Response-set size effects in recall and recognition

MAY F. D'AMATO and AGNES MARCHESE
Brooklyn College, City University of New York, Brooklyn, New York

(Susan K. Manning, *sponsor*)

A retroactive inhibition design was employed to study the role of rules in paired-associate learning and retention. Subjects first learned one of three unmixed lists of paired nonsense syllables. A rule applied to all items in two of the lists: Pair members changed first letter (rhymed) or changed middle letter. In the third list, pair members were not related. A common second list was learned, and either a first-list recall test or a four-alternative forced-choice recognition test followed. First-list learning and recall were superior for lists with rules. Recognition, which did not show a ceiling effect, was not superior for lists with rules when both target and distractors followed a rule. The results supported the hypothesis that rules facilitate retrieval rather than storage, and that they do so to the extent that they reduce response-set size.

Rhyming can effectively aid verbal performance. Rhyming pairs, for example, are consistently easier than pairs of unrelated words (Bower & Bolton, 1969). One hypothesis proposed to account for the advantage of rhyming paired associates is that the presence of a rhyming rule restricts the number of potential responses of each list pair. With reduced response-set size comes improved performance. Hunt, Elliot, and Spence (1979), for example, found that, with words of low meaningfulness, rhyme cues led to better recall following a single study trial than did associative cues. The response-set size for these rhyme cues was smaller than that for the associative cues. More recently, Nelson, McEvoy, and Friedrich (1982) showed that when rhymes were used as cues for recall following a single study trial, the smaller the set size, the better the recall. Thus, evidence suggests that performance is inversely related to response-set size and that the effectiveness of rhyming stems from the reduced set size it entails.

The response-restriction hypothesis need not apply to only a rhyming rule. Any rule that reduces response-set size should improve performance. The hypothesis also implies that the extent of improvement should vary directly with the extent of reduction in the size of the response set. D'Amato and Rubenstein (1981) tested this prediction by comparing a list of rhyming pairs with other lists that followed progressively more restrictive rules. They did indeed find that the greater the reduction in the size of the set defined by the rule, the better the paired-associate learning.

The response-restriction hypothesis of rule superiority may be readily extended to paired-associate retention. As with learning, retention should be inversely related to the size of the response set defined by the rule. Using a retroactive inhibition design, D'Amato and Guber (1982) obtained results quite consistent with this prediction. Recall of the first-list responses progressively improved as rules became more restrictive. In addition, a forced-choice recognition procedure that effectively equated response-set size under all conditions failed to show retention differences among the groups. These results suggested that superior recall for lists with rules was not due to superior storage of the individual pairs, but rather to the superior retrieval provided by the rule that stemmed from the reduced size of the response set relevant to each list pair. The recognition results of the D'Amato and Guber study had to be interpreted cautiously, however, because recognition was at or close to ceiling for all groups.

The present study sought to replicate and extend the D'Amato and Guber (1982) experiment while avoiding a ceiling effect. This was accomplished by lowering the criterion of first-list learning and by using nonsense syllables instead of words so that the number of distractors accompanying each target on the recognition test could be increased.

METHOD

Subjects
Sixty-three students participated in order to fulfill a requirement for introductory psychology. They were tested individually and were assigned randomly in blocks of 7, with each condition occurring once in each block.

Materials
Four unmixed lists, with each containing 10 pairs of nonsense syllables, were used. The same nonsense syllables were employed as stimuli in all four lists. These nonsense syllables were: NEF, HIX, KER, SAQ, TUF, BUZ, COZ, JEK, LOH, and VIB. Glaze association values of these syllables ranged from 13% to 93%. With but one exception, the three responses selected for each stimulus in the three lists used for original learning were of equal Glaze value, ranging from 47% to 60%. For

The authors' mailing address is: Department of Psychology, Brooklyn College, City University of New York, Brooklyn, NY 11210.

two of these three initially learned lists, a general rule applied: Either each pair rhymed (e.g., NEF-SEF),or the middle letter changed (e.g., NEF-NIF). The third list consisted of unrelated pairs (e.g., NEF-LUZ). Each subject learned one of these three first lists. The fourth list, which was the same for all subjects, served for interpolated learning. It consisted of the first-list stimuli paired with unrelated responses (e.g., NEF-JUN). The Glaze value of these responses ranged from 67% to 80%. Four lists were used for the forced-choice recognition test. In all of them, each first-list stimulus appeared along with four alternatives, the first-list response to the stimulus plus three distractors. For rule lists, distractors as well as targets followed the rule. Thus, targets and distractors rhymed when the rhyme list was tested (e.g., NEF-SEF, PEF, DEF, VEF), and the middle letter changed when the middle-letter-changed list was tested (e.g., NEF-NIF, NOF, NUF, NAF). The recognition test for the unrelated list took one of two forms. In one list, all distractors were unrelated to each other and to the target syllable; in the other, the distractors rhymed with the target for half the items, and the middle letter changed for the remaining half. In selecting distractors for each target, association values were equated across the four lists insofar as was possible. Target position varied within each list. Between lists, target position remained the same.

Procedure

The study employed a retroactive inhibition design. Prior to the first trial on the original list, standard written paired-associate instructions were presented. The subjects learning the rhyming and middle-letter-changed lists were informed of the rule applying to their list. On the first trial, all subjects repeated aloud the experimenter's pronunciation of each syllable as it appeared. Each list was presented on a Stowe memory drum paced at a 2:2-sec rate with a 4-sec intertrial interval for as many anticipation trials as were necessary to achieve a criterion of 4/10 correct. All lists followed the same sequence of three different random orders of the pairs. Approximately 2 min elapsed between first- and second-list learning; during that 2 min, the subjects read instructions for the interpolated list. Following 10 anticipation trials on the interpolated list, the subjects were tested for retention of first-list responses. They were presented with a pencil and a sheet of paper on which the entire set of stimuli appeared, and were told to write down the syllable that went with each one on the first list, and to guess if necessary to avoid omitting any items. The same sequence of stimuli was used on all the retention tests. For recognition, each stimulus appeared along with four alternatives followed by a blank space, where the response was to be entered. Three minutes were allowed for the retention test. In sum, there were seven groups in all, with nine subjects in each group. Three groups were given a recall test of their first list, which had been composed of pairs that rhymed, that had the middle letter changed, or that were unrelated. Of the four groups receiving the forced-choice recognition test, one group had a first list of rhyming pairs, one of middle-letter-changed pairs, and two of unrelated pairs. The latter two groups differed only in the type of distractors employed, these being either unrelated or related to the target.

RESULTS AND DISCUSSION

Trials to reach first-list criterion yielded the following means: middle letter changed, 2.78; rhyme, 3.89; and unrelated, 10.15. Lists significantly affected performance [F(2,60) = 29.89, p < .001]. Orthogonal comparisons of these means showed that the two rule lists combined differed significantly from the unrelated list (p < .001) but not from one another (p > .05). The number of correct anticipations on the last trial of interpolated learning did not differ [(F2,60) = 2.20, p > .05], indicating that the groups were comparable in their performance on this second list.

The mean numbers of first-list responses correctly recalled on the recall test of retention were 6.00, 4.33,

and 1.00 for middle-letter-changed, rhyming, and unrelated lists, respectively. These means were significantly different [F(2,24) = 25.93, p < .001]. Orthogonal comparisons revealed that the middle-letter-changed condition was significantly superior to the rhyming condition (p < .05) and that both these rule lists were significantly superior to the unrelated list condition (p < .001). Forced-choice recognition scores produced these mean numbers of first-list responses correctly recognized: middle letter changed, 5.78; rhyming, 7.33; unrelated-related (unrelated list with rule following distractors), 7.67; and unrelated-unrelated (unrelated list with unrelated distractors), 9.33. These means were significantly different [F(3,32) = 5.04, p < .01]. Orthogonal comparisons showed no difference (p > .05) between the rhyming and middle-letter-changed groups or between these two rule groups combined and the unrelated-related group. The difference between these three groups and the unrelated-unrelated group was significant (p < .01). In order to determine whether rhyming and middle-letter-changed alternatives presented a comparably difficult recognition task, performance of the unrelated-related group on items with rhyming and middle-letter-changed alternatives was compared. No difference was found [t(8) = 1.92, p > .05].

The results of this study replicated and extended the D'Amato and Guber (1982) findings under conditions that avoided the ceiling effects that limited interpretation of the results of that study. Thus, added support has been obtained for the response-restriction hypothesis of the effectiveness of rules, which maintains that rules benefit paired-associate performance to the extent that they reduce the size of the set of potential responses to the stimulus. First-list learning showed that the two paired-associate lists in which a rule applied were significantly easier to learn than a list of unrelated pairs. And, once again, the middle-letter-changed rule proved superior to the rhyming rule, although this difference was not significant. The response-restriction hypothesis would predict greater ease of learning for a list of pairs in which the middle letter changed than for a list of rhyming pairs, because the rhyming rule would reduce response-set size less than the middle-letter-changed rule would.

As expected, first-list recall yielded results that were similar to those for first-list learning. Lists with rules were recalled significantly better than was the list of unrelated pairs, and the middle-letter-changed rule was superior to the rhyming rule, in this case significantly so. Of greater interest are the recognition results. As mentioned, the recognition findings of D'Amato and Guber (1982) had to be interpreted rather cautiously, because a ceiling effect was evident. The present study clearly has avoided a ceiling effect for the two rule groups and for the unrelated group with rule-following distractors. Their mean percentage of first-list responses chosen from among the four possible alternatives ranged from 58% to 77%. The near-ceiling recognition for the unrelated group with unrelated distractors is consistent with results found by others

showing greater ease of recognition when distractors are not related graphemically and phonemically to target items (e.g., Gillund & Shiffrin, 1984). This group will not be further considered. The two rule groups and the unrelated-related group used alternatives that were similar, so that comparisons among these three groups are the appropriate ones for testing the response-restriction hypothesis. What do these comparisons show? They show no difference among the means. This was, of course, anticipated, because the response-set size was effectively equal for these three groups—a condition that, according to the hypothesis, makes rules ineffective.

The results suggest, then, that rules benefit retention, as they do learning, when they enable set size to be diminished. Added support for this conclusion comes from comparisons of recall with recognition scores. For the middle-letter-changed group, there is virtually no difference— 60% and 58% for recall and recognition, respectively. This would be anticipated, since response-set size was, in all likelihood, roughly equal under recall and recognition testing. With rhyming, there is a large difference between recall and recognition—43% and 73%, respectively. This is due most likely to the greatly reduced set size enjoyed by the rhyming list in recognition testing. Finally, if recall of the unrelated list is compared with recognition for either one of the recognition tests that were employed, even larger differences than with rhyming are observed, reflecting the relatively greater restriction of set size under recognition testing for the unrelated list.

The results of the present study are consistent with any one of a number of retrieval models of memory. For example, the SAM model (Gillund & Shiffrin, 1984) maintains that in recall the cues used to probe memory establish a search set that effectively restricts the domain of information in memory that need be considered. A search process is carried out on this set of information until an item that the subject decides is the response to the stimulus is sampled or until the search is terminated. As search-set size decreases, the probability increases that the target will be sampled when a particular search is performed. In the present study, it may be assumed that search-set size varied directly with the size of the set defined by the particular rule. In conjunction with the stimuli of the list pairs acting as probe cues, the particular rule established a corresponding search set. The rule yielding the smaller search set for each list stimulus increased the probability of sampling the response to that stimulus. Thus, the middle-letter-changed rule, which defined sets smaller than those defined by the rhyming rule, resulted in superior performance in recall and acquisition (which may be assumed to involve the recall processes described in the model). Lists following either rule would reduce search-set size considerably more than would a list of unrelated pairs. Thus, performance on the rule lists should indeed have surpassed that of the unrelated list.

The SAM model claims that recognition may be determined by the outcome of a familiarity process set in mo-

tion by the cues. A positive recognition response would occur if the familiarity value for an item were above criterion or were greater than that for the other alternatives. In the present study, in which set size was the same, rules could have aided recognition if they had increased target familiarity. This they apparently did not do, because recognition performance did not differ among the groups. Taken together, the results of this study suggest that the presence of a rule in a list does not result in superior storage of individual pairs. Rather, rules facilitate performance by reducing response-set size and thus increasing retrieval success.

Of course, it may be argued that the present results are due to some factor other than set size that covaried with the rules. For example, there may have been a difference in the similarity of pairs that rhymed, that had changed middle letters, or that were unrelated, a difference that could have created a difference in list difficulty. In order to determine whether there were indeed systematic similarity differences, 60 students, not otherwise involved in the study, were presented with one of the three lists used for first-list learning and were asked to rate each pair for similarity. A significant difference was found among the three means [$F(2,57) = 15.13$, $p < .01$]. The pairs of the unrelated list were rated as being significantly less similar than were those of each of the two rule lists, but a Tukey test revealed no difference between the two rule lists ($p > .05$). Apparently, then, differences in pair similarity are not responsible for the observed differences in acquisition and retention.

Attempts have been made to relate paired-associate learning to the frequency with which the two interior letters of a pair occur together in the language (e.g., Underwood, 1982). Did the lists employed in this study differ in this frequency? Apparently not, for when bigram frequency of list pairs was examined, no systematic differences among these lists emerged. It remains possible that some other factors that covaried with the rules employed in this study exist. However, an explanation of the findings in terms of the response-restriction hypothesis seems most satisfactory at the present time.

Although we have stated that the lists differed in the rules that they employed, it may be claimed instead that they differed in the retrieval cues that they encouraged. That is, in the middle-letter-changed list, first and last letters of the stimuli were used as retrieval cues, and these were superior to the middle and last letters that served as retrieval cues in the rhyming list. Such an interpretation is indeed plausible for the present study. However, in order to explain the differences in effectiveness of these retrieval cues, we would still have to resort to a response-restriction account such as has been presented here, because, without such an account, it would be difficult to explain why first and last letters were superior as retrieval cues for acquisition and recall but not for recognition. We have used the term *rules* to indicate that the interpretation presented should not be restricted to such physical

information as letter position, but should apply more generally whenever lists comprise pairs of related items—be they physically or otherwise related.

REFERENCES

BOWER, G. H., & BOLTON, L. S. (1969). Why are rhymes easy to learn? *Journal of Experimental Psychology*, **82**, 453-461.

D'AMATO, M. F., & GUBER, S. (1982). Rules and paired-associate retention. *Perceptual and Motor Skills*, **54**, 110.

D'AMATO, M. F., & RUBENSTEIN, V. (1981). Response-set size and paired-associate learning. *Psychological Reports*, **48**, 172-174.

GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.

HUNT, R. R., ELLIOT, J. M., & SPENCE, M. J. (1979). Independent effects of process and structure on encoding. *Journal of Experimental Psychology: Human Learning and Memory*, **5**, 339-347.

NELSON, D. L.,, McEVOY, C. L., & FRIEDRICH, M. A. (1982). Extralist cuing and retrieval inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **8**, 89-105.

UNDERWOOD, B. J. (1982). Paired associate learning: Data on pair difficulty and variables that influence difficulty. *Memory & Cognition*, **10**, 610-617.