-DEF, HIJ-NOP, TUV-ZOB) to assure understanding of instructions. S was encouraged to guess and was required to spell the CVC stimulus and response terms on each occasion of their appearance and to anticipate or recall the CVC response by spelling it. A 4-4 rate was used, with a thirty-second inter-trial interval (for recall only four seconds intervened between presentation of the pairs and presentation of the stimuli, i.e., the intra-trial interval was four rather than thirty seconds as in Experiment 1).

The experiment, then, consists of three levels of intra-pair similarity, the two methods, and the two sexes, and 10 Ss were run in each cell, there being 120 in all. Learning was to criterion (one perfect trial) or through twenty trials, whichever came first.

Results. Table 3 shows the total and mean numbers of correct responses achieved over twenty trials by the various groups (where criterion was reached before twenty trials the S was given credit for perfect performance thereafter[2]). Inspection of the totals and the means reveals that more correct responses occur under recall than under anticipation at each similarity level but that the only substantial difference is for the medium similarity case. In the analysis of variance both levels of similarity and methods yield significant values [$F(2,108) = 95.32$, $p < .01$] for similarity and [$F(1,108) = 4.40$, $p < .05$] for methods, but the interaction of these factors is not significant (F is less than 1). The significant methods difference presumably arises mainly because of the medium similarity comparison.

Further inspection of Table 3 reveals evidence suggestive of sex differences. Thus, females perform somewhat less well than males at low and high similarity levels under anticipation and better than males at medium similarity. Under recall

females perform better than males at medium and high similarity and are roughly comparable at low similarity. As a main effect, sex does not reach significance [$F(1,108) = 3.33$, $p < .10$], but the sex by similarity level interaction is significant [$F(2,108) = 4.92$, $p < .01$]. However, the sex by methods interaction just fails to reach significance [$F(1,108) = 3.83$, with 3.95 required at $p = .05$]. (The triple interaction of sex by similarity levels by methods yields an F of less than 1.)

There is no apparent reason for the sex differences which appear, roughly, to show that females do consistently better under recall than anticipation, whereas males are inconsistent, sometime showing no difference (LS), sometimes favoring recall (MS), and sometimes favoring anticipation (HS). The differences in methods for males are smaller than for females. For total correct responses the difference is two responses for LS, 119 for MS, and 96 for HS, whereas for females the corresponding differences are 101, 268, and 240.

As to the over-all methods differences, the small separation of the two methods for the LS list is consistent with prior work with trigrams. The similarly small difference for the HS list may arise from the fact that it is an easy list. The advantage for recall with the MS list resembles results found with two-digit responses. Perhaps the partial identity (one letter) of the S and R terms in each pair in this case provides for ready response integration, parallel to that in the case of digits.

Experiment 3

In this experiment, the two methods were compared for a list having inter-stimulus similarity and for a list having inter-response similarity.

Materials. Two eight-item lists were made up from CVC's with less than 50% association value (Stevens, 1951, p. 545) and are shown in Table 1. List I items were composed of only the consonants J, K, Q, V (plus all the vowels) and is thus high in inter-item similarity. In List II, 15 consonants were used (H appearing twice), and J, K, Q, V (and B) were excluded; all the vowels are represented in this list. Inter-item similarity is minimized in List II. When high inter-stimulus similarity was involved, List I was used for the stimulus terms and List II for the response terms; for the reverse case of inter-response similarity, List II was used for the stimulus terms and List I for the response terms.

Subjects. Forty male and forty female Ss, comparable to those used in Experiments 1 and 2, were employed and were assigned to conditions roughly in order of appearance at the laboratory.

Procedure. Procedure was identical to that employed in Experiment 2.

Results. Table 4 shows the data for total and mean number of correct responses obtained in Experiment 3. Analysis of variance reveals that the only significant main effect is for sex [$F(1,72) = 5.41$, $p < .05$], the effects of similarity and methods producing F's of less than one. The only significant interaction is between sex and stimulus and response similarity [$F(1,72) = 7.57$, $p < .01$]. As Table 4 indicates, females performed much better than males under conditions of stimulus similarity for both methods and performed about equally to males under response similarity. The performance of females under recall is superior to that under

anticipation (slightly for stimulus similarity, more so for response similarity), whereas males do better under anticipation for stimulus similarity and for recall under response similarity.

Discussion

The major findings with respect to the recall and anticipation methods observed in this series of experiments may be summarized as follows: with word-number pairs, recall produced more rapid learning than anticipation under three presentation intervals and for both sexes (over-all); with CVC pairs varying in intra-pair similarity, recall was superior to anticipation over-all, but the chief locus of advantage (more correct responses) was for the medium similarity condition; for CVC pairs varying either in stimulus or response similarity, females performed somewhat better under recall than anticipation, but males were inconsistent (and the sex by method interaction was not significant and neither was the methods main effect).

In general, these findings are consistent with those in the literature which have found recall to be superior to anticipation under certain conditions and have found no differences under other conditions. Sex differences have not been reported in the prior studies, and a close review of the relevant papers indicates that in no case was the sex composition of the groups used reported. In order to have a clearer picture of the sex differences found here, we have tabulated learning scores across all conditions by sex for all three experiments. Table 5 shows the results. It is clear that, for females, the recall procedure generates more rapid learning than the anticipation procedure over all experiments and that, for males, the methods differences are smaller and occasionally reversed.

After the results of Experiment 3 were obtained, an effort was made to find in various characteristics of the groups of male and female Ss attributes correlated with learning performance. The available data, obtained from University files, were scores on Aptitude tests, English and Mathematics Placement tests and their sub-parts. None of these available scores showed relationships to the learning measures obtained in either sex group or differentiated the sex groups successfully.

The findings of materials and sex differences related to performance by the anticipation and recall methods seriously limits the generality of the conclusion that recall produces faster P-A learning than anticipation. However, anticipation is seldom superior to recall. It is possible that where recall is superior to anticipation, response integration occurs faster than when it is not. On the other hand, the separation of stimulus presentations from pair presentation or the separation of the learning task from the recall task in the recall method, where the recall method is superior, may be the important factors. Further experiments are being conducted in this laboratory to get at these problems. It appears now to be necessary, in conducting experiments on these methods, to keep the data on the sexes separate and to look for different effects of variables on males and females.

Summary

Three experiments comparing recall and anticipation methods in paired associate learning are reported. In the first, word-number pairs were used with three

exposure intervals, and learning was superior under recall. In the second, CVC trigram pairs were employed, with three levels of intra-pair (S-R) similarity. Recall was again superior, over-all, the major difference appearing at moderate similarity. Sex differences appear. In the third experiment, inter-stimulus or inter-response similarity was varied, with CVC trigram pairs. The major effect here was from sex differences. Across all experiments, females tend to perform better on recall than on anticipation, whereas males tend to be inconsistent and to show smaller differences when they occur.

References

Battig, W. F., & Brackett, H. R. Comparison of anticipation and recall methods in paired-associate learning. Psychol. Rep., 1961, 9, 59-65.

Battig, W. F., & Brackett, H. R. Transfer from verbal discrimination to paired-associate learning: II. Effects of intralist similarity, method and percentage occurrence of response members. J. exp. Psychol., 1963, 65, 507-514.

Battig, W. F., Brown, S. C., & Nelson, D. Constant vs. serial order in paired-associate learning: Psychol. Rep., 1963, 12, 695-721.

Battig, W. F., & Wu, R. D. Comparison of recall and anticipation paired-associate procedures within mixed aurally-presented lists. Psychonomic Science, (1965, in press).

Lockhead, G. R. Methods of presenting paired associates. J. verbal Learn. verbal Behav., 1962, 1, 62-65.

Schild, M. E., & Battig, W. F. Directionality in paired-associate learning. J. verbal Learn. verbal Behav., (in press).

Stevens, S. S. (Ed.) Handbook of experimental Psychology. New York: Wiley, 1951.

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

Footnotes

[1]Experiment 1 was carried out by Miss Florence Diamond for a senior thesis in 1964.

[2]In several cases, learning was taken one trial beyond criterion, and no S made an error on the additional trial.

[3]In a paper which came to our attention after this manuscript was written, Battig & Wu (1965) have found recall superior to anticipation for word-bigram pairs. Their experiment involved a mixed list, in which for half the pairs the anticipation method and for the other half the recall method were used. Over all trials to criterion, there were 44.8 errors on the recall pairs and 49.5 errors on the anticipation pairs.

Table 1

Lists Used in the Three Experiments

Experiment 1

GET	12
DID	29
RED	19
OUR	52
MAY	63
SAW	69
PUT	51
HOW	17
LET	56
CAR	28
ACT	22
USE	64

Experiment 3

LIST I	LIST II
KAQ	YUN
VIJ	FOH
QOJ	XAD
QUK	TAH
KIV	MEP
JUQ	RIW
VEK	SOZ
JEV	GIC

Experiment 2

Common Stimuli	Responses		
	High Similarity	Medium Similarity	Low Similarity
DAJ	FAJ	KUJ	KEB
YAF	MAF	VOF	MIV
NOJ	NIJ	NAF	VUD
PIB	POB	PEJ	ZOT
VEF	PEF	KIF	BUP
BIW	BUW	BUP	JEX
KEX	KEB	BEW	ZIH
VUK	VUD	ZIK	JAT
TIJ	TOJ	HAJ	YOD
YIK	YIN	YOB	HAJ

Table 2

Total and Mean Trials to Criterion in Experiment 1 for the Sexes, Intervals and Methods

	Interval		
Anticipation	**1-1**	**3-3**	**5-5**
Male - total	188	121	98
Female - total	336	119	49
Total	524	240	147
Mean	32.75	15.00	9.18
Recall			
Male - total	192	73	45
Female - total	177	56	46
Total	369	129	91
Mean	23.06	8.06	5.69

Table 3

Total and Mean Number of Correct Responses Achieved in 20 Trials in Experiment 2. High Similarity Designated by HS, Medium Similarity MS, and Low Similarity by LS.

	LS	MS	HS
Anticipation			
Male	759	590	1521
Female	638	820	1392
Total	1397	1410	2913
Mean	69.85	70.50	145.65
Recall			
Male	757	709	1425
Female	739	1088	1632
Total	1496	1797	3057
Mean	74.80	89.85	152.85

Table 4

Total and Mean Number of Correct Responses in Experiment III. The Column marked Stimulus refers to the list inter-stimulus similarity and the one marked Response to the list with inter-response similarity.

	Stimulus	Response
Anticipation		
Male	648	619
Female	821	608
Total	1469	1227
Mean	73.45	61.35
Recall		
Male	462	743
Female	888	719
Total	1350	1462
Mean	67.50	73.10

Table 5

Sums of Trials (Experiment 1) and Correct Responses (Experiments 2 and 3) Across All Conditions in Each Experiment by Sex for the Anticipation$^{(A)}$ and Recall Methods$^{(R)}$

	Males		Females	
	A	R	A	R
Experiment 1 (trials)	407	310	504	279
Experiment 2 (correct responses)	2870	2891	2850	3459
Experiment 3 (correct responses)	1267	1205	1429	1607