NATIONAL LIBRARY
OTTAWA

BIBLIOTHÈQUE NATIONALE
OTTAWA

CANADA

NAME OF AUTHOR....Peggy Anne Runquist........

TITLE OF THESIS. Organizational Factors.....
..in...Paired-Associate..Learning

....

UNIVERSITY...University...of Alberta.....

DEGREE FOR WHICH THESIS WAS PRESENTED...Ph.D.............

YEAR THIS DEGREE GRANTED..1971..........................

Permission is hereby granted to THE NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

(Signed)..P.A. Runquist...

PERMANENT ADDRESS:
#1907-10883 Sask. Drive
..Edmonton.......
....

DATED April..26....1971

NL-91 (10-68)

THE UNIVERSITY OF ALBERTA

ORGANIZATIONAL FACTORS IN PAIRED-ASSOCIATE LEARNING

by

© PEGGY ANNE RUNQUIST

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

SPRING, 1971

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Organizational Factors in Paired-Associate Learning", submitted by Peggy Anne Runquist in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

....
Supervisor

....

....

....

....
External Examiner

Date April 16, 1971

....

Abstract

This experiment investigated two aspects of response-term grouping in paired-associate learning: (1) the nature of the relationship among the responses that are grouped together, and (2) the effect of this grouping on acquisition performance.

Eight groups of 24 Ss learned two 12-item categorized paired-associate lists to a criterion of 12/12 correct on a single trial. After a 2-min. interval filled with arithmetic problems, Ss were given two 1.5-min. free recall tests in which they were asked to recall as many of the response terms as possible. A cued-recall test for the stimulus terms followed immediately. A ninth group learned only the second list, then followed the same recall procedure.

For all conditions, the paired-associate lists consisted of three instances of four different conceptual categories as stimulus terms and 12 unrelated adjectives as response terms. In the transfer conditions, the stimuli in the two lists were either identical, different instances of the same conceptual categories, or instances of completely different conceptual categories. The response terms in the two lists were always identical but were re-paired in the second list either with stimuli from the same first-list category (within-category re-pairing), with stimuli from a different first-list category (between-category re-pairing) or with stimuli from three different first-list categories (across-category re-pairing).

The three re-pairing schemes were designed to disrupt different types of relationships that could exist among the response terms that were grouped together during first-list learning and hence to affect the amount of grouping of the same responses in second-list learning. Within-category re-pairing was not expected to disrupt any associations among the responses; between-category re-pairing was expected to disrupt associations mediated by stimulus categories and associations mediated by specific stimuli; and across-category re-pairing to disrupt both types of mediated associations that might have developed among the response terms themselves.

The results of the experiment showed an increase in second-list grouping only when response terms were re-paired with stimuli from the same conceptual categories as in the first list, regardless of whether the specific stimuli in the two lists consisted of the same or new instances of these categories. Such a result indicates the importance of specific category-mediated associations among the grouped responses. No evidence was found that the responses within a category became directly associated with one another. The second major finding was that grouping of the response terms into sets did not facilitate the acquisition of specific associations in the second list. It was suggested, however, that this failure to find facilitation may be a result of the potency of the category name as a mediator.

Acknowledgements

I would like to thank the members of my supervisory committee, Dr. W. Rozeboom, Dr. A. Dobbs and Dr. J. Sampson, for their advice and critical comments on the writing of this thesis.

Thanks are also due to Marva Blackmore who critically typed many drafts of the thesis and to Willie who nagged me until it was done.

The research was supported in part by Research Grant APA-88 from the National Research Council of Canada to Dr. W. N. Runquist.

Table of Contents

Table of Contents (continued):

List of Tables

Chapter 1

Introduction

The purpose of this research was to investigate the nature of an organizational process in paired-associative learning whereby stimuli and responses from a number of pairs are grouped into sets on the basis of interrelationships among the stimulus items. It has been argued (Battig, 1966, 1968; Runquist, 1966, 1968b) that not only does such a process exist but that it is a primary means by which interpair interference can be reduced.

The existence of such a process is interesting for two reasons. Organization refers to the processes whereby discrete items are grouped together or unitized on the basis of perceived relationships among the items and thence treated as single units. This view has largely come from research in free recall learning, which, since the early demonstrations of clustering, has been the standard task for the study of organization in memory. In this task, Ss are presented a series of items, one at a time, and asked simply to recall as many of these items as possible. Since there are no restrictions placed on the order in which items are to be recalled, the order of output is assumed to reflect the order in which S thinks of the items. Consequently, the recall of items in an order different from the one in which they were presented, either as clusters of related words (clustering) or as consistent sequences of two or more words over several trials (subjective organization) is interpreted as reflecting S's tendency to group or organize the material into larger units on the basis of pre-existing conceptual relationships among the items to be recalled (Tulving, 1968). Organization as thus defined has been demonstrated over a wide variety of stimulus materials and has almost invariably been accompanied by an increase in the amount recalled (Shuell, 1969; Tulving, 1968). The unitization hypothesis (Miller, 1956), with elaborations to account for variations in the free recall paradigm (e.g., Cohen, 1966; Mandler, 1967; Tulving, 1964), has been the major explanation of both the occurrence of organization and its effect on the amount recalled. Miller (1956) suggested that since immediate memory is limited by the number of items to be recalled rather than the amount of information contained within the items, recall could be increased by recoding the input into chunks of items, labelling these chunks in some way that symbolizes the information contained within them, then remembering these chunks rather than specific items. Recall of specific items could then proceed via the recall and subsequent decoding of these recoded units. Since there are fewer chunks than discrete items, this grouping would lead to an increase in the amount recalled. Furthermore, since item recall occurs via the decoding of a chunk, the items within that chunk would be recalled together.

In paired-associate (PA) learning on the other hand, there are restrictions in the order in which items are recalled. In this task, several pairs of verbal items are presented to the S with the instructions that he is to learn to associate the two members of each pair so that given one member (the stimulus term), he will be able to recall the item associated with it (the response term). Although some kind of unitization may occur within a pair where stimulus and response terms are related, e.g., redintegrative learning (Horowitz and Prytulak, 1969), or within low meaningful response terms, e.g., strings of letters or numbers (Johnson, 1970), and lead to better performance, unitization of items from several pairs should interfere with the learning and recall of specific associations within pairs.

The existence of such an interpair grouping process also poses some difficulties for the traditional view of PA learning as the acquisition of several relatively isolated associations. Although Ss learn several pairs simultaneously, theoretical accounts of PA learning have assumed that the establishment of associative connections within a single pair is relatively independent of the other pairs within the list (Battig, 1968). Furthermore, although similarity and other interrelationships among the stimulus terms are known to affect PA learning, the usual conception is that such interrelationships must be overcome if learning is to proceed.

According to this view, stimulus similarity results in competition between correct and generalized response tendencies (the latter resulting from stimulus generalization), which retards the rate at which associations are formed but does not interact with this process in any other way. As learning proceeds, these error tendencies are extinguished as a result of nonreinforcement, and the correct responses strengthened, so that at the end of learning each pair represents a single associative unit which is independent of all the other pairs in the list. Evidence that during learning, interpair similarity relationships are maintained or even strengthened rather than eliminated, and facilitate rather than depress acquisition performance, suggests the inadequacy of such an account of PA learning (Battig, 1966; Runquist, 1968b).

Although few contemporary psychologists would deny that there is more involved in PA learning than the formation of stimulus-response associations within individual pairs of items, the acceptance of an additional process in acquisition or recall demands that it be distinguished both conceptually and operationally from processes that have already been identified. With respect to the first of these criteria, interpair grouping does not fare too badly. For the most part, this process has been conceptualized as a stage, prior to the learning of specific associations, during which S learns to group or categorize into sets those response terms which are possible correct responses to corresponding sets of interrelated stimulus terms (Runquist, 1970). The result of this grouping is a "hierarchical organizational structure, whereby the overall list is subcategorized into subsets or groups of pairs with some common property or interrelationship among them (Battig, 1968, p. 163)."

Conceptually, the process of grouping can be distinguished from the process of generalization in that the former involves categorization of pairs into sets sharing some common characteristic whereas the latter involves a failure to discriminate between error tendencies and the correct S-R associations. Whereas generalization would lead to competition among correct and error tendencies and hence retard the rate of acquisition, grouping, by restricting the number of items that interfere with learning a single pair to those belonging to the same subset, should facilitate acquisition performance.

Even if one accepts the distinction between grouping and generalization at the conceptual level, there remains the necessity of specifying how these two processes can be distinguished at the operational level. It is with respect to this criterion that the acceptance of grouping as a component process of PA learning is least justified. The primary difficulty with the evidence for grouping, in fact, is that the kinds of behaviors from which grouping has been inferred are also interpretable on the basis of other processes. The occurrence of generalized errors in learning PA lists in which the stimulus terms can be categorized into sets of similar stimuli has, for example, been used to indicate the existence of both grouping and generalization. Usually a generalized or within-category error is said to occur when the S responds to a given stimulus item with the response term that is appropriate to another stimulus belonging to the same stimulus category. The occurrence of this type of error could therefore reflect equally well either a failure to discriminate among responses paired with similar stimuli (generalization) or the delimitation of a set of responses as appropriate to a corresponding set of interrelated stimuli (grouping). Other criteria have also been used to distinguish grouping from non-grouping behavior, but like the generalized error data, they, too, when taken alone, need not implicate grouping to the exclusion of other processes. Further evidence for grouping has come largely from investigations of quite different problems, e.g., the effects of degree of learning and intralist stimulus similarity on acquisition performance.

If the acceptance of grouping as a valid subprocess in PA learning is to be contingent upon being able to specify operations by which it may be uniquely assessed, it would be useful to review the experiments from which grouping has been inferred, and to examine alternate interpretations of their results.

Grouping on the basis of degree of learning. The classification of pairs as learned and unlearned in experiments involving manipulations of degree of learning has been suggested as one type of interpair grouping in PA learning. This process is said to occur in intermediate stages of acquisition when only some of the pairs have been given correctly, and consists of the classification of pairs into sets of learned and unlearned items. Presumably, the effect of this classification is to restrict the interference in learning a single pair to other pairs in the same stage of learning (Battig, 1966, 1968). Evidence for this type of grouping has come largely from studies in which the ease of differentiating among pairs on the basis of degree of learning was manipulated by either transferring pairs that had already been learned to a new list, or by modifying some aspect of pairs in a single list once they had been learned to some criterion.

Brown (1964) found that when first-list pairs were substituted for half the new pairs in the second-list, performance on the new pairs could either be facilitated or interfered with depending upon whether the degree of learning on the old pairs was high or low, respectively. Presumably, the lower the degree of learning on the first-list pairs, the harder it would be to classify the second-list pairs as new or old and to learn them as sets. Support for this explanation was found in a second experiment (Brown and Battig, 1966) in which first-list pairs of varying degrees of learning were added to a constant number of second-list pairs. The results again showed that there was less interference in learning the second-list pairs when the first-list pairs had been learned to a high criterion rather than a low criterion.

It has also been shown that any cue which will facilitate the discrimination of learned and unlearned pairs, and, hence, the grouping of pairs into these categories, will also lead to better performance. Brown and Battig (1962, Exp. I) have shown that fewer errors are made in PA learning if pairs which have been given correctly are held constant in the same serial position in which the first correct response was made and the serial position of the incorrect pairs varied from trial to trial. A similar finding was shown in an experiment by Battig, Brown, and Nelson (1963, Expt. V) in which pairs were held in a constant serial position until the first correct response, then varied on successive trials. Both this contingent condition and a constant condition in which pairs were always presented in the same serial order, were superior to a condition in which the order of items always varied from trial to trial. Another study (Schild and Battig, 1966) showed that if a shift from bidirectional presentations (each item serving as both a stimulus and a response term) to a constant unidirectional pairing is made contingent upon responding to the pair correctly, performance on the remaining pairs will be facilitated. Similarly, Brown, Battig, and Pearlstein (1965) showed that when new letters were added to a single stimulus term following varying numbers of correct responses on each pair, fewer errors were made in reaching criterion than when the stimuli contained all letters throughout. In all these studies, the cues of serial position, bidirectionality, and added letters presumably served to aid the differentiation of pairs into learned or unlearned categories and thereby restrict the interference from all pairs in the list to just those in the same state.

The main problem with these studies is that degree of learning is confounded with item difficulty. Because degree of learning was determined by how quickly a pair was learned, pairs that were learned rapidly must have been easier to learn than pairs which were acquired more slowly. Since whatever characteristics would make a pair easy to learn in the first place should also make them more discriminable from other pairs, easy (i.e., discriminable) pairs would be expected to interfere less than pairs that are difficult to discriminate from one another.

Runquist (1965, 1967), however, was able to manipulate degree of learning independent of item difficulty and still find grouping effects. His lists were composed of two types of pairs: critical items that appeared only once per trial and interference items that appeared either once or three times per trial. Critical items were learned much faster when interference items occurred three times as often as the critical items than when they occurred equally often. Runquist suggested that Ss can group pairs on the basis of task variables, in this case, degree of learning, amount of practice, or frequency of occurrence. Further support for this interpretation was based on the finding that the facilitation of the critical items under asymmetrical presentation was accompanied by an increase in the proportion of generalized or confusion errors between items of the same low frequency of presentation. The occurrence of generalized errors in this case was interpreted as representing the categorization of pairs into sets--those that occurred three times and those that occurred once per trial--with interference being restricted to items within the same category. Grouping on the basis of variable frequency of appearance has been found with both high intralist stimulus similarity lists (Runquist, 1967) and low intralist stimulus similarity lists (Runquist, 1965, 1967). Runquist (1968a) has also provided evidence that the facilitation of performance with variable frequency of appearance is permanent, the facilitation being maintained when asymmetry was removed (all items appearing only once per trial) or reversed (critical items appearing more frequently and interference items less frequently). Reversing the asymmetry of items is particularly relevant to considerations of grouping since categorization based on the verbal label "more frequently" or "less frequently" should be disrupted when the switch is made (Runquist, 1968a, p. 1060). The within-category or generalized error data did, in fact, indicate some disruption after switching, with the proportion of these errors decreasing on later trials. Because the correct responses were not affected by the switch, however, Runquist suggested that grouping information is utilized in learning and recall only when direct associations are poorly learned, being of little value when those associations are well learned.

Whereas the asymmetry of frequency presentations could have produced grouping, the data in these studies need not be interpreted in this way. Under the asymmetry conditions, interference items appeared more frequently and were, as one would expect, learned faster than the critical items. Once learned, each interference pair would be easily discriminable from the critical pairs and, hence, would not be expected to interfere with them. The generalized errors would then simply reflect interference among the unlearned pairs, each interference pair being removed as a potential source of interference, one at a time, as it was learned, rather than representing the categorization of pairs into groups on the basis of differential frequency of appearance. Strong support for the grouping interpretation, on the other hand, would require evidence that the interference items had indeed been classified together as a group, as would be the case if a high proportion of generalized errors had occurred among the frequently appearing items as well as among the less frequently appearing errors. Error data for these items, however, were not presented.

Grouping on the basis of stimulus similarity. Grouping on the basis of stimulus similarity has been conceptualized, for the most part, as the categorization of response terms into sets on the basis of similarity among the stimulus terms with which they are paired. This process occurs early in learning and represents a stage of learning in which S has learned to categorize responses into groups appropriate for a set of similar stimuli but has not yet learned the specific S-R associations. It has been suggested that this grouping can facilitate performance either by reducing the amount of interference in learning a single pair or by limiting the number of response alternatives for any stimulus term.

Underwood, Ekstrand, and Keppel (1965) had Ss learn PA lists composed of nouns from varying numbers of conceptual categories as stimulus terms and double letters as response terms. They found that errors in learning tended to be restricted to confusions among responses paired with nouns from the same categories much more than would be expected by chance and that rate of learning was directly related to the number of categories used. The error data were interpreted as evidence for "S-R limitation" (or grouping) on the basis of stimulus category membership, the effect of this process being to limit the number of possible responses which could be correct for any given stimulus, and, thereby, to reduce interference.

Runquist (1966) found evidence for grouping in two studies in which list length and formal stimulus similarity were varied factorially. The S learned either high or low similarity lists consisting of either 8 or 16 S-R pairs. The stimulus terms for the high similarity lists consisted of either 4 or 8 pairs of similar CVC trigrams, while for the low similarity lists, the trigrams were all dissimilar. When unrelated nouns were used as response terms in the first experiment, pairwise stimulus similarity had a detrimental effect on the acquisition of short lists but not with the long lists. Since over half the errors, after the first correct response, were generalized errors between similar pairs, Runquist suggested that the failure to find an effect of similarity with the longer list was a result of grouping. Grouping, in this study, was conceptualized as stage of learning in which S had learned which two responses were paired with each set of similar stimuli, but did not know as yet which response was paired with which stimulus. It was suggested that grouping could provide an advantage in learning a structured high similarity list in either of two ways: first, it could reduce the interference from other pairs in the list, and, secondly, it could enable S to learn the response to one stimulus and then give the correct response to the other stimulus because it is the only alternative; i.e., he could learn partially by elimination. Further support for this interpretation was gained in the second task when the necessity of response learning was eliminated by having the S associate the stimuli with one of two push-buttons. Under these conditions there was evidence that the long high similarity lists were learned even faster than the long low similarity lists.

Joinson and Runquist (1968) compared the retention of lists varying in formal intralist similarity after one week following learning to each of three performance criteria. Stimulus terms were again pairwise similar trigrams paired with unrelated nouns. The major finding was that, at medium degree of learning, the high similarity groups forgot fewer items than the low similarity groups and, at the same time, made more generalized errors during recall than any of the other groups. No differences due to similarity were found at low or high performance citeria. The facilitation of long term retention was attributed to the retention of grouping information which would enable the S to recall which two responses were paired with each pair of similar stimuli even though the specific associations had been forgotten. Recall at the low and high criteria was not affected, presumably, because at the low criterion there was not sufficient practice for group membership to be learned, and at the high criterion, there was so little associative forgetting that grouping had only a slight effect.

Inadequacies of these data as evidence of grouping. The basic evidence for grouping in these studies was the preponderance of generalized errors together with a facilitation in performance. Neither factor by itself, however, constitutes sufficient evidence for grouping. The preponderance of generalized errors in learning a list with high stimulus similarity, for example, can be explained without postulating a grouping stage. According to the traditional account of stimulus similarity, associative strength would be built up via stimulus generalization between a response and a stimulus similar to the one with which it is paired, but not between a response and a dissimilar stimulus. Since associative strength determines whether a response will be given, overt errors would be expected to be predominately generalized errors. According to this view, however, the occurrence of a large proportion of these errors ought to be accompanied by a depression rather than facilitation in performance. That is, generalized errors reflect competition between the correct and generalized error tendencies, with the error tendency, being stronger, giving rise to an overt response. Learning, in this case, would involve both the gradual weakening of generalized error tendencies by non-reinforcement of incorrect responses and the strengthening of the correct response tendencies by confirmation of the correct responses. In no case could a competing response notion of grouping lead to a prediction of equal or better performance on a groupable list (high similarity list) than on a low similarity list; such grouping should always lead to more interference for the former. (This argument would also hold true for the degree of learning studies, although there remains the difficulty in specifying how the similarity between stimulus terms developed.) The evidence with respect to facilitation, however, is not very strong. It consists, in the case of stimulus similarity, essentially of failing to find a difference favoring the learning of a low similarity list over the learning of a high similarity categorizable list, the latter (by virtue of stimulus generalization) being expected to produce much more interference than the former. There is, however, some question about the appropriateness of a low similarity control list. Even if grouping restricted the interference in learning each pair to items in the same group, there should still be more interference than in a low similarity list. A noncategorizable high similarity list in which each stimulus is equally similar to every other stimulus term, on the other hand, is just as inappropriate since categorizing the stimuli effectively reduces the overall similarity in a list, each stimulus being similar only to some of the other stimuli rather than to all.

In effect, then, since generalized errors could reflect processes other than the categorization of the responses, and since facilitation in performance due to grouping has not been clearly demonstrated, there is little evidence in these studies to demand the existence of a grouping process.

Two studies have provided somewhat more promising operations to assess the existence and effects of grouping on the basis of stimulus similarity in PA learning. Runquist (1968b), employing a transfer task, had <u>Ss</u> learn two 8-item PA lists in succession, each consisting of the same two sets of four similar stimuli and the same unrelated response terms but paired differently in the second list. Re-pairing in the second list was such that responses for half the pairs in each set of stimuli were reversed (reversal re-pairing) and half were paired with two items from the other set of stimuli (non-reversal re-pairing). The rationale was that if the S did group the responses on the basis of the two sets of similar stimuli, re-pairing a response paired with a stimulus from the same set should result in faster learning than re-pairing it with a stimulus from a different set. The results supported this hypothesis: reversal re-paired items were learned faster than the non-reversal items. Furthermore, it was suggested on the basis of overt error data that the superiority of the reversal re-pairing resulted from positive transfer in the reversal pairs rather than negative transfer in regrouping the non-reversal items, there being a larger proportion of list-2 generalized errors for the non-reversal items. These data suggest, then, that the existence and effect of grouping on performance can be detected in a transfer situation.

The second study (Runquist, 1970) incorporated a free recall trial following PA learning to assess the occurrence of grouping during acquisition. In this experiment, Ss learned PA lists consisting of either categorized or noncategorized stimulus terms paired with the same unrelated response terms and then recalled the stimulus and response terms on separate free recall trials. The hypothesis was that, if the response terms were categorized into groups corresponding to the sets of similar stimuli, these response groups, as well as the related stimuli, ought to be manifested as clusters on the free recall tests. The results confirmed this hypothesis: groups learning the categorized lists showed significantly above chance clustering of both the stimulus and the response terms when evaluated with respect to the categories to which the paired stimuli belonged; the noncategorized list groups, on the other hand, showed no tendency to cluster these same responses. Furthermore, as expected from previous studies, generalized or within-category errors accounted for more than half the total overt errors made by the categorized list conditions.

While the occurrence of grouping in PA learning would demand that the response terms cluster in free recall, the converse need not be true. If, for example, S recalled the responses by recalling each of the stimulus terms and giving the response term associated with it, response term clustering could simply reflect the organization of the stimulus terms by category membership at the time of recall rather than the categorization of the responses into sets during acquisition. According to this interpretation, the recall of one stimulus (S_{A1}) from a given category (A) would lead to the recall of both the response term that was paired with it (R_1)

and another stimulus term (S_{A_2}) from the same category. Similarly, S_{A_2} would lead to the recall of R_2 and S_{A_3}, and so on. Clustering of the responses would thus be observed to the extent that the similar stimuli were recalled together. Even this type of explanation would not be necessary, however, if it were assumed that generalized error tendencies still existed in some strength at the end of acquisition. In this case, recall of any one stimulus term might be expected to lead to the recall of all the response terms paired with similar stimuli simply by generalization.

If, on the other hand, the clustering did reflect the grouping of responses into sets during learning, there are at least two ways in which response clustering could have occurred. First, if S had categorized the responses into sets of appropriate responses for a given category of stimuli, e.g., animals, in addition to learning the specific associations to stimuli within that category, recall of the category name alone would be sufficient for response terms to be recalled in clusters. Furthermore, if, in addition, the response terms forming a common set were to become directly associated with one another by virtue of having been categorized together during learning, the recall of any one response should be sufficient to lead to the recall of the other responses in the same set, and clustering would result.

To the extent, then, that free recall of the responses occurs via direct associative connections between specific stimuli and responses, response term clustering need not implicate the occurrence of grouping in PA learning. If, on the other hand, responses were recalled either through stimulus category name recall or through recall of other responses, response-term clustering could indeed reflect the categorization of responses into sets during learning. In the absence of any conclusive evidence about the mechanism by which clustering of the responses occurs, both alternative explanations are viable.

The research to be reported here was designed, in part, to provide evidence bearing on this issue.

The rationale for the design. In this experiment, Ss learned two categorized PA lists, conforming basically to the A-B, A-B_r transfer paradigm, then recalled the response terms on a free recall test following the learning of the second list. The rationale was that if the response terms were categorized into sets on the basis of stimulus categories with which they were associated in the first list, the amount of categorization in the second list should increase if the response groupings are the same (compatible) but decrease if they are different (incompatible) from those in the first list, as reflected by response term clustering during free recall.

Compatibility of the response groupings in the two lists was determined by the way in which the response terms that were paired with a given category of stimuli in the first list were re-paired in the second list. Responses were re-paired with other stimuli from the same category, with the stimuli from a single different category, or with stimuli from each of several different categories. If response grouping reflected the delimitation of a set of responses to a category name, re-pairing within that stimulus category is the only way in which compatible groupings in the two lists could occur. On the other hand, if response terms forming a set become directly associated with one another, compatible groupings would result from re-pairing the responses either with stimuli from a different category or with stimuli from within the same category. The third type, re-pairing responses with stimuli from several different categories, would produce response groupings that are incompatible with the first list from either view. If the mechanism for response clustering in free recall is recall via the stimulus category name only, re-pairing within a category would consistently lead to response clustering; if it occurs via direct associations among responses, re-pairing with stimuli from the same category or with stimuli from a different, intact category should both produce clustering, whereas, re-pairing with stimuli from several different categories should not.

A second purpose of the experiment was to assess the effect of grouping on performance. If categorization of responses into sets does facilitate the learning of specific associations, one would expect more rapid acquisition on the second list if the response groupings in the two lists were the same than if they were different. A test of this hypothesis with the A-B, $A-B_r$ paradigm, would be difficult to interpret, however, since specific negative transfer from first-list associations may be stronger than any positive transfer from compatible response groupings. If, on the other hand, one were to use similar rather than identical stimulus terms in the second list, such as different instances of the same conceptual category, the specific interference from first-list associations would be reduced and the effect of grouping could be assessed. The design of this experiment thus included conditions which varied in degree of interlist stimulus similarity as well as type of re-pairing in the second list.

Chapter 2

Method

Design. There were eight basic transfer conditions in the experiment. Labelled on the basis of the relation of the first list to the second (A-B) list, the conditions were: $A\text{-}B_{WC}$, $A\text{-}B_{BC}$, $A\text{-}B_{AC}$, $A'\text{-}B_{WC}$, $A'\text{-}B_{BC}$, $A'\text{-}B_{AC}$, $C\text{-}B_{BC}$, and $C\text{-}B_{AC}$. The first letter indicates the similarity of the stimulus terms in the two lists: stimuli were identical (A), different instances of the same category (A'), or instances of different conceptual categories (C). The remaining letters indicate three different S-R re-pairing schemes: responses were re-paired with stimuli from the same category (Within-Category or WC re-pairing), with the stimuli from a different category (Between-Category or BC re-pairing), or with stimuli from three different categories (Across-Category or AC re-pairing). In addition, there was one control group (A-B) that learned only the second list. The types of S-R pairings for the first and second lists for each condition are shown in Table 1.

With two sets of lists used under each of the conditions, there were a total of 18 groups.

Materials. Two identically constructed sets of sixteen 12-item paired-associate lists were used as materials (Appendix A). All lists consisted of three nouns from each of four different

Table 1

S-R Pairings

List-1 Pairs								List-2 Pairs
$A-B_{WC}$	$A-B_{BC}$	$A-B_{AC}$	$A'-B_{WC}$	$A'-B_{BC}$	$A'-B_{AC}$	$C-B_{BC}$	$C-B_{AC}$	$A-B$
A_1-R_2	A_1-R_7	A_1-R_4	A_4-R_2	A_4-R_7	A_4-R_4	E_1-R_7	E_1-R_4	A_1-R_1
A_2-R_3	A_2-R_8	A_2-R_8	A_5-R_3	A_5-R_8	A_5-R_8	E_2-R_8	E_2-R_8	A_2-R_2
A_3-R_1	A_3-R_9	A_3-R_{12}	A_6-R_1	A_6-R_9	A_6-R_{12}	E_3-R_9	E_3-R_{12}	A_3-R_3
B_1-R_6	B_1-R_1	B_1-R_{11}	B_4-R_6	B_4-R_1	B_4-R_{11}	F_1-R_1	F_1-R_{11}	B_1-R_4
B_2-R_4	B_2-R_2	B_2-R_2	B_5-R_4	B_5-R_2	B_5-R_2	F_2-R_2	F_2-R_2	B_2-R_5
B_3-R_5	B_3-R_3	B_3-R_9	B_6-R_5	B_6-R_3	B_6-R_9	F_3-R_3	F_3-R_9	B_3-R_6
C_1-R_9	C_1-R_{10}	C_1-R_{10}	C_4-R_9	C_4-R_{10}	C_4-R_{10}	G_1-R_{10}	G_1-R_{10}	C_1-R_7
C_2-R_7	C_2-R_{11}	C_2-R_6	C_5-R_7	C_5-R_{11}	C_5-R_6	G_2-R_{11}	G_2-R_6	C_2-R_8
C_3-R_8	C_3-R_{12}	C_3-R_1	C_6-R_8	C_6-R_{12}	C_6-R_1	G_3-R_{12}	G_3-R_1	C_3-R_9
D_1-R_{11}	D_1-R_4	D_1-R_7	D_4-R_{11}	D_4-R_4	D_4-R_7	H_1-R_4	H_1-R_7	D_1-R_{10}
D_2-R_{12}	D_2-R_5	D_2-R_5	D_5-R_{12}	D_5-R_5	D_5-R_5	H_2-R_5	H_2-R_5	D_2-R_{11}
D_3-R_{10}	D_3-R_6	D_3-R_3	D_6-R_{10}	D_6-R_6	D_6-R_3	H_3-R_6	H_3-R_3	D_3-R_{12}

conceptual categories as stimulus terms and 12 unrelated adjectives as response terms. The nouns were selected from the 12 most frequent responses to category names as determined by the Battig and Montague (1969) norms, and were assigned to lists in such a way that mean taxonomic frequency was equated across lists. The adjectives were two- or three-syllable words of A or AA frequency of occurrence according to the Thorndike-Lorge (1944) G count.

Each set of 16 lists was constructed from four sets of stimulus terms paired in four ways with the same set of 12 adjectives. Two of the sets of stimuli contained nouns representing different instances of one group of four categories and two contained nouns representing a different four categories. All four sets of nouns were used as stimulus terms for first and second lists equally often under all conditions. S-R pairings for first and second lists, however, were not counterbalanced. Instead, a basic pairing for the second (A-B) list was selected for each set of stimuli so that the same three adjectives were paired with nouns from the same categories in the two similar sets. The remaining S-R pairings, used in the first lists, were generated from the A-B lists by re-pairing these groups of three adjectives within, between, or across categories according to the scheme presented in Table 1. In constructing the specific pairings, care was taken to avoid obvious associations between the stimulus and response members and among the responses paired with stimuli from the same category.

Subjects. The Ss were 216 students from an introductory psychology course, serving in order to obtain course credit. The Ss were assigned to the nine conditions in order of their appearance at the laboratory according to a prearranged scheme randomizing the order of conditions within blocks containing the four sets of stimuli used under each condition. The two sets of lists were run consecutively with 12 Ss in each replication. An additional 28 Ss were discarded from the experiment. Of these losses, four were due to apparatus failures, 14 to E errors, and 10 to failure of S to finish the experiment within the time allotted.

Apparatus and Procedure. All learning was by the anticipation method at a 2:2-sec. rate with a 4-sec. intertrial interval. Words were presented on two IEE Series 80,000 digital display units mounted immediately adjacent to one another on the wall of a booth in which S was seated. Stimuli were programmed by paper tape which advanced every 4-sec., paced by a motor-driven cam. The stimulus term appeared on the left-hand unit as soon as the tape advanced, while the response term, appearing on the right-hand unit, was delayed for 2-sec. by a Hunter electronic interval timer. Five orders of item presentation were used for all lists. Selection of the particular orders attempted to minimize possible systematic effects of presentation order during learning on the order of output at recall. The criteria used in determining these orders were:

(a) no two items should be immediately followed by another item belonging to the same category; (b) no two items should follow one another more than once; (c) no item should appear in the same serial position more than once; and (d) the last item in one order should not be the first in the succeeding order. Starting order for first-list learning was determined systematically prior to the experiment, while, for second-list learning, it was the one which would have occurred next if List-1 acquisition had continued. For both lists, the Ss started to anticipate the response terms on the second trial and continued until they reached a criterion of 12/12 correct. The interval between lists was approximately 2-min.

Immediately following List-2 learning, Ss were given a series of arithmetic problems to solve within a 2-min. period. This filled interval following learning was designed to prevent rehearsal of items and to minimize recency effects on the two free-recall tests that followed.

On the free recall trials, Ss were instructed to write down on a test sheet as many of the 12 adjectives as they could recall in the order in which they recalled them. The Ss were allowed 2-min. for each recall test, with 1.5-min. between tests. The Ss continued to work on the arithmetic problems during this interval.

Following the free-recall test, Ss were given a sheet of paper which showed the 12 adjectives typed down one side in a random order with two blanks opposite each adjective. The Ss were instructed to write down the two nouns with which the adjective had been paired during learning and to indicate whether that noun had appeared with the adjectives in the first-list or second-list.

Chapter 3

Results

All means reported in this section are based on the combined data from the two sets of lists used under each condition, since, with one exception (number recalled on the second free recall trial), there were no interactions between lists and conditions. Statistical comparisons between groups were made by analyses of variance and Duncan's Multiple Range tests. Only the overall *F* values and results of comparisons of theoretical interest between groups will be reported in the text. Complete summaries of the results of Duncan's tests for second list analyses (tested for significance at the .05 level for all analyses) are presented in Appendix B.

First-list learning. Table 2 presents the mean number of trials to criterion (TTC) and the mean number of errors to the first correct response for each pair of items (ETFC) for each of the eight transfer conditions. Neither measure showed any significant differences between conditions in the first-list learning either by analysis of variance (Table 3, $F(7,176) < 1$ for both measures) or by Duncan's test. Any differences in performance in later tasks, therefore, cannot be attributed to differences in degree of first-list learning.

Table 2

Mean Number of Trials to Criterion and Errors to the First Correct Response in List-1 Learning

	$A-B_{WC}$	$A-B_{BC}$	$A-B_{AC}$	$A'-B_{WC}$	$A'-B_{BC}$	$A'-B_{AC}$	$C-B_{BC}$	$C-B_{AC}$
TTC	9.96	11.17	9.46	10.83	9.88	9.71	9.42	8.46
ETFC	28.17	29.96	27.25	29.91	30.25	28.83	32.04	25.00

The percentage of overt errors that represented confusions among response terms paired with stimuli from the same category was used as an index of response grouping in list-1 acquisition. On the basis of chance, these generalized errors ought to account for only 18% of the total overt errors. Table 4 shows the obtained mean percent generalized errors (%GE), and the number of <u>S</u>s who made at least one overt error (n) for each of the eight conditions. As with previous studies employing categorized lists, within-category errors consistently accounted for a greater proportion of total overt errors than would be expected simply on the basis of chance.

Table 3

Summaries of Analyses of Variance for

List-I Performance Measures

Trials to Criterion

Source	SS	df	MS	F	p
Conditions (C)	120.25	7	17.18	.30	>.05
Replications (R)	29.30	1	29.30	1.17	>.05
C x R	166.25	7	23.75	.95	>.05
Error	4419.42	176	25.11	-	-
Total	4735.22	191	-	-	-

Errors to the First Correct Response

Source	SS	df	MS	F	p
Conditions (C)	775.56	7	110.79	.30	>.05
Replications (R)	1045.33	1	1045.33	2.78	>.05
C x R	2353.92	7	336.27	.90	>.05
Error	66146.17	176	375.83	-	-
Total	70320.98	191	-	-	-

Table 4

Mean Percent Generalized Errors in List-1 Learning

	A-B_{WC}	A-B_{BC}	A-B_{AC}	A'-B_{WC}	A'-B_{BC}	A'-B_{AC}	C-B_{BC}	C-B_{AC}
Mean %G.E.	43.1	58.1	43.8	59.0	49.9	52.6	60.1	62.6
n	24	24	22	22	22	20	22	23

Transfer. Performance on List-2 was examined at three stages of practice for transfer effects. Early transfer was assessed by the number of errors on the first test trial (the second anticipation trial), transfer at a medium degree of learning by the number of errors to the first correct response for each pair, and late transfer by the number of errors after the first correct response for each pair. Overall transfer effects were assessed by the number of trials to criterion. The means for these data are presented in Table 5.

Analyses of variance (Table 6) revealed a significant difference between conditions for each measure, $F(8,198) = 5.081$ for errors on trial 1, 4.810 for errors to the first correct response, 9.934 for errors after the first correct response, and 7.568 for trials to criterion, $p < .005$ for all measures. The interesting

Table 5

Mean Errors During Stages of List-2 Learning and Mean Trials to Criterion

	A-B	$A\text{-}B_{WC}$	$A\text{-}B_{BC}$	$A\text{-}B_{AC}$	$A'\text{-}B_{WC}$	$A'\text{-}B_{BC}$	$A'\text{-}B_{AC}$	$C\text{-}B_{BC}$	$C\text{-}B_{AC}$
First Test Trial	9.00	9.00	9.25	9.33	6.96	8.08	8.25	6.33	6.79
To First Correct Response	35.42	31.00	35.00	31.33	21.54	22.46	23.83	16.38	16.46
After First Correct Response	14.79	23.71	21.04	15.67	8.46	6.83	6.83	4.92	4.25
Trials to Criterion	11.70	12.54	12.17	11.00	7.88	8.46	8.33	6.67	5.96

Table 6

Summaries of Analyses of Variance
for List 2 Performance Measures

Errors on Trial 1

Source	SS	df	MS	F	p
Conditions (C)	254.91	8	31.86	5.09	<.01
Replications (R)	6.00	1	6.00	< 1	-
C x R	78.59	8	9.82	1.57	>.05
Error	1241.83	198	6.27	-	-
Total	1581.33	215	-	-	-

Errors to the First Correct Response

Source	SS	df	MS	F	p
Conditions (C)	10653.09	8	1331.64	4.81	<.01
Replications (R)	880.07	1	880.07	3.18	>.05
C x R	2361.26	8	295.16	1.07	>.05
Error	54814.67	198	276.84	-	-
Total	68709.09	215	-	-	-

Errors after the First Correct Response

Source	SS	df	MS	F	p
Conditions (C)	9983.83	8	1247.98	9.93	<.01
Replications (R)	28.17	1	28.17	< 1	-
C x R	942.17	8	117.77	< 1	-
Error	24873.83	198	125.63	-	-
Total	35828.00	215	-	-	-

Trials to Criterion

Source	SS	df	MS	F	p
Conditions (C)	1182.50	8	147.81	7.57	<.01
Replications (R)	20.17	1	20.17	1.03	-
C x R	105.00	8	13.13	< 1	-
Error	3866.83	198	19.53	-	-
Total	5174.50	215	-	-	-

comparisons at each stage, however, are between groups with the same interlist stimulus similarity but different S-R pairings, between groups with the different interlist stimulus similarity for each type of S-R pairing, and between the transfer groups and the A-B control.

For the most part, type of S-R re-pairing had little effect on performance at any stage of List-2 learning. Of all the comparisons made, only in errors after the first correct response were any significant differences revealed, i.e., the $A\text{-}B_{AC}$ made significantly fewer errors than $A\text{-}B_{WC}$. Even so, it is interesting to note that when group means are ranked from poorest to best performance, there is considerable consistency revealed both within a stage and in the changes between stages, regardless of the stimulus similarity. On trial 1, WC re-pairing was easier than BC re-pairing, which, in turn, was easier than AC re-pairing. WC re-pairing was still easiest in terms of errors to the first correct, but BC and AC re-pairings for the A stimuli were reversed. After the correct response, the ordering of groups was completely reversed from that on Trial 1 with AC being easiest and WC, the most difficult. Overall performance reflects this last ordering with the exception of the A' re-pairing in which WC re-pairing led to better performance than BC or AC re-pairings.

The rank ordering of means for groups differing in interlist stimulus similarity remained the same at all stages of learning and for all types of re-pairings. Performance was best with stimuli representing completely different categories, second with new stimuli from the same categories, and worst when the same stimuli were used in both lists. All differences between A and C stimuli and 8 of the 12 differences between A and A' stimuli were statistically significant (the exceptions were the BC and AC comparisons on Trial 1 and the WC and AC comparisons of errors to the first correct response). Of the differences between A' and C stimuli, only the BC comparison on the first trial reached significance.

Comparisons of the transfer groups with the A-B control showed significant positive transfer for all groups with A' and C stimuli in overall performance and at all the stages but Trial 1 and after the first correct response. The $A'\text{-}B_{BC}$ and $A'\text{-}B_{AC}$ groups did not differ significantly from the control on the first trial nor did the $A'\text{-}B_{WC}$ after the first correct response. Although the groups with A stimuli tended toward negative transfer, none of the differences reached significance.

Some evidence for the transfer of grouping information may be obtained from an analysis of overt errors during List-2 learning. The results of this analysis are shown in Table 7 in terms of the percentage generalized errors of the total overt errors during various stages of List-2 acquisition, the generalized errors now being defined in terms of second-list pairings. Since there were a large number of Ss who made only a few or no overt errors, the error data are simply pooled over all Ss in each condition. Despite the unknown

Table 7

Mean Percentage Generalized Errors

During Stages of List-2 Learning

	A-B	A-B_{WC}	A-B_{BC}	A-B_{AC}	A'-B_{WC}	A'-B_{BC}	A'-B_{AC}	C-B_{BC}	C-B_{AC}
Trial 1	.27	.83	.25	.18	.83	.12	.13	.34	.40
Before first correct	.42	.68	.35	.29	.73	.34	.40	.40	.36
After first correct	.52	.82	.30	.31	.76	.33	.41	.75	.44
to criterion	.47	.76	.32	.30	.73	.34	.41	.52	.39

reliability of the differences, an overall pattern can easily be detected. The only groups that made a considerably higher than chance proportion of generalized errors on the first trial were the A-B_{WC} and A'-B_{WC} groups, both of which made over three times as many of these errors as the A-B control group. Although the proportion of these errors increased and remained slightly above chance before and after the first correct response for the other groups, only with groups A-B_{WC}, A'-B_{WC} and C-B_{BC} (after the first correct response) did the proportion ever exceed that of the A-B control.

Free recall tests: number recalled. The mean numbers of adjectives recalled on the first and second free recall trials by the nine groups are shown in Table 8. There was no evidence of any differences between conditions on either trial, $F(8,198) = 1.604$ $p > .05$, for the first trial, and 1.294, $p > .05$ for the second trial (Table 9). An overall mean difference of .36 items between the two list sets did reach significance on the second free recall trial, $F(1,198) = 4.913$ but lists did not interact with conditions.

Free recall tests: clustering. Clustering of the response terms was measured in relation to the stimulus categories used in the A-B lists. Thus, the three adjectives paired with each set of categorized stimuli defined the response categories for all conditions.

Table 8

Mean Number of Adjectives Recalled on Free Recall Tests

	A-B	$A\text{-}B_{WC}$	$A\text{-}B_{BC}$	$A\text{-}B_{AC}$	$A'\text{-}B_{WC}$	$A'\text{-}B_{BC}$	$A'\text{-}B_{AC}$	$C\text{-}B_{BC}$	$C\text{-}B_{AC}$
Test 1	10.04	10.71	10.63	10.71	10.29	10.41	9.71	10.46	9.92
Test 2	10.41	10.83	10.21	10.67	10.71	10.33	10.04	10.46	10.17

Table 9

Summaries of Analyses of Variance for

Number Recalled on Free Recall Trials

Free Recall: Trial I

Source	SS	df	MS	F	p
Conditions (C)	24.92	8	3.12	1.60	>.05
Replications (R)	2.90	1	2.90	1.49	>.05
C x R	6.72	8	.84	< 1	-
Error	384.42	198	1.94	-	-
Total	418.96	215	-	-	-

Free Recall: Trial II

Source	SS	df	MS	F	p
Conditions (C)	14.84	8	1.86	1.29	>.05
Replications (R)	7.04	1	7.04	4.91	<.05
C x R	4.33	8	.54	< 1	-
Error	283.75	198	1.43	-	-
Total	309.96	215	-	-	-

Assessment of clustering was based on Bousfield and Bousfield's (1966) deviation measure, whereby the number of repetitions expected by chance, E(SCR), is subtracted from the number that was actually obtained, O(SCR), to yield a difference score, D(SCR), the amount of clustering not attributable to chance. The quantity O(SCR) was determined for each S by summing over the four categories, the number of times in recall that an item in one category was immediately followed by an item belonging to the same category. The E(SCR) values were calculated for each S by substitution in the formula, $E(SCR) = ([m^2_1 + \ldots + m^2_4]/n) - 1$, where the m_4 is the number of words recalled in the fourth category, and n is the total number of words recalled.

The mean O(SCR) and D(SCR) scores for the first and second free recall trials for the nine groups are shown in Table 10 along with the t values (df = 23). The only groups to show clustering significantly above chance on both trials were the $A\text{-}B_{WC}$, $A'\text{-}B_{WC}$, and $C\text{-}B_{BC}$ groups; the A-B control group showed clustering on the first trial only.

Since response groupings for $A\text{-}B_{AC}$, $A'\text{-}B_{AC}$, and $C\text{-}B_{AC}$ differed in the first and second lists, clustering was also measured in relation to the first-list pairings for these groups. The $\bar{D}(SCR)$ values were -.57, .37, and .14 for the first recall trials, and -.40, -.04, and .09 for the second for groups $A\text{-}B_{AC}$, $A'\text{-}B_{AC}$, and $C\text{-}B_{AC}$, respectively. The .37 difference for the $A'\text{-}B_{AC}$ was significant at the .05 level, one-tailed test.

Table 10

Mean Obtained and Difference Clustering Scores on Free Recall Trials

	Trial 1			Trial 2		
	$\bar{O}$(SCR)	$\bar{D}$(SCR)	*t*	$\bar{O}$(SCR)	$\bar{D}$(SCR)	*t*
A-B	2.50	.86	3.805**	2.08	.39	.796
$A\text{-}B_{WC}$	3.25	1.45	5.513**	3.13	1.31	3.619**
$A\text{-}B_{BC}$	1.96	.21	.894	1.92	.25	1.025
$A\text{-}B_{AC}$	1.50	-.26	- .963	2.08	.34	1.459
$A'\text{-}B_{WC}$	2.42	.73	2.332*	2.63	.86	2.663*
$A'\text{-}B_{BC}$	2.08	.37	1.217	1.96	.34	1.596
$A'\text{-}B_{AC}$	1.58	.00	.016	1.92	.28	1.191
$C\text{-}B_{BC}$	2.46	.67	2.120*	2.96	1.21	5.042**
$C\text{-}B_{AC}$	1.96	.34	1.377	2.08	.38	1.297

* difference significant at .025 level of probability, one-tailed test

** difference significant at .005 level of probability, one-tailed test

Stimulus term recall. The mean numbers of correctly recalled nouns from the first and second lists for each group are shown in Table 11 and the analysis of these data in Table 12. List-2 noun recall was significantly superior to List-1 recall over all conditions $F(1,184) = 232.79$, $p < .001$, as well as within each specific condition. With lists combined, there was no overall effect of conditions, $F(7,184) = 1.84$, $p < .05$, but there was a significant interaction between list recalled and condition, $F(7,184) = 2.72$, $p < .05$. Further analyses found no differences between conditions in List-2 recall, but significantly poorer recall for the A'-B_{BC} and A'-B_{AC} groups than for any of the other groups in List-1 recall.

Since the number of incorrect list identifications were approximately equal for all conditions and both lists, these data will not be presented.

Table 11

Mean Noun Recall From First and Second Lists

	A-B	$A\text{-}B_{WC}$	$A\text{-}B_{BC}$	$A\text{-}B_{AC}$	$A'\text{-}B_{WC}$	$A'\text{-}B_{BC}$	$A'\text{-}B_{AC}$	$C\text{-}B_{BC}$	$C\text{-}B_{AC}$
List 1	-	8.83	8.38	8.12	8.38	7.29	7.04	8.95	8.33
List 2	11.33	11.25	10.88	11.21	10.17	10.25	10.50	10.79	10.67

Table 12

Summary of Analysis of Variance of Number

of Stimulus Terms Recalled

Source	SS	df	MS	F	p
Conditions (C)	75.10	7	10.73	1.84	>.05
Error (a)	1071.04	184	5.82	-	-
Lists (L)	622.71	1	622.71	232.79	<.01
C x L	50.98	7	7.28	2.72	<.05
Error (b)	492.79	184	2.68	-	-
Total	2312.62	383	-	-	-

Chapter 4

Discussion

This experiment investigated the response-term grouping process in paired-associate learning whereby the response terms are subcategorized into sets corresponding to similarity groupings among the stimulus terms. The two questions to which the research was addressed were: (1) what is the nature of the relationships among the responses within each set once grouping has occurred, and (2) what is the effect of this grouping on subsequent acquisition performance. The findings that are relevant to each of these questions will be presented and discussed in separate sections following a brief review of the procedure and rationale of the experiment.

Procedure and rationale. In the experimental conditions *Ss* learned two categorized PA lists in which the stimulus terms in the two lists were identical (A), different instances of the same categories (A'), or instances of different categories (C). Response terms were identical in both lists but re-paired in second-list learning in three different ways. In WC or within-category re-pairing, the three responses paired with stimuli from a given category in the first list were re-paired with stimuli from the same category in the second list. In BC or between-category re-pairing the three responses were re-paired with stimuli from a different first-list category. Finally, in AC or across-category re-pairing, the responses were re-paired with stimuli from three different first-list categories.

The three re-pairing schemes were designed to disrupt different kinds of relationships that could exist among the responses that were grouped together in first-list learning. Three such relationships were suggested. First, responses in the first list could be grouped on the basis of their common association (correct or incorrect) with specific stimulus terms. Second, the responses could be grouped on the basis of their common associations with the conceptual category represented by the specific stimulus terms. Third, the responses, by virtue of either of these two types of relationships, could become directly associated with one another, and thereafter be independent of the specific stimulus terms or stimulus categories with which they had been paired.[1] With identical stimuli in the two lists, all three types of relationships among the first-list grouped responses would be maintained among the second-list grouped responses following WC re-pairing. BC re-pairing would be expected to disrupt associations among the responses that

---

[1] The three types of relationships may be illustrated as follows: The response terms "loud", "willing", and "complete" were paired respectively with the stimulus terms "zinc", "copper", and "silver" in the first list. The first relationship is exemplified by all three response terms being associated with each of the specific stimulus terms, i.e., zinc, copper, and silver. The second relationship is exemplified by all three responses being associated with the same category, "metal". The third relationship is exemplified by the responses "loud", "willing", and "complete", becoming directly associated with one another.

were mediated by either specific stimulus terms or by stimulus categories so that only direct associations among the grouped responses would be the same in the two lists. Finally, AC re-pairing would be expected to disrupt all three types of relationships with the result that neither mediated nor direct associations among the grouped responses in the two lists would be the same. Similar relations between the grouped responses in the two lists would be expected with A' or C stimuli with the exception that the associations mediated by specific stimuli would be different in the two lists. WC re-pairing would keep any category-mediated or direct associations intact, BC re-pairing would keep only direct associations intact, and AC re-pairing, neither. Thus, by investigating the conditions under which transfer of groupings to the second list occurs, it should be possible to determine the type of relationships that existed among the responses at the end of first-list acquisition.

The nature of the relationship between responses. The amount of grouping transfer was assessed by two measures: the proportion of generalized errors during list-2 learning and the occurrence of response-term clustering in free recall tests following acquisition.

With respect to the first of these measures, it was expected that maintaining the same response groupings in the two lists would result in an increase in the proportion of generalized errors (evaluated in terms of second-list pairings) at all stages of acquisition, when compared to the errors for a group learning only the second list. On the other hand, if first-list groupings were incompatible with those in a second list, it was expected that fewer such errors would be made until after the new groupings had been acquired (indicated arbitrarily in this study by the occurrence of the first correct response).

An examination of the second-list error data in view of these predictions showed that positive transfer of response groupings occurred for only the within-category re-paired conditions, and that this result was independent of whether the specific stimuli within these categories represented the same or new instances in the two lists. That is, only the A-B_{WC} and the A'-B_{WC} conditions showed a substantially larger proportion of generalized errors than the control group at all stages of learning. Although there were some exceptions in the C-B_{BC} and C-B_{AC} conditions, between-category and across-category re-pairing conditions, for the most part, showed the same trend toward negative grouping transfer, the proportion of generalized errors under these conditions being smaller than in the control group. The major apparent exception to this trend was found in the C-B_{BC} condition which showed a considerably larger proportion of generalized errors after the first correct response than did the control (i.e., .75 vs. .52 respectively). However, since only a relatively small number of errors were made by Ss in this condition, the reliability of this finding is questionable. (In this condition, a total of 65 overt errors were made by only 14 Ss, and, of these, 25 were contributed by a single S.) Since no consistent differences were found in the proportion of generalized errors following AC or BC re-pairing in the second list, there is little evidence in these data to suggest that response terms within a set had become directly associated to one another during grouping in the first list. On the other hand, there is considerable evidence for category-mediated associations among the responses within a set. WC re-pairing led to an increase in the proportion of generalized errors and BC re-pairing led to a decrease of these errors, regardless of whether the stimuli in the two lists consisted of the same or different instances on the same categories.

The second measure of grouping transfer, clustering of the responses according to second-list pairings, showed a similar pattern of results. The rationale for using this measure was that if the occurrence of response term clustering in free recall does reflect the categorization of responses into sets via grouping, disruption of the critical relationship among a set of responses should reduce the amount of clustering that occurs following second-list learning. More positively, one would expect clustering only if the critical relationship among the response terms within a set was maintained in the two lists.

There is little evidence to suggest that responses associated with stimuli from the same category become directly associated with one another. If this were the case, one would expect response-term clustering to occur, after both within-category and between-category re-pairing, regardless of interlist stimulus similarity since the same response terms would comprise each set in the two lists in both these conditions. Although within-category re-pairing did, in fact, produce this clustering, two of the three between-category re-pairing conditions, A-B_{BC} and A'-B_{BC}, did not. Although clustering on the part of the C-B_{BC} condition might suggest that response recall is independent of specific stimuli or stimulus categories, such an interpretation would demand that the two other between-category re-pairing conditions also show this effect.

An alternate interpretation is that response-term clustering occurs via recall of specific S-R associations (either forward or backward). As will be recalled, such an interpretation would demand that S make use of specific stimulus terms as cues to recall the responses that had been correctly or incorrectly associated with them. Whereas this mechanism could clearly account for the occurrence of clustering in the A-B_{WC} condition, it is difficult to account for the same findings in the A'-B_{WC} and C-B_{BC} conditions, both of which used different stimulus terms in the second list. The failure of the A-B_{BC} condition to produce significant clustering, on the other hand, is quite compatible with the mechanism of recall by specific associations.

Since every stimulus term in this condition was associated with responses from two different categories in the two lists, recall of the responses by specific stimuli should not be expected to produce consistent clustering according to second-list pairings. One would expect, instead, that a response associated with a given stimulus in the first list would be recalled immediately preceding or following a response associated with it in the second list. An examination of the recall protocols for the $A\text{-}B_{BC}$ conditions gives some fairly strong support for this interpretation: of a total of 230 response terms that preceded or followed another response term on the free recall test, 68 or approximately 30% of these had been associated with the same stimulus term in first- and second-list learning. By chance, only 2/11 or 18% of these responses should have been of this type.

The same type of analysis, however, also lends some support to an interpretation that stimulus categories, as opposed to specific stimuli, mediate response recall. Since, in the between-category re-pairing conditions, each stimulus category was associated with two different sets of three responses in the two lists, recall via stimulus category would not be expected to produce consistent clustering for either the $A\text{-}B_{BC}$ or $A'\text{-}B_{BC}$ conditions. The relevant data, in this case, are the frequencies with which a response associated with a stimulus from a given category in the first list is immediately preceded or followed by a response associated with any stimulus from the same category in the second list. In the A-B_{BC} condition, 132 or 57% of the responses were of this type as were 130 or 57% of the 228 possible responses in the A'-B_{BC} condition. On the other hand, since approximately 55% (6/11) of the responses would be expected to be of this type of chance alone, these data by themselves do not constitute very strong evidence for an interpretation of response clustering mediated by stimulus categories alone. The major support for this interpretation comes from the occurrence of clustering for the A'-B_{WC} and C-B_{BC} conditions. With the A'-B_{WC} condition, the responses within each set were paired with the same stimulus category even though specific stimuli within the categories were different in the two lists. Recall of the responses via category names would thus be expected to proceed in clusters of responses associated with these stimulus categories. In the C-B_{BC} condition, on the other hand, each set of responses was associated with two different categories of stimuli in the two lists. However, since unlike the A-B_{BC} and A'-B_{BC} conditions, the C-B_{BC} group involved completely new stimulus categories in the second list, recall via category names (from either the first or second list) would still be expected to produce consistent clustering.

In general, taking both the generalized error and the response clustering data into account, the results of this experiment would seem to indicate that response-term categories formed on the bases of their associations with similar stimuli are maintained through their association with the category represented by the stimuli. There was no evidence in this study that the responses within a set became directly associated with one another.

The effect of grouping on performance. The second purpose of this experiment was to determine the effect of grouping on performance. Generally, one would expect that if response grouping does facilitate the acquisition of specific associations by limiting the number of items that interfere in the learning of a single pair, then the conditions which led to positive grouping transfer in this experiment would also have led to more rapid acquisition of the second-list pairs. The results of this experiment, however, do not support this hypothesis. Despite some evidence of an advantage early in learning, re-pairing responses within the same stimulus categories did not lead to more rapid acquisition than re-pairing between or across categories, even though within-category re-pairing did produce greater evidence of grouping in the second list. In fact, although Ss did show evidence of restricting the response alternatives for stimuli within a category in terms of proportion of generalized errors, if anything this restriction seemed to make it even more difficult to learn the specific associations within each set, particularly in the $A\text{-}B_{WC}$ condition.

This result seems contrary to the findings of Runquist (1968b) that reversal re-pairing (re-pairing the response with a similar stimulus) in an A-B, A-B_r transfer task, leads to faster acquisition. Runquist, however, did not continue second-list learning until Ss reached a criterion of all correct on a single trial and, thus, the negative effects of grouping on differentiating between pairs within a set would probably not yet have appeared. The slight suggestion of early facilitation in the present study (fewer errors on the first test trial and to the first correct response in second-list learning) for the within-category re-pairing conditions suggests that this indeed might be the case.

With respect to the effect of grouping on peformance, then, the transfer data of this experiment suggest that, while grouping may lead to an early advantage in learning by increasing the probability of a correct response by guessing, it will not lead to more rapid acquisition of specific associations under all conditions.

Conclusions. The findings of this experiment support the following two conclusions: (1) The response-term groupings acquired on the basis of stimulus similarity in paired-associate learning are dependent upon the categories to which the paired stimuli belong rather than the specific stimulus terms within the categories. Furthermore, response terms within each set do not become directly associated with one another, but rather, depend entirely upon their relationship with specific stimulus categories.

(2) Although grouping the response terms into sets does restrict the interference in learning each S-R pair to other items within that set, such grouping does not facilitate the acquisition of specific associations.

The latter conclusion, however, may be limited to situations where conceptual similarity relationships among the stimuli are the basis for response-term grouping. As Underwood (1964) has pointed out, with conceptually similar nouns as stimuli, it could be the case that "the gross mediating responses (category names) were so strong that they interfered with the development of more precise discriminations (p. 72)." If this were indeed the case, the use of other types of stimulus similarity relationships which provide clearly defined rules for differentiating among pairs within subsets of similar stimuli and grouped responses might be expected to show facilitation following grouping. In the absence of such data, there is little basis for concluding that grouping per se, under any conditions, facilitates learning of specific associations.

References

Battig, W.F. Evidence for coding processes in "rote" paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 177-181.

Battig, W.F. Paired-associate learning. In T.R. Dixon and D.L. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Battig, W.F., Brown, S.C., and Nelson D. Constant vs. varied serial order in paired-associate learning. Psychological Reports, 1963, 12, 695-721.

Battig, W.F. and Montague, W.E. Category norms for verbal items in 56 categories. Journal of Experimental Psychology Monograph, 1969, 80, No. 3, Part 2.

Bousfield, A.K. and Bousfield, W.A. Measurement of clustering and of sequential constancies in repeated free recall. Psychological Reports, 1966, 19, 935-941.

Brown, S.C. Interpair interference as a function of level of practice in paired-associate learning. Journal of Experimental Psychology, 1964, 67, 316-323.

Brown, S.C. and Battig, W.F. Partial serial-position constancy in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1962, 1, 42-47.

Brown, S.C. and Battig, W.F. Second-list paired-associate facilitation produced by addition of previously learned first-list pairs. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 320-321.

Brown, S.C., Battig, W.F., and Pearlstein, R. Effect of successive addition of stimulus elements on paired-associates learning. Journal of Experimental Psychology, 1965, 70, 87-93.

Cohen, B.H. Some-or-none characteristics of coding behavior. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 182-187.

Horowitz, L.M. and Prytulak, L.S. Redintegrative memory. Psychological Review, 1969, 2, 1-21.

Johnson, N.F. The role of chunking and organization in the process of recall. In G.H. Bower (Ed.), The psychology of learning and motivation, Vol. 4. New York; Academic Press, 1970.

Joinson, P.A. and Runquist, W.N. Effects of intralist stimulus similarity and degree of learning on forgetting. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 554-559.

Mandler, G. Organization and memory. In K.W. Spence and J.T. Spence (Eds.), The psychology of learning and motivation, Vol. 1. New York: Academic Press, 1967.

Miller, G.A. The magical number seven, plus or minus two: Some limits in our capacity to process information. Psychological Review, 1956, 63, 81-97.

Runquist, P.A. Clustering in free recall following paired-associate learning. Journal of Experimental Psychology, 1970, 83, 260-265.

Runquist, W.N. Order of presentation and number of items as factors in paired-associate verbal learning. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 535-540.

Runquist, W.N. Intralist interference as a function of list length and interstimulus similarity. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 7-13.

Runquist, W.N. Effects of variable frequency of presentation on paired-associate learning as a function of interstimulus similarity. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 470-475.

Runquist, W.N. Permanent effects of variable frequency of presentation in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 1060-1064 (a).

Runquist, W.N. Reversal vs. non-reversal re-pairing in categorized paired-associate lists. Canadian Journal of Psychology, 1968, 22, 285-293 (b).

Schild, M.E. and Battig, W.F. Directionality in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 42-49.

Shuell, T.J. Clustering and organization in free recall. Psychological Bulletin, 1969, 72, 353-374.

Thorndike, E.L. and Lorge, I. The teacher's word book of 30,000 words. New York: Teacher's College, Columbia University, Bureau of Publications, 1944.

Tulving, E. Intratrial and intertrial retention: Notes towards a theory of free recall verbal learning. Psychological Review, 1964, 71, 219-237.

Tulving, E. Theoretical issues in free recall. In T.R. Dixon and D.L. Horton (Eds), Verbal behavior and general behavior theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Underwood, B.J. The representativeness of rote verbal learning. In A.W. Melton (Ed.), Categories of Human Memory. New York: Academic Press, 1964.

Underwood, B.J., Ekstrand, B.R., and Keppel, G. An analysis of intralist similarity in verbal learning with experiments on conceptual similarity. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 447-462.

Appendix A

Materials

Set 1

Stimulus Terms				Response Terms			
1	2	3	4	R	WC	BC	AC
SALT	PEPPER	CARROT	PEA	SHARP	DIRECT	MARRIED	EASY
CINNAMON	SUGAR	LETTUCE	SPINACH	DIRECT	GENERAL	RECENT	RECENT
GARLIC	VANILLA	TOMATO	CORN	GENERAL	SHARP	OPPOSITE	TOTAL
SHIRT	DRESS	CAT	PIG	EASY	UPPER	SHARP	HOLY
SHOES	PANTS	ELEPHANT	HORSE	FALSE	EASY	DIRECT	DIRECT
COAT	SOCKS	BEAR	LION	UPPER	FALSE	GENERAL	OPPOSITE
RUM	WINE	DRUM	TRUMPET	MARRIED	OPPOSITE	ABLE	ABLE
WHISKEY	GIN	VIOLIN	PIANO	RECENT	MARRIED	HOLY	UPPER
BEER	VODKA	FLUTE	GUITAR	OPPOSITE	RECENT	TOTAL	SHARP
KNIFE	BOMB	TRUCK	TRAIN	ABLE	HOLY	EASY	MARRIED
RIFLE	CLUB	BICYCLE	BUS	HOLY	TOTAL	FALSE	FALSE
SWORD	GUN	AIRPLANE	BOAT	TOTAL	ABLE	UPPER	GENERAL

Appendix A (continued)

Materials

Set 2

Stimulus Terms				Response Terms			
1	2	3	4	R	WC	BC	AC
ORANGE	BANANA	SILVER	ALUMINUM	SIMPLE	EARLY	HAPPY	COMPLETE
CHERRY	PLUM	ZINC	GOLD	EARLY	UNKNOWN	LOUD	LOUD
LEMON	GRADE	COPPER	TIN	UNKNOWN	SIMPLE	NORTHERN	WILLING
FRANCE	CHINA	ANT	FLY	COMPLETE	GENTLE	SIMPLE	PRIVATE
GERMANY	RUSSIA	BEE	SPIDER	THIN	COMPLETE	EARLY	EARLY
SPAIN	ENGLAND	MOSQUITO	BEETLE	GENTLE	THIN	UNKNOWN	NORTHERN
SWIMMING	TENNIS	TABLE	LAMP	HAPPY	NORTHERN	BRIEF	BRIEF
FOOTBALL	HOCKEY	BED	SOFA	LOUD	HAPPY	PRIVATE	GENTLE
GOLF	BASEBALL	COUCH	CHAIR	NORTHERN	LOUD	WILLING	SIMPLE
LEG	HAND	WINDOW	DOOR	BRIEF	PRIVATE	COMPLETE	HAPPY
NOSE	FOOT	ROOF	WALL	PRIVATE	WILLING	THIN	THIN
STOMACH	EYE	CEILING	FLOOR	WILLING	BRIEF	GENTLE	UNKNOWN

Appendix B

Summaries of Duncan's Tests for List-2 Learning

Trials to criterion

	C-B_{BC}	A'-B_{WC}	A'-B_{AC}	A'-B_{BC}	A-B_{AC}	A-B	A-B_{BC}	A-B_{WC}
	6.67	7.88	8.33	8.46	11.00	11.75	12.17	12.54
C-B_{AC}	.71	1.92	2.37	2.50	5.04*	5.79*	6.21*	6.58*
C-B_{BC}		1.21	1.66	1.79	4.34*	5.09*	5.51*	5.83*
A'-B_{WC}			.45	.58	3.12*	3.42*	4.29*	4.66*
A'-B_{AC}				.13	2.67*	3.29*	3.84*	4.21*
A'-B_{BC}					2.54*	3.87*	3.71*	4.04*
A-B_{AC}						.75	1.17	1.54
A-B							.42	.79
A-B_{BC}								.37

Shortest significant range: 2.50, 2.63, 2.72, 2.79, 2.84, 2.88, 2.92, 2.95

* difference significant at .05 level of probability, two-tailed test

Errors on first test trial

	C-B_{AC}	A'-B_{WC}	A'-B_{BC}	A'-B_{AC}	A-B	A-B_{WC}	A-B_{BC}	A-B_{AC}
	6.79	6.96	8.08	8.25	9.00	9.00	9.25	9.33
C-B_{BC}	.46	.63	1.75*	1.92*	2.67*	2.67*	2.92*	3.00*
C-B_{AC}		.17	1.29	1.46	2.21*	2.21*	2.46*	2.54*
A'-B_{WC}			1.12	1.29	2.04*	2.04*	2.29*	2.37*
A'-B_{BC}				.17	.92	.92	1.17	1.25
A'-B_{AC}					.75	.75	1.00	1.08
A-B						.00	1.25	.33
A-B_{WC}							1.25	.33
A-B_{BC}								.08

Shortest significant range: 1.42, 1.49, 1.54, 1.58, 1.61, 1.63, 1.65, 1.67

* difference signficant at .05 level of probability, two-tailed test

Errors to the first correct response

	C-B_{AC}	A'-B_{WC}	A'-B_{BC}	A'-B_{AC}	A-B_{WC}	A-B_{AC}	A-B_{BC}	A-B
	16.46	21.54	22.46	23.83	31.00	31.33	35.00	35.42
C-B_{BC}	.08	5.16	6.08	7.45	14.62*	14.95*	18.62*	19.04*
C-B_{AC}		5.08	6.00	7.37	14.54*	14.87*	18.54*	18.96*
A'-B_{WC}			.92	2.29	9.46	9.79	13.46*	13.88*
A'-B_{BC}				1.37	8.54	8.87	12.54*	12.96*
A'-B_{AC}					7.17	7.50	11.17*	11.59*
A-B_{WC}						1.33	4.00	4.24
A-B_{AC}							3.67	3.09
A-B_{BC}								.42

Shortest significant range: 9.41, 9.91, 10.25, 10.49, 10.68, 10.84, 10.98, 11.09

* difference significant at .05 level of probability, two-tailed test

Errors after the first correct response

	C-B_{BC}	A'-B_{AC}	A'-B_{BC}	A'-B_{WC}	A-B	A-B_{AC}	A-B_{BC}	A-B_{WC}
	4.92	6.83	6.83	8.46	14.79	15.67	21.04	23.71
C-B_{AC}	.67	2.58	2.58	4.21	10.50*	11.42*	16.79*	19.46*
C-B_{BC}		1.91	1.91	3.54	9.87*	10.75*	16.12*	18.79*
A'-B_{AC}			.00	1.63	7.96*	8.84*	14.21*	16.88*
A'-B_{BC}				1.63	7.96*	8.84*	14.21*	16.88*
A'-B_{WC}					6.33	7.21*	12.58*	15.25*
A-B						.88	6.25	8.92*
A-B_{AC}							5.37	8.04*
A-B_{BC}								2.67

Shortest significant range: 6.34, 6.68, 6.90, 7.07, 7.20, 7.31, 7.39, 7.47

* difference significant at .05 level of probability, two-tailed test