Paired associate learning predicted from indices of free association strength¹

S. JOYCE BROTSKY, *San Fernando Valley State College, Northridge, Calif. 91324* and MARIGOLD L. LINTON, *San Diego State College, San Diego, Calif. 92115*

Following a word-association test (WAT), 88 Ss studied a list of paired associates composed of 68 stimulus words (Experiment 1) or 60 stimulus words (Experiment 2) and their primary normative responses (R1). The tendency to give R1s in the test was a function of the preexperimental free-association strength (FAS) between R1 and its eliciting stimulus, with significantly more R1s given to stimuli in the high-FAS group than to stimuli in the low-FAS group. The amount of facilitation produced by training varied inversely with initial FAS. The results were attributed to the elicitation and reinforcement of a mediational rule which was most effective if R1 was a strong normative primary.

If free-association norms provide valid estimates of strength of language habits, the probability of occurrence of any response to a given normative stimulus (FAS) should provide the basis for differential predictions of paired-associate learning (PAL). Wicklund, Palermo, & Jenkins (1964) have confirmed this hypothesis with children, but the PAL of adults has not appeared sensitive to variations in FAS between stimulus and normative primary. Rather than finding that facilitation varied as a function of FAS, Jenkins and his associates (Jenkins, 1963) found that all word pairs with nonzero FAS showed significantly more facilitation after training than the word pairs used as controls.

Following a strategy suggested by Jenkins (1963), an attempt was made to provide optimal conditions for Ss to form, and subsequently to apply, a mediational rule or response set that would lead to a high proportion of correct answers following a minimum of training. In order to produce minimal interference with Ss' response sets, a modified recall PAL procedure was developed. The recall PAL method provides for the presentation of the stimulus-response (S-R1) pairs, one at a time, without interspersed test trials. After the study trial, a test consisting entirely of stimulus items is presented, and S responds without feedback (Battig & Brackett, 1961). A procedural variation in the present research was the presentation of a word-association test (WAT) prior to the study trials. The WAT was employed to obtain a direct measure of changes in

response probabilities in addition to the inferences of changes in response probabilities that are possible when free-association norms are used. A second variation was the use of two training trials rather than a single training trial prior to the presentation of the test stimuli.

A second experiment was conducted in an attempt to replicate the results of Experiment 1, and to explore the effect of training trials on the frequency of occurrence of R1 in the test.

GENERAL METHOD

Stimulus Materials

The stimulus words were chosen from the Connecticut Free Associational Norms (Bousfield, Cohen, Whitmarsh, & Kincaid, 1961). In Experiment 1, R1 was the normative primary as defined by the Connecticut Norms, while in Experiment 2, R1 was the normative primary as defined by the San Fernando Valley State College (SFVSC) Memory Drum Norms (Brotsky & Linton, 1968). The stimulus words were typed in upper-case letters on vellum tape and presented to S via a Stowe memory drum.

Procedure

In both experiments the three phases followed one another, interrupted only by brief instructions. In the first phase (WAT), the stimulus words were presented serially on the memory drum, and Ss were given standard free-association instructions which stressed speed of response. An electric timer was activated by the rotation of the drum as each stimulus appeared and stopped when E tapped a telegraph key at the onset of S's response.²

The second phase consisted of one or two study trials during which S was asked to study, as pairs, the stimulus words that had been presented in the WAT and their R1s, reading the response member of the pair aloud as it appeared in the drum window. After the word pairs had been presented once, those Ss who were to have a second study trial were told, "Let's go through the list once more to be sure that you have the pairs clearly in mind before we go on to the test list; remember to pay close attention to the word you say aloud, because you will be asked to recall it later." The Ss who were given only one study trial went on to the third phase immediately. In the third phase, the stimulus words were presented serially in a different order than had been used in training, and S was instructed to respond to each of the words as quickly as possible with

the word that had been paired with it during the training phase.

Subjects

The Ss were undergraduates between the ages of 17 and 23 who had been born in the continental United States and had learned English as their first language. Half of the Ss in each group were males and the other half were females. Participation in the experiment fulfilled a course requirement at SFVSC.

EXPERIMENT 1

Stimulus Materials

A distribution was made of the normative frequency of the most popular response (R1) to each of the 400 words in the Connecticut Norms. The distribution approached normality with a range of 8% to 91%, the greatest concentration of scores occurring between 20% and 25%. A frequency distribution was also made for the second most popular response (R2). This distribution, though more leptokurtic than the R1 distribution, approached normality with a range of 1% to 40%, the greatest concentration of scores falling between 9% and 11%. Each distribution was divided into three parts: high, medium, and low. Four groups of stimulus words were constructed: Group 1 where R1 was high and R2 was low, Group 2 where both R1 and R2 were high, Group 3 where R1 was low and R2 was high, and Group 4 where both R1 and R2 were low. Sixteen words were chosen from each of the four groups by a method of restricted randomization. Four filler words were added whose R1 and R2 were from the medium part of their respective distributions. When the normative probabilities of occurrence of R1 and R2 were plotted for the 68 words in the sample, the range was somewhat restricted for the two bimodal distributions. Two randomizations of the list were developed.

Procedure

The 68 stimulus words in the WAT and the test were presented to S at a 6-sec rate, although responses were accepted for scoring only if they occurred within the first 5 sec following stimulus presentation. All Ss were given two study trials during the training phase, with the 68 word pairs presented at a 4-sec rate.

Subjects

The Ss were 24 males and 24 females.

Results

WAT data. The mean probability of occurrence of R1 was .72 for Group 1, .57 for Group 2, .23 for Group 3, and .11 for Group 4. An alternative way of organizing the WAT data was to divide the 68 stimulus words into four groups of 17 words each on the basis of the simple frequency of occurrence of R1 in the Connecticut Norms. The mean probability of occurrence of R1 was .68 for Group A (high FAS), .54 for

Group B (med-high FAS), .27 for Group C (med-low FAS), and .11 for Group D (low FAS). The 68 words used in the experiment were also organized into four groups of 17 words each on the basis of the frequency of occurrence of R1 in the WAT itself. The mean probability of occurrence of R1 was .73 for Group A' (high FAS), .51 for Group B' (med-high FAS), .27 for Group C' (med-low FAS), and .10 for Group D' (low FAS). Females gave 40.9% R1s in the WAT, while males gave 38.7%, a difference that was not statistically reliable.

Test data. Training produced an increase in the probability of occurrence of R1 to every stimulus word. The mean probability of occurrence of R1 in the test was .98 for Group 1, .97 for Group 2, .88 for Group 3, and .85 for Group 4. An analysis of variance indicated that the main effect for groups was reliable [$F(3,60) = 8.74, p < .05$]. The Duncan Multiple Range Test (Edwards, 1960) used to compare the means of the four groups indicated that neither the difference between Group 1 and Group 2 nor the difference between Group 3 and Group 4 was statistically reliable, although all other differences were statistically significant ($\alpha = .05, df = 60$).

When the 68 words were divided into four groups on the basis of the frequency of occurrence of R1 in the Connecticut Norms, the mean probability of occurrence of R1 in the test was .98 for Group A, .95 for Group B, .92 for Group C, and .82 for Group D. The main effect for groups was reliable [$F(3,64) = 11.76, p < .05$]. The Duncan Range Test indicated that the mean number of R1s in Group B did not differ reliably from the mean number of R1s in Group A or Group C. All other differences were statistically significant ($\alpha = .05, df = 64$).

When the 68 words were divided into four groups of 17 words each on the basis of the simple frequency of occurrence of R1 in the WAT itself, the probability of occurrence of R1 in the test was .98 for Group A', .97 for Group B', .92 for Group C', and .81 for Group D'. The main effect for groups was reliable [$F(3,64) = 15.62, p < .05$]. The Duncan Range Test indicated that Group B' did not differ reliably from Group A' or Group C', although all other comparisons were statistically significant ($\alpha = .05, df = 64$).

Female Ss gave significantly more R1s in the test than male Ss (93.8% vs 90.1%) [$F(1,46) = 5.17, p < .05$].

The relationship between the frequency of R1 to each of the 68 words in the test and the frequency of R1 in the Connecticut Norms was statistically reliable ($r = .53, df = 66, p < .01$). The amount of facilitation produced by training varied inversely with the initial FAS of R1, since

training produced substantial facilitation and a consequent clustering of the groups toward the ceiling of 100% production of R1s in the test. The same conclusions were reached on the basis of the relationship between the frequency of R1 to each of the 68 words in the test and the frequency of R1 in the WAT ($r = .60, df = 66, p < .01$).

EXPERIMENT 2 Stimulus Materials

Eight words were deleted from the set of 68 words used in Experiment 1 in order to discard homonyms (2), words which elicited no R1s in the WAT of Experiment 1 (3), and to equate the size of word groups (3). The R1 was defined as that response which was given most frequently in a WAT administered to the 48 Ss who participated in Experiment 1 and to an additional 61 Ss who provided the data for the SFVSC Memory Drum Norms. The 60 stimulus words were divided into three groups of 20 words (high FAS, moderate FAS, and low FAS) on the basis of the normative frequency of R1 in the drum norms.

Procedure

The Ss were randomly assigned to two groups. Group 1 was given a single study trial, while Group 2 was given two study trials in the training phase. All stimuli were presented at an 8-sec rate, and Ss were given the full 8 sec to respond.

Subjects

The Ss were 20 male and 20 female undergraduates.

Table 1
Mean Proportion of R1s in the WAT and the Test for High, Moderate, and Low FAS Word Groups for the 20 Ss in Group 1 and the 20 Ss in Group 2

FAS	Group 1		Group 2	
	WAT	Test	WAT	Test
High	.74	.98	.67	.98
Moderate	.42	.94	.30	.96
Low	.20	.87	.17	.90
Mean	.46	.93	.38	.95

Results

WAT data. The mean probability of occurrence of R1 was .71 for the high-FAS words, .36 for the mod-FAS words, and .19 for the low-FAS words. The mean probability of occurrence of R1 in both WAT and test for the three word groups is presented in Table 1 for the Ss in Groups 1 and 2. Although 10 male and 10 female Ss had been randomly assigned to Group 1 and to Group 2, the stimuli elicited more R1s prior to training from Ss in Group 1 than from Ss in Group 2 (45.6% vs 38.1%). A 2 (groups) by 2 (sex) analysis of variance indicated that this difference was significant [$F(1,36) = 9.60, p < .05$]. Neither the difference between males and females

(40.2% vs 43.4%) nor the interaction between group and sex was statistically reliable.

Test data. The probability of occurrence of R1 to every stimulus word increased following training. The data presented in Table 1 indicate that there was a monotonic increase in the mean number of R1s for the three FAS word groups for both Group 1 and Group 2 Ss in the test. For Group 1, the main effect of word groups was reliable [$F(2,57) = 13.74, p < .05$]. A Duncan Range Test indicated that the high-FAS group did not differ reliably from the mod-FAS group, although all other comparisons were statistically significant ($\alpha = .05, df = 57$). The main effect of word group was also reliable for Group 2 [$F(2,57) = 7.07, p < .05$]. As in Group 1, the mean number of R1s in the high-FAS group did not differ from the mean number of R1s in the mod-FAS group, although all other differences were statistically reliable.

The difference between the mean percentage of R1s in the test for Group 1 and Group 2 was small and statistically unreliable (93.2% vs 94.8%) as was the difference in the percentage of R1s in the test for males and females (93.6% vs 94.4%).

The relationship between the frequency of R1 to each of the 60 words in the test and the frequency of R1 in the SFVSC Drum Norms was statistically reliable ($r = .58, df = 58, p < .01$). As in Experiment 1, the amount of facilitation produced by training varied inversely with the initial FAS of R1. The same conclusions were reached on the basis of the relationship of the frequency of R1 to each of the 60 words in the test and the frequency of R1 to each of the 60 words in the WAT ($r = .58, df = 58, p < .01$).

GENERAL DISCUSSION

Training substantially increased the number of R1s in the test, and the amount of increase varied inversely with the initial normative frequency of R1. The magnitude of the difference among the four (Experiment 1) or three (Experiment 2) FAS groups was small due to the constraint produced by a ceiling effect and the fact that the FAS of the stimulus words varied on a continuum which had been arbitrarily divided in order to establish separate FAS groups. A larger word pool would increase the likelihood that reliable differences would be found between word groups of intermediate FAS values. Nonetheless, the main effect for groups was reliable with statistically significant differences between the high-FAS word group and the low-FAS word group in both experiments.

The present studies provided no support for the hypothesis that a second training trial increases the level of R1 responding in the test. Although it is possible that a single 8-sec training trial is sufficient to produce

asymptotic R1 responding in the test, the present data cannot be adequately interpreted, since significantly more R1s were given in the WAT by Group 1 Ss than by Group 2 Ss (Experiment 2). Thus, the effect of multiple training trials must be examined in the context of further research.

The methods of analysis employed provided no evidence that direct assessment of FAS produces predictions that are more precise than those based on normative data. These results may be due to the highly restricted range of R1 test frequencies which occurred as a result of the great amount of facilitation produced by the training procedure. A single training trial of short duration may produce sufficient effect to demonstrate the relationship between FAS and PAL without the tendency to push R1 responding in the test to 1.00 for all words.

The place of complex indicators in predicting PAL remains unclear. Although it was predicted (Experiment 1) that training would produce more facilitation for a group of words where R1 was strong and R2 was weak, as opposed to a group of words where there was a strong competing secondary response, the prediction was not supported.

Although the demonstrated relationship between FAS and PAL confirmed a hypothesis that is a clear deduction from widely held association or S-R mediation theory, the results were not consistent with those studies which have failed to show a general effect on PAL resulting from differential FAS with adult Ss (cf. Jenkins, 1963; Martin, 1964; Palermo, 1963). The present study differed in many respects from the earlier studies of Jenkins and his associates. The slow presentation rate, the large number of word pairs, and the use of only one or two training trials represented substantial procedural variations. Moreover, the WAT prior to the study trials and the use of the recall PAL method produced a very different set of antecedent conditions from those provided by other investigators. Finally, the use of a restricted sample of stimuli from the Connecticut Norms represented a departure from most studies that have been directed to related problems

in the last decade, for the majority of studies have used samples of words taken from the 100-word Kent-Rosanoff list (Russell & Jenkins, 1954).

The amount of facilitation observed in the present study was quite extraordinary, in view of the fact that either 60 or 68 words pairs were presented only once or twice. Had Ss been presented 60 pairs of nonassociated words, some Ss would certainly have commented on the *impossibility* of the task. However, the typical S tended to behave as if there were a rule, and, to the extent that the rule could be made to work for him, a *correct* response might be given to all stimuli. That is, having instructed himself that the pairs were "words that belong together" or "words that most people give together," S apparently did not attempt to develop a separate mediator for each word pair, a process that would have required more training than was provided in the studies discussed in this report.

Simply adopting the correct mediational rule might account for general facilitation, but the fact that the amount of facilitation varied as a function of FAS suggested that there was considerable variation in the success achieved through the application of the response set. Since a response with high FAS is one that, by definition, "goes with" the stimulus for many Ss, imposing the mediational rule would result in correct responses when there is high FAS between stimulus and response. Since there is no single response that "goes with" a stimulus when FAS is low, correct responses were less likely to occur despite the application of the correct response set.

In most experiments designed to test some hypothesis relating FAS to learning, there has been a tendency to include unrelated S-R pairs as controls. Since it is well established that very small amounts of associative strength facilitate PAL (Palermo, 1963), the inclusion of unrelated word pairs seemed quite unnecessary. Thus, there was no instance in which the appropriate mediational rule, once adopted, failed to be confirmed.

REFERENCES

BATTIG, W. F., & BRACKETT, H. R. Comparison

of anticipation and recall methods in paired-associate learning. *Psychological Reports*, 1961, 9, 59-65.

BOUSFIELD, W. A., COHEN, B. H., WHITMARSH, C. A., & KINCAID, W. D. The Connecticut free associational norms. Technical Report No. 35. Contract Nonr 631 (00), Office of Naval Research, 1961.

BROTSKY, S. J., BUTLER, D. C., & LINTON, M. L. Association time commonality, and the test-retest reliability of free association responses. *Psychonomic Science*, 1967, 9, 319-320.

BROTSKY, S. J., & LINTON, M. L. San Fernando Valley State College memory drum norms. Technical Report No. 2, 1968, San Fernando Valley State College, Research Grant 4.259.01, SFVSC Foundation.

EDWARDS, A. L. *Experimental design in psychological research*. New York: Rinehart, 1960.

JENKINS, J. J. Mediated associations: Paradigms and situations. In C. N. Cofer and B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963. Pp. 210-257.

MARTIN, J. G. Associative strength and word frequency in paired-associate learning. *Journal of Verbal Learning & Verbal Behavior*, 1964, 3, 317-320.

PALERMO, D. S. Word associations and children's verbal behavior. In L. P. Lipsitt and C. C. Spiker (Eds.), *Advances in child development and behavior*. Vol. 1. New York: Academic Press, 1963. Pp. 31-68.

RUSSELL, W. A., & JENKINS, J. J. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test. Technical Report No. 11, 1954, University of Minnesota Contract N 8 Onr 66216.

WICKLUND, D. A., PALERMO, D. S., & JENKINS, J. J. The effects of associative strength and response hierarchy on paired-associate learning. *Journal of Verbal Learning & Verbal Behavior*, 1964, 3, 413-420.

NOTES

1. This research was supported by a research grant to the first author from the SFVSC Foundation and by a faculty research award to the second author from the San Diego State College Foundation. The authors are indebted to Dr. David Parker for assistance in the construction of the apparatus and to Mr. Donald Bizarri and Mr. William America who ran the Ss. The final draft of this report was written during the first author's tenure as NIH Special Fellow, Harvard University.

2. Association time data were collected during the course of this study, but will not be referred to again in this report. These data are reported, in part, elsewhere (Brotsky, Butler, & Linton, 1967).

ERRATUM

WHITE, MURRAY J. Identification and localization within digit and letter spans. *Psychonomic Science*, 1969, 14 (6), 279-280.—Page 279, 2nd column, 3rd line from bottom of 2nd paragraph: "Position by Material" should read "Position by Exposure."