

The function of one-word mediators in the recall of word pairs*

FRANCIS S. BELLEZZA† and ALEX J. POPLAWSKY

Ohio University, Athens, Ohio 45701

The problem of demonstrating that natural language mediators play a role in learning and are not epiphenomena resulting from learning is an important problem in cognitive learning theories. Using a cued-recall and a free-recall learning task, Ss were requested to add a one-word mediator to some of the pairs of concrete nouns presented. The mediated pairs were learned better than the control pairs in both tasks. Both words were recalled only when the mediator was also recalled. Also, one-word mediators were the most effective recall cues and were the best recalled words in free recall. A two-stage learning model adequately described the data. However, a counterargument can be made which considers the mediator to be a high associate of one of the words presented and actually has no direct link to the other presented word. A possible experimental resolution of the problem is discussed.

The possibility that elaborative strategies are routinely used by the human learner is receiving increasing attention in verbal learning theories. These elaborative strategies can involve conceptual mediation (Mandler, 1963), verbal elaboration (Montague, 1972), or visual imagery (Paivio, 1971). How necessary an elaborative strategy is to verbal learning is still in question, but the concept does play some role in many different theories (Melton & Martin, 1972). One technique used in the study of verbal elaboration in learning is natural language mediation (Montague, Adams, & Kiess, 1966; Prytulak, 1971), in which Ss are presented with to-be-remembered material such as CVCs or CVC word pairs and are requested to transform the material into meaningful words, phrases, or sentences in order to better acquire and retain it. These transformations are the natural language mediators. It has been shown that correct recall of the presented material is strongly related to the recall of the natural language mediators used to code it (Montague et al, 1966).

One problem in the study of natural language mediators is to demonstrate that they play a role in learning and are not merely epiphenomena accompanying learning (Adams & McIntyre, 1967; Underwood, 1972). In Experiment I Ss were presented with 24 pairs of nouns; half of these pairs were simply read aloud by the Ss and studied, whereas, upon presentation of the other half of the pairs, the Ss were asked to add a third word which described, connected, or associated the two words presented. For recall of the control pairs, S was cued with either the first word of the pair (A) or the second word (B) and had to respond with the other word of the pair. For recall of the mediated pairs, which included the one-word mediator, the Ss were cued with A, B, or the mediator word (M) and had to respond with the other two words. If the act of generating a mediator facilitates learning, then

performance should be better on the mediated pairs than on the control pairs. If it is true that the mediator provides the only connecting link between the two words presented, as shown in Fig. 1, then the mediator should always accompany recall of one of the presented words when the other is used as a cue. Also, M should be an effective cue for the recall of A and B, but how effective a cue M will be depends on whether or not the links between A and M and between B and M are bidirectional. It simplifies matters if it can be assumed that, if A is recalled when M is presented, then M will be recalled when A is presented. The question of associative symmetry in paired-associate learning is still unresolved (Ekstrand, 1966). However, two procedures were used in Experiment I which maximized the bidirectionality of the links between A and M and between M and B. The Ss were given a practice list that included both mediated and nonmediated pairs, and all three word types were used as recall cues. This informed the Ss that the word pairs or word triads could be cued by any of their components. Second, the words used were concrete nouns. Associative bonds between high-imagery words are approximately bidirectional (Yarmey & O'Neill, 1969).

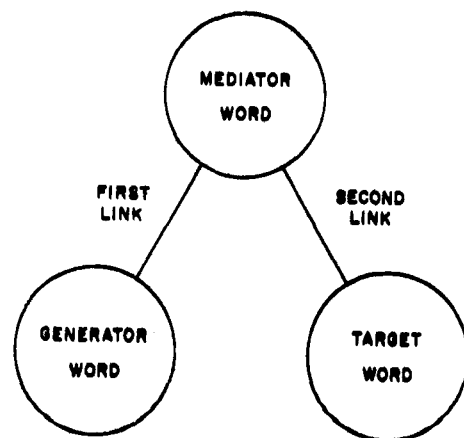


Fig. 1. Structure of the mediated pair.

*This research was supported in part by Grant OURC 437 to the first author by the Ohio University Research Committee.

†Requests for reprints should be sent to Francis S. Bellezza, Department of Psychology, Ohio University, Athens, Ohio 45701.

In order to make a definite prediction about how effective a recall cue M will be compared to A and B, the mediating process is proposed to occur in the following manner: The S, when presented with a word pair to which he must add a one-word mediator, chooses one of the words as the generator word and generates associates of it until a word is selected that can act as a mediator word to connect the generator word to the other word in the pair, the target word. Although the S can always give a third word when requested to do so, the probability that the word generated is an effective mediator is *f*. Generating a mediator represents the formation of the first link, as shown in Fig. 1. The probability that the S remembers the transformation rule connecting the mediator with the target word is *s*. This forms the second link in Fig. 1. The transition matrix giving the probabilities of the formation of the links for any mediated pair on any given presentation trial can be represented as follows:

$$\begin{matrix}
 & L2 & L1 & L0 \\
 L2 & \left[\begin{matrix} 1 & 0 & 0 \\ s & 1-s & 0 \\ fs & f(1-s) & 1-f \end{matrix} \right] & & \\
 L1 & & & \\
 L0 & & &
 \end{matrix} \quad (1)$$

where State L0 represents the pair as having no links, State L1 represents the pair as having the first link, and State L2 represents the pair as having two links. The model represented by Eq. 1 is a two-process Markov model similar to the three-process model discussed by Atkinson and Crothers (1964). It is assumed in the present model that no intertrial forgetting occurs; if an S's response during a test trial indicates that a link has been formed, then that link is assumed not to be forgotten on subsequent trials. In order to generate a formula which expresses the expected number of words recalled when any element in Fig. 1 is used as a cue, assume that on any given test trial *t* for any given pair the probability that the first link has been formed is *f*(*t*) and that the second link has been formed is *s*(*t*). If A is presented as a cue, then the expected number of words recalled from the word triad is

$$E_t(R|A) = .5 \left\{ 2f(t)s(t) + 1f(t)[1 - s(t)] + 0[1 - f(t)] \right\} + .5 \left\{ 2f(t)s(t) + 0f(t)[1 - s(t)] + 0[1 - f(t)] \right\}$$

where the expression in the first braces is the expected number of correct responses if S chose A as the generator word, which is assumed to occur with probability equal to .5. The expression in the second braces is the expected number of responses if A was chosen as the target word. Similarly, $E_t(R|B) = E_t(R|A)$.

If the one-word mediator is presented as the recall cue, then

$$E_t(R|M) = 2f(t)s(t) + 1f(t)[1 - s(t)] + 0[1 - f(t)].$$

Therefore

$$E_t(R|M) = E_t(R|A) + .5f(t)[1 - s(t)] \quad (2)$$

$$E_t(R|M) > E_t(R|A) = E_t(R|B).$$

In Experiment I this prediction will be tested and the parameters *f* and *s* will be estimated.

EXPERIMENT I

Method

Subjects. The Ss were 32 volunteers drawn from a pool of introductory psychology students at Ohio University. The Ss received course credit for their participation.

Materials. Two lists of 24 word pairs were made using 96 one-syllable concrete nouns having Thorndike and Lorge general frequency counts of 10 to 40 per million. The pairs were constructed such that the two words bore no obvious relation to each other. Examples are ranch-web, tin-beard, and film-pear. Twelve of the pairs were randomly drawn from each list and a red X was printed two spaces to the right of each pair. A blue O was printed two spaces to the right of each of the other 12 pairs on the list. A practice list of three pairs was also made with two pairs followed by an X and one pair followed by an O.

Procedure. The Ss were run individually using a Lafayette IBM memory drum. Each S was presented the practice list for three trials and one of the two experimental lists for three trials. The lists of pairs were randomized with each presentation. Each pair was presented for 10 sec, during which time the S read aloud the two words. For those pairs followed by an X, S was instructed to add a third word which described, connected, or associated the two words. E then recorded this word as the one-word mediator for that pair. For those pairs followed by an O, S merely read aloud the two words and added no third word. The procedure was counterbalanced so that half of the Ss gave a one-word mediator for the O pairs. A small card placed on the memory drum reminded S of the meaning of the X and O codes. After each presentation of the list, E read aloud one word from each control pair and one word from each mediated pair. For each cue S was given 10 sec to respond verbally with the other word or with the other two words. No information was given regarding the type of pair each cue represented. For the 12 control pairs, six first words and six second words were used as cues on each trial. For the 12 mediated pairs, four first words, four second words, and four mediated words were used as cues. Over the three trials, the first word (A) was used as a cue twice for some control pairs and, for the other control pairs, the second word (B) was used twice. For each mediated pair A, B, and M were each used once as a cue over the three trials. After each presentation of the list, the first three pairs tested were randomly drawn from the first half of the presented list to prevent responding from short-term store.

Analysis. The first analysis performed was a 2 by 3 by 2 by 2 analysis of variance, with the specific factors being list form, trials, mediation condition, and position of cue word. Only list form was a between-Ss factor. In this analysis, data involving the mediator word was ignored and only the proportions of recall of A given B and of B given A for the control and mediated pairs were analyzed. The second analysis was a 2 by 3 by 3 analysis of variance, with the specific factors being list form, trials, and

position of cue word. This analysis was limited to data from mediated pairs. The third analysis used Subroutine Stepit (Chandler, 1965) to estimate the parameters f and s of Eq. 1 using data from recall of the mediated pairs.

Results

The first analysis of variance showed significant effects due to mediation condition [$F(1,30) = 11.30, p < .005$], trials [$F(2,30) = 175.87, p < .001$], and position of cue word [$F(1,30) = 5.38, p < .05$]. For the mediated pairs the proportion of correct recalls was .35, .65, and .85 over the three trials. For the control pairs the proportions were .19, .53, and .74, respectively. The A words were a more effective cue than were the B words, with $P(B|A) = .57$ and $P(A|B) = .53$. In the second analysis of variance, trials was significant [$F(2,60) = 166.03, p < .001$], but the position of the cue word effect was only marginally significant [$F(2,60) = 2.87, p < .10$]. However, a priori contrasts showed that M was a more effective cue than were A and B [$F(1,60) = 5.69, p < .025$]. There was no difference in the effectiveness of A and B. The recall results are shown in Fig. 2. Table 1 gives the recall performance with each cue for each list for all three trials. The symbol E in this context represents an error of omission or of intrusion. It can be seen from Table 1 that the probability of recalling B given A or of recalling A given B for the mediated pairs is zero unless the mediator is also recalled. Also, there was little evidence of forgetting. In the first list, there was forgetting in only .06 of the items and on the second list in only .07.

In order to estimate the parameters f and s , the data points $P(E,E|W)$, $P(M,E|W)$, $P(E,E|M)$, and $P(W,E|M)$ were computed for each trial. The symbol W represents either A or B, since it has been assumed that each is

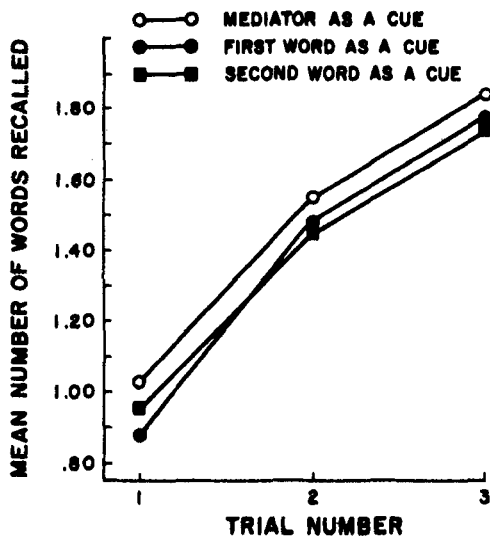


Fig. 2. Recall of the mediated pairs using the different words of the pair as a cue.

Table 1
Order of Words Recalled for Each Recall Cue

Words	Probability of Recall $P(j,k i)$					
	List 1 Trial			List 2 Trial		
	1	2	3	1	2	3
Mediated Pairs						
A E M	.16	.02	.03	.09	.11	.06
B E M	.20	.11	.06	.41	.14	.06
A B M	.22	.42	.63	.11	.39	.48
B A M	.13	.33	.27	.16	.22	.36
B E A	.00	.00	.00	.00	.00	.00
M E A	.16	.17	.06	.14	.14	.06
B M A	.30	.48	.56	.20	.30	.58
M B A	.11	.23	.33	.13	.31	.23
A E B	.00	.00	.00	.00	.02	.00
M E B	.33	.14	.11	.28	.23	.16
A M B	.19	.34	.52	.20	.33	.38
M A B	.19	.34	.33	.08	.35	.39
Control Pairs						
B A	.23	.56	.75	.22	.57	.72
A B	.12	.50	.77	.20	.45	.74

equally likely to be the generator word. These data points are plotted in Fig. 3. Using the model presented

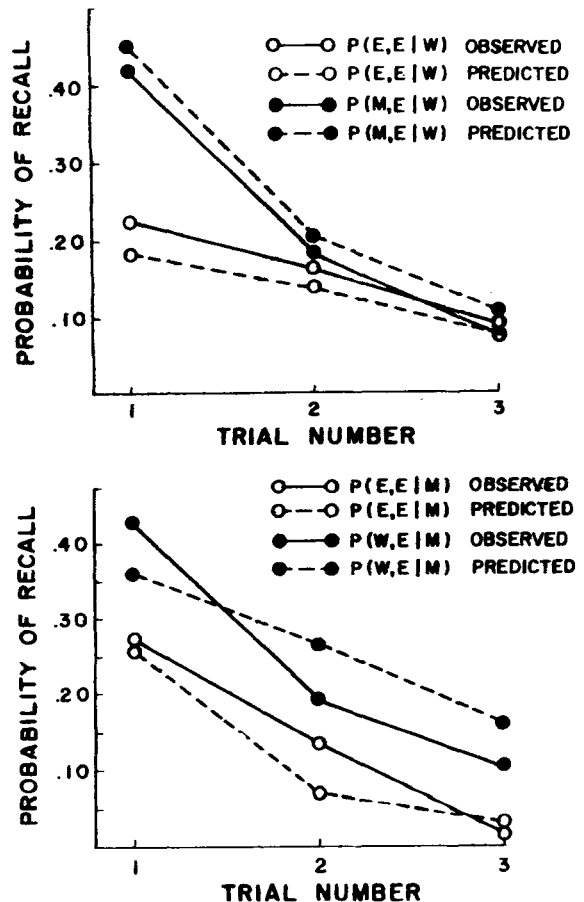


Fig. 3. Observed and predicted cued-recall performance.

in Eq. 1, the expressions for $P(E,E|M)$, etc., on any trial t in terms of f and s are:

$$\begin{aligned} P_t(E,E|M) &= P_t(L0) \\ P_t(W,E|M) &= P_t(L1) \\ P_t(E,E|W) &= P_t(L0) + .5P_t(L1) \\ P_t(M,E|W) &= .5P_t(L1) \end{aligned}$$

where $P_t(L0)$ is the probability of any item being in State L0 on trial t , etc. Also,

$$P_t(L0) = (1 - f)^t$$

$$P_t(L1) = f(1 - s)[(1 - f)^t - (1 - s)^t]/(s - f), \text{ for } f \neq s.$$

And

$$P_t(L2) = 1 - P_t(L0) - P_t(L1).$$

Using Subroutine Stepit, the minimum chi-square estimates for f and s were found to be $\hat{f} = .73$ and $\hat{s} = .51$. Although the distribution of differences only approximates a chi-square distribution (Atkinson & Wickens, 1971), the chi-square value found by comparing the observed and predicted points was $\chi^2(10) = 11.97$, which indicates that the observed and predicted points are not significantly different. The predicted data points are also plotted in Fig. 3.

The results of Experiment I indicate that the mediator is more than an epiphenomenon accompanying learning, for it can be used to cue the presented pair, and it is a significantly better recall cue than either of the two words presented. Also, the two-stage learning model presented in Eq. 1 provides a reasonably good fit to the data. In Experiment II the procedure used was identical to that used in Experiment I, except that after each list presentation the Ss were asked to free recall all the words presented, including the one-word mediators. As in Experiment I, more words should be recalled from the mediated pairs. Because the mediators are supplied by the S, they should be recalled better than the words presented. Also, few words from the mediated pairs should be recalled without their accompanying mediator word.

EXPERIMENT II

Method

Procedure. The procedure was the same as that used in Experiment I, except that after each presentation of the list each S was asked to count backward by sevens as fast as he could for 20 sec. This was done to eliminate any effects of short-term store. S was then given 240 sec to try to recall as many of the words and mediators as he could. The same procedure was used on the practice list, but S was given only 30 sec for recall. The same two lists were used as in Experiment I, with 16 Ss run on each list.

Analysis. A 2 by 3 by 2 by 2 analysis of variance was performed, with the specific factors being list form, trials, mediation condition, and word position. Only list form was a

between-Ss factor. The dependent variable was the proportion of presented words recalled. The second analysis of variance was a 2 by 3 by 3, with the factors being list form, trials, and word position, which included the mediated word. This analysis was performed on the recall data from the mediated pairs only.

Results

The first analysis of variance showed that the mediated A and B words were recalled significantly better than the control A and B words [$F(1,30) = 43.75$, $p < .001$]. The proportion of words recalled from the mediated pairs were .30, .53, and .69 over the three trials, and the proportions for the control pairs were .12, .37, and .54, respectively. The main effect of trials was also significant [$F(2,60) = 263.43$, $p < .001$], as was the main effect of word position [$F(1,30) = 20.21$, $p < .001$]. The A words were recalled .40 of the time over all trials, whereas the B words were recalled .45 of the time. In addition, the Mediation Condition by Word Position interaction was significant [$F(1,60) = 6.17$, $p < .025$], with .47 of the A words and .55 of the B words recalled from the mediated pairs. In the control pairs, the proportions were .33 and .35, respectively. Because the Ss read aloud the words in the order A, B and then added M in the mediated pairs, it could be that B was more likely to be the generator word and would be more likely to be recalled. However, in Experiment I there was no difference in the cued recall of A and B. The List Form by Word Position interaction was also significant [$F(1,30) = 11.30$, $p < .001$], with .37 of the A words and .46 of the B words recalled from List 1. In List 2 the proportions were .43 and .44, respectively. The only explanation for this interaction is that the B words in List 1 were easier to learn.

The second analysis of variance showed significant main effects due to trials [$F(2,60) = 113.47$, $p < .001$] and word position [$F(2,60) = 21.42$, $p < .001$]. Recall performance on the three word types is shown in Fig. 4. A priori orthogonal contrasts showed that M words were recalled better than A and B words [$F(1,60) = 22.67$, $p < .001$] and that B words were recalled better than A words [$F(1,60) = 20.16$, $p < .001$]. The one other significant source of variation was the List Form by Word Position interaction [$F(2,60) = 5.82$, $p < .025$]. As was shown in the first analysis, B words were recalled better than were A words in List 1, but recall was about the same in List 2. Using the parameter values $f = .73$ and $s = .51$ estimated in Experiment I, the proportions of A and B words that should be recalled on each trial were computed from the proportion of M words recalled. This predicted performance is indicated by the dotted line in Fig. 4. The line should fall halfway between the recall levels of A and B because these words were predicted to be recalled equally well. As can be seen from Fig. 4, more recall is occurring than can be predicted from the recall of the mediator words. It is possible that the free-recall task changes the Ss' learning strategy somewhat from that used in cued recall, and it

may not be appropriate to try to predict performance using the parameter values found in Experiment I. Table 2 shows what words were recalled together from the word pairs and the word triads. Only a small proportion of A and B words were recalled from the mediated pairs without the recall of the corresponding mediator word.

As predicted, the mediator words were significantly more accessible in free recall than were any of the presented words. In addition, the presented words that were mediated were recalled better than words in the control pairs. As in cued recall, few words were recalled from the mediated pairs without the recall of the corresponding mediator word.

DISCUSSION

The results of Experiments I and II seem to provide support for the hypothesis that the one-word mediators facilitate and play a functional role in learning the word pairs. The results show that the mediated pairs were learned better than the control pairs; the results also show that the mediator is the most effective recall cue and is most accessible in free recall. In both cued recall and free recall, Ss rarely recalled a word from a mediated pair without also recalling the mediator word or having it presented. The word added by S to the mediated pair was not merely an epiphenomenon resulting from the learning of the pair. If it were, the added word could not function as a cue for the pair. However, a

Table 2
Free Recall of Words from Control and Mediated Pairs

Words	Probability of Recall P(i, j, k)					
	List 1 Trial			List 2 Trial		
	i	j	k	1	2	3
Mediated Pairs						
A E E	.02	.03	.02	.02	.03	.03
B E E	.03	.04	.04	.02	.02	.01
M E E	.07	.02	.01	.06	.03	.01
A B E	.02	.00	.00	.01	.00	.00
A M E	.04	.05	.03	.05	.03	.03
B M E	.16	.16	.19	.07	.05	.08
A B M	.18	.39	.54	.19	.47	.64
Control Pairs						
A E	.06	.05	.06	.07	.06	.06
B E	.06	.11	.11	.06	.09	.05
A B	.05	.26	.42	.07	.31	.51

counterargument can be made stating that the mediator in these experiments represents additional learning and does not facilitate learning the word pair. It is possible that S perceived the mediated pairs as being more important and tried harder to learn these pairs, perhaps selectively rehearsing them during the presentation of the control pairs to which S did not have to add a mediator. Also, if mediation facilitated learning, then the mediated pairs should have been learned at a faster rate, which would have resulted in a significant Mediation Condition by Trials interaction. This interaction was never significant. The one-word mediator could be merely a word that S adds to the pair rather than a representation of a cognitive process by which the pair is learned. The model given in Eq. 1 may represent a learning sequence in which the additional word is given by S only because E requests it. The S does this by generating a high associate to the generator word with probability f such that $P(G|M) = P(M|G) = 1$, where G represents the generator word and M represents the high associate. After he does this, S learns the presented pair GT with probability s , where T represents the target word. This means that GM can be recalled without T, and that GT will never be recalled without M. The hierarchical structure represented in Fig. 1 is now MGT rather than GMT. However, Eqs. 1 and 2 remain the same, and M is still the best recall cue. This equivocation in interpretation is unfortunate. If the one-word mediator does represent the coding of new material in memory, then the processes of acquisition and retention of the to-be-remembered material could be better traced and studied. An experiment requiring one group of Ss to give a high associate to one word of each presented pair compared to a group of Ss required to give a mediator might show the second group superior and, at the same time, equate the processing time the mediated and control pairs receive. At the present time, there is still doubt as to what the one-word mediator represents.

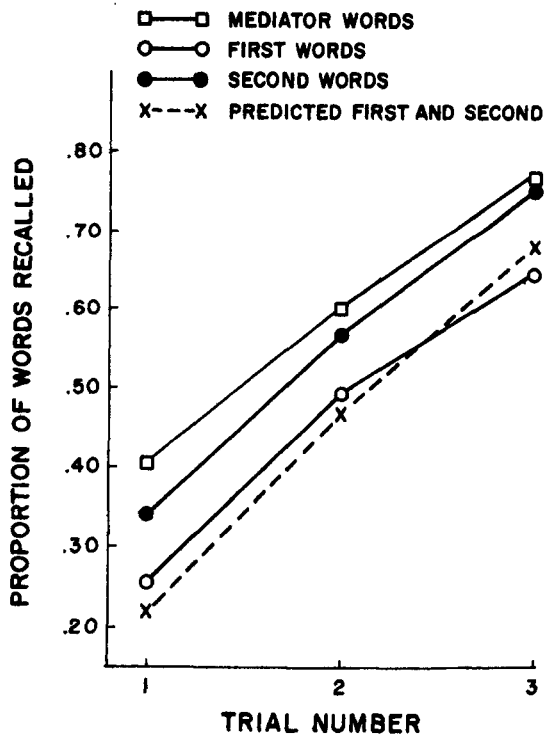


Fig. 4. Free recall of the different words in the mediated pairs.

REFERENCES

- Adams, J. A., & McIntyre, J. S. Natural language mediation and all-or-none learning. *Canadian Journal of Psychology*, 1967, 21, 436-449.
- Atkinson, R. C., & Crothers, E. J. A comparison of paired-associate learning models having different acquisition and retention axioms. *Journal of Mathematical Psychology*, 1964, 1, 285-315.
- Atkinson, R. C., & Wickens, T. D. Human memory and the concept of reinforcement. In G. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press, 1971.
- Chandler, J. P. Subroutine Stepit. Program QCPE 66. Quantum Chemistry Program Exchange, Indiana University, 1965.
- Ekstrand, B. R. Backward associations. *Psychological Bulletin*, 1966, 65, 50-64.
- Mandler, G. Comments on Professor Jenkins' paper. In C. N. Cofer and B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963.
- Melton, A. W., & Martin, E. (Eds.), *Coding processes in human memory*. Washington, D.C: Winston, 1972.
- Montague, W. E. Elaborative strategies in verbal learning and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*. Vol. VI. New York: Academic Press, 1972.
- Montague, W. E., Adams, J. A., & Kiess, H. O. Forgetting and natural language mediation. *Journal of Experimental Psychology*, 1966, 72, 829-833.
- Paivio, A. *Imagery and verbal processes*. New York: Holt, 1971.
- Prytulak, L. S. Natural language mediation. *Cognitive Psychology*, 1971, 2, 1-56.
- Underwood, B. J. Are we overloading memory? In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C: Winston, 1972.
- Yarney, A. D., & O'Neill, B. J. S-R and R-S paired-associate learning as a function of concreteness, imagery, specificity, and association value. *Journal of Psychology*, 1969, 71, 95-109.

(Received for publication September 24, 1973;
accepted November 25, 1973.)