Psychological Bulletin
1969, Vol. 72, No. 5, 353–374

CLUSTERING AND ORGANIZATION IN FREE RECALL [1]

THOMAS J. SHUELL [2]

State University of New York at Buffalo

Research on clustering and subjective organization (SO) in free recall is reviewed and evaluated. Various indexes developed to measure clustering and SO are evaluated, and two intercorrelation matrices among clustering measures and the number of words recalled are presented. The existence of a large negative bias in the correlation between the ratio of repetition (*RR*) measure and recall is demonstrated. Various theoretical issues which have developed from the study of organization in free recall are presented and discussed.

Psychologists have long been interested in the way organization affects memory. Over the years there have been several, rather different, approaches to the problem. Probably the most productive to date has been the study of clustering and subjective organization in free recall since an attempt has been made to quantify the organizational process.

In free recall the subject is presented a list of words to learn, and he is told to recall the items in the order in which he thinks of them. Interestingly, certain regularities appear in the order in which the items are recalled. For example, there is a tendency for items which are somehow related to one another to be recalled together even though these items were not contiguous during presentation. This discrepancy between the order in which the items were presented and the order in which they were recalled is presumed to represent a tendency on the part of the subject to organize his recall on the basis of various second-order habits, that is, preexperimental associations or conceptual relationships. This tendency for related items to be recalled together has been termed clustering.

Clustering was apparently first observed by Bousfield and Sedgewick (1944) while they were studying sequential characteristics of associative responses. When they asked subjects to list items in specified categories, they noticed that the subjects tended to respond with sequences of related items. For example, if asked to list a series of animals, the subjects might first write down several felines, then several canines, etc. Later, Bousfield (1953) developed a technique for quantifying clustering. During the ensuing 15 years, a great deal of research has been done on clustering in free recall. The main purpose of this research has been to explicate the relationship between organization and memory. This purpose can best be represented by a quote from Bousfield's (1953) original paper

> The theoretical significance of this undertaking derived in part from the assumption that clustering is a consequence of organization in thinking and recall. If clustering can be quantified, we are provided with a means of obtaining additional information on the nature of organization as it operates in the higher mental processes [p. 229].

The purpose of the current paper is to review and evaluate this research, to determine how well its original purpose has been fulfilled, and to raise some questions as to what direction future research in this area might take.

Three paradigms have been developed for the study of organization in free recall. These differ primarily in the experimental situation used for inducing clustering. Two of these paradigms are similar in that the basis of organization is determined by the experimenter. The first of these is called *categorical clustering*. So far, the bulk of the research in the area, including all of the early studies

[1] An earlier draft of this paper was written during the author's tenure as a predoctoral National Institutes of Health trainee at the Institute of Human Learning, University of California, Berkeley. The Institute is supported by grants from the National Science Foundation and the National Institutes of Health.

[2] Requests for reprints should be addressed to Thomas J. Shuell, Educational Psychology Department, State University of New York, Buffalo, New York 14214.

by Bousfield, has used this paradigm. In this situation the stimulus list is comprised of words representing two or more mutually exclusive conceptual categories. The other paradigm is called *associative clustering* and refers to the situation in which the stimulus list is comprised of associatively related words (as determined from associative norms) which are not members of the same conceptual category.

Subjective organization differs from the other two paradigms in that the basis of organization is not predetermined by the experimenter. Rather, the stimulus list is comprised of so-called unrelated words, that is, a random sample of words in which the experimenter has made no attempt to include words which are categorically or associatively related. In fact, he frequently tries to minimize or eliminate such relationships among the words. Thus, the subject is more or less free to organize the words in any way that he wishes. Organization is determined by the extent to which the subject recalls the words in the same order on two successive trials. This differs from categorical and associative clustering which compare the order of recall on any given trial with the organization, as defined by the experimenter, which is present in the stimulus list. Therefore, a multitrial experiment is required in order to measure subjective organization while categorical and associative clustering can be measured after a single trial.

It seems most probable that all three paradigms are dealing with the same basic psychological processes. The main differences are the experimental procedures which are employed and the extent to which some type of organization or relatedness is apparent to the subject. Any paradigm will, of course, influence the types of questions which are asked by investigators using that paradigm. The present situation is no exception, although these influences are not always made explicit. A case in point, which will be discussed in more detail later, is the single-trial versus multitrial nature of the various paradigms (cf. Cofer, 1967; Tulving, 1968).

The present paper is divided into three major sections. The first section deals with the various techniques which have been developed for measuring organization in free recall. The ability to quantify organization is a major reason for the interest in clustering and subjective organization. The second section discusses some of the major findings of the various research studies on organization in free recall. The final section considers the major theoretical interpretations of the relationship between organization and memory which have resulted from the study of clustering and subjective organization.

The Measurement of Organization

A variety of measures have been developed to measure organization in free recall. It is possible to classify these measures in several different ways. For purposes of this paper, the major distinction is made in terms of whether they were designed to measure organization as determined by the experimenter or organization as determined by the subject. Then, within each of these two classes the various measures are discussed in terms of the assumptions underlying the method.

Predetermined Basis of Organization

The procedure typically followed in the study of categorical clustering can be exemplified by Bousfield's original experiment (1953). A list of 60 nouns containing 15 examples from each of four mutually exclusive categories (i.e., animals, names, professions, and vegetables) was constructed. These words were then read to the subjects in a random order. Following the presentation of the list, the subjects were given 10 minutes to write down as many of the words as they could remember. The order in which the words are recalled is used to determine the amount of organization present.

All of the indexes used to measure categorical clustering, with the exception of the first one discussed, are based on the number of repetitions which occur during recall. A repetition is said to occur any time two items from the same category are recalled one after the other. The number of repetitions for a sequence of items from a given category is always the number of items in the sequence minus one, since the first item in the sequence is not a repetition. Summing over all of the sequences in the protocol for one sub-

ject gives the total number of repetitions for that subject.[3]

Cluster size. If subjects organize the words to be remembered on the basis of the categories present in the stimulus list, then one reasonable way to measure the amount of this organization would be to count the number of words appearing in clusters representing one of the categories. If there is little or no organization, then most of the words would probably appear by themselves. If, however, organization does occur, then there should be a relatively large number of words appearing in clusters. Bousfield (1953) tabulated the number of times single (unclustered) words and clusters of varying numbers of words occurred in the subjects' recalls. He made a similar tabulation for a parallel, artificial experiment in which 100 sequences, matched in length with those of the actual subjects, were drawn without replacement from a box containing 15 capsules each of four different colors. While the results of the artificial experiment exhibited a greater occurrence of single words and clusters of two words than were produced by the real subjects, the subjects produced many more clusters of larger sizes than were obtained from the artificial experiment. While this tabulation appears to be a valid measure of clustering, it seldom has been used in subsequent papers. Probably the main reason it has not been used more frequently is that it does not provide a single, summary measure.

Items equally available. If one is willing to assume that all of the words presented for learning are equally available for recall, then it is possible to apply a probability model to determine if the amount of clustering observed exceeds a chance level. Using the normal approximation to the binomial, Bousfield (1953) developed what he called the index of repetition (*IR*). The *IR* is based on the probability of a given word being followed by another word in the same category if the order of recall occurs on a random basis. The index can be represented by the following formula:

$$IR = \frac{cr - n}{\sqrt{n(c-1)}} \qquad [1]$$

in which r is the number of repetitions observed, n is the total number of sequences in a given protocal and is always the total number of words recalled minus one, and c is the number of categories in the stimulus list.[4] The *IR* is normally distributed with a mean of zero and a standard deviation of one, and it is independent of the number of words recalled. The extent to which the obtained mean IR deviates from the expected value of zero reflects the extent to which clustering occurred during recall. However, the model is one of sampling with replacement, and the validity of this model may be seriously questioned particularly when short lists are used. The *IR* has not been used a great deal in subsequent papers.

The measure which has been used most frequently in studies of clustering is the ratio of repetition (*RR*—Bousfield, 1953). The *RR* is a simple ratio of the number of obtained repetitions to the number of repetitions possible for the words recalled. This can be represented by

$$RR = \frac{r}{N-1} \qquad [2]$$

where N is the total number of words recalled.[5]

Tables of chance values of *RR* for various types of lists have been constructed by Cohen, Sakoda, and Bousfield (1954). They used a model of sampling without replacement from a pool comprising all the items in a given

[3] In computing the various measures of clustering, categorical intrusions (i.e., words recalled which were not presented to the subject but which belong to one of the categories represented in the stimulus list) have generally been treated the same as correct items. The effect which this procedure has on the measures is probably minimal, but it has never been compared with the procedure in which categorical intrusions are excluded.

[4] The assumption made here, of course, is that all of the items, and therefore all of the categories, are equally available for recall. If one wished to assume that unless at least one item from a given category is recalled, the category is not available for recall and should not be counted in computing the probability of obtaining a repetition, then c would be the number of categories represented by the words recalled.

[5] The formula presented in the original 1953 paper was r/N. However, since $r/(N-1)$ is the formula which has been used most frequently in subsequent papers, it is the one presented.

list. These tables were used in many subsequent papers to determine the amount of clustering that could be expected by chance for a given categorized list.

The main advantage of *RR* over *IR* is the ease with which the measure can be computed. However, the *RR* does not take into account the transition from one category to another which occurs even with perfect clustering. Because of this, it is not independent of the number of categories represented by the words recalled. For example, the only situation in which *RR* can equal one is when only a single category is recalled.

Robinson (1966) developed the item-clustering index (*ICI*) in order to allow for this possibility of perfect clustering in spite of the gaps in repetitions resulting from categorical transitions. The formula which he presents is

$$ICI = \frac{r}{c(Wc - 1)} \qquad [3]$$

in which r is the total number of repetitions occurring in recall, c is the number of categories represented by the words recalled, and the Wc is the number of items per category in the stimulus list. Chance values were not computed. The *ICI*, however, does not take into consideration the possibility of perfect clustering occurring for the words actually recalled; that is, the only time *ICI* can equal one is when all of the words in a certain number of categories are recalled. Thus, the assumption is made that all of the items within a given category are available for recall given that at least one word in that category is recalled.

The methods which have been developed for measuring associative clustering (Jenkins, Mink, & Russell, 1958; Jenkins & Russell, 1952) also make the assumption that all of the items are equally available for recall. Clustering was said to occur when the two words in a given stimulus-response pair occurred together in recall. Jenkins and Russell (1952) used the occurrence of arbitrary pairs, that is, a stimulus word followed not by its own response but by another specified and randomly selected response from the list as a chance base line against which the occurrence of both forward and backward associations was compared. Jenkins, Mink, and Russell (1958) developed an alternative procedure in which the opportunity to cluster is taken into consideration. This method is used for comparing groups receiving different experimental manipulations and does not use a chance base line. Two studies (Jenkins, Mink, & Russell, 1958; Wicklund, Palermo, & Jenkins, 1965) have reported very similar conclusions when both of the methods are used for analyzing the same data.

All of the measures discussed in this section are based on the assumption that all of the items presented for learning are equally available for recall. However, this is not a valid assumption. For example, there is substantial evidence that a serial position effect is obtained in free recall under a wide variety of conditions (Murdock, 1962; Shuell, 1967; Tulving & Patterson, 1968). If all of the items are equally available such an effect should not be obtained. Therefore, the usefulness of these measures is limited since they are based on a faulty assumption.

Limited availability of items. Since it does not appear reasonable to assume that the items are equally available for recall, the question arises as to how one might proceed in developing a more valid model for the measurement of clustering. One such attempt has been made by Bousfield (Bousfield & Bousfield, 1966; Bousfield & Puff, 1964). The difference between the obtained and expected number of repetitions is used as the measure of clustering; the larger this difference the greater the clustering or organization. A distinction is made between item properties, that is, what items are recalled, and order properties, that is, what is the sequential ordering of those items which are recalled. In determining the number of repetitions that can be expected on the basis of chance, one is concerned with order properties rather than item properties. Therefore, the following assumptions are made: (*a*) When a subject begins his recall a certain number of items are *not* available and therefore can have no effect on the incidence of clustering, and (*b*) at every stage of recall all of the words *remaining* to be recalled are equally available and are chosen without replacement. The expected number of repetitions calculated on

the basis of this model is

$$E(r_k) = \frac{m(m-1)}{N} \quad [4]$$

where m is the number of words recalled in a category k, and N is the total number of words recalled. Summing over all of the categories in a particular list, the expected number of repetitions in the list as a whole can be represented by

$$E(r) = \frac{\sum_{i=1}^{k} m_i^2}{N} - 1 \quad [5]$$

This deviation is, of course, computed for each subject.

Bousfield and Bousfield (1966) also present a formula for calculating the expected number of times an item in one category, for example, Category i will be followed by an item from another specified category, for example, Category j. This formula is

$$E(r_{ij}) = \frac{m_i m_j}{N} \quad [6]$$

for $i \neq j$.

One shortcoming of the deviation measure is that no allowance is made for the maximum amount of clustering which can be obtained with the words recalled. This limitation becomes particularly important when there are fairly large discrepancies in the number of words recalled by subjects, for example, when lists of different length are used. Perfect clustering could be obtained in a short list, yet the value of the deviation measure could be less than that obtained in a longer list in which clustering is not perfect. In such situations it would be relatively simple to convert the deviation measure into a proportion similar to the one developed by Fagan (1968) for measuring subjective organization.

Comparison of the various measures. The various measures were empirically compared in a study by the present author (Shuell, 1967). Specifically, 336 college students received four alternate study-recall trials on a 35-item list in which there were five words representing each of seven conceptual categories. The words were selected from the 10 most frequent responses to the category names as reported in the norms of Cohen, Bousfield, and Whitmarsh (1957). Intrusions were disregarded in counting repetitions. The indexes used were the actual number of repetitions obtained, the observed minus the expected (Equation 5) number of repetitions (O − E), the *IR* (Equation 1), the *RR* (Equation 2), and the *ICI* (Equation 3). The intercorrela-

TABLE 1

INTERCORRELATIONS OF VARIOUS CLUSTERING MEASURES

Measure	O−E[a]	*RR*	*IR*	*NR*	*ICI*	*W/C*
Trial 1						
Words recalled	.76	.44	.73	.79	.61	.75
Words per category (*W/C*)	.70	.63	.65	.78	.84	
Item clustering index (*ICI*)	.91	.93	.89	.94		
Number of repetitions (*NR*)	.99	.87	.98			
Index of repetition (*IR*)	.99	.89				
Ratio of repetition (*RR*)	.87					
Trial 4						
Words recalled	.89	.62	.85	.90	.78	.85
Words per category	.85	.74	.79	.88	.92	
ICI measure	.95	.94	.93	.96		
Number of repetitions	1.00	.89	.99			
Index of repetition	.99	.92				
Ratio of repetition	.89					

[a] O−E is the observed minus the expected number of repetitions.

tion matrices for these measures, the number of words recalled, and the number of words recalled per category are presented for both Trial 1 and Trial 4 in Table 1. It can be seen that the correlations among the clustering measures are all quite high. Since all of the measures are primarily based on the number of repetitions obtained and differ only in the way corrections for various factors are applied, this finding is not very surprising. However, there are substantial differences in the degree to which the various measures correlate with the number of words recalled. For example, the correlation coefficient between the traditional *RR* measure and the number of words recalled is substantially lower than the coefficients for the other measures on both trials. The values of the correlation between *RR* and recall are in the range typically reported in other studies.

There are at least two possible explanations for these differences. First, it has already been noted that the assumption of equal availability which underlies the *RR* measure is not a reasonable assumption. However, the same assumption is made for the *IR* measure, and the correlation between *IR* and recall appears to be much larger than the one between *RR* and recall. Second, there is a statistical artifact in the correlation coefficient between *RR* and recall. All of the measures make some adjustment for differential recall. There is, of course, a direct relationship between the total number of words recalled and the opportunities for repetitions to occur. The *RR* makes this correction by calculating a ratio of the number of repetitions obtained to the total number of repetitions possible given the number of words recalled $(N - 1)$.[6] In effect, an attempt is made to provide a base line for comparison by partialing out the number of words recalled. In such a case, *RR* could not be expected to correlate highly with recall since the effect of the latter variable has already been removed. In fact, it can be shown that the expected value of the correlation coefficient between two random variables X and Y/X, assuming that X and Y are independent, is not zero but some negative value which is dependent upon the parameters of X and Y.[7] This expectation can be expressed approximately by

$$E(r_{x(y/x)}) \approx \frac{-C_x}{\sqrt{C_x^2 + C_y^2}} \qquad [7]$$

where C, the coefficient of variation, is equal to the standard deviation divided by the mean. Using the present data to estimate the value of C_x and C_y, the expected value of the correlation coefficient between *RR* and recall is $-.43$ for Trial 1 and $-.48$ for Trial 4. Thus, the obtained values of $+.44$ and $+.62$ are substantial in view of the large negative bias inherent in the relationship.

It appears most likely that this statistical artifact is the main reason for obtaining the lower values for the *RR* measure. Thus, while the *RR* measure is obviously able to detect clustering, it is clear that at least under certain conditions, use of this measure will result in somewhat different conclusions than will be reached from use of the other measures. However, there may be situations in which it is desirable to use a proportion measure, for example, when comparing lists of different lengths. In these situations, it may be best to convert the deviation measure into a proportion. It is interesting to note that at least for the conditions used in this study the various measures, with the exception of *RR*, appear to be roughly equivalent. The measure used will, of course, depend in part on the purpose of the investigator and the assumptions he is willing to make. Nevertheless, the present author considers the deviation measure, or some form of it, to be the best measure currently available since the model on which it is based seems to be the most valid.

Subject-Determined Basis of Organization

Probably the biggest drawback in using predetermined basis of organization is the impossibility of determining the extent to which the subjects are in fact using the organization present in the stimulus list. It is possible that bases of organization other than those

[6] It has already been pointed out that this is a biased measure since no allowance is made for transitions from one category to another, but this is not the main concern here.

[7] Y/X is taken as a reasonably close approximation to $Y/(X - 1)$.

specified by the experimenter are being used by the subjects to organize their recall. Also, to the extent that other bases of organization are being used, clustering, or the use of predetermined bases of organization, underestimates the amount of organization being used by the subjects. All of the measures discussed so far have been based exclusively on predetermined aspects of organization. Thus, it is possible that many interesting facets of the organizational process have been overlooked.

The study of subjective organization in free recall is an attempt to investigate those bases of organization which the subjects are actually using. No assumption is made concerning the organization present in the stimulus list. Rather, organization is measured by the subject's tendency to recall words in the same order on successive trials. The measurement of subjective organization was first developed by Tulving (1962b). Since then Bousfield and his associates (Bousfield & Bousfield, 1966; Bousfield, Puff, & Cowan, 1964) have developed alternative methods for measuring subjective organization.

While the measures of clustering are based on repetitions, the measures of subjective organization are based on intertrial repetitions (*ITR*). In determining the number of *ITR*s which occurred during recall, a matrix is constructed for each subject. All of the words in the stimulus list are represented along both the rows and the columns of the matrix. The rows represent the nth word recalled while the columns represent the $(n + 1)$th word recalled. The matrix can represent any number of trials from two to the total number of trials given. There is one slight difference between the matrix used by Tulving and the one used by Bousfield. The matrix used by Tulving has extra positions for no word immediately preceding the first word and for no word immediately following the last word recalled. The matrix used by Bousfield, however, does not have these additional blank positions. This is done since repetitions cannot be formed for these two situations. The matrix is used for making a tabulation of the frequency with which given pairs occur adjacent to one another on two immediately successive trials. An *ITR* is said to occur whenever a given cell is checked twice in succession as a consequence of recording the data for a given pair on Trial n and Trial $n + 1$.

Items equally available. Bousfield, Puff, and Cowan (1964) have developed a deviation measure of subjective organization which is similar in many respects to the deviation measure used in the study of clustering. The model used in calculating the expected value is one in which two successive random drawings are made from the same pool of W items in which W is the total number of words presented for learning. Each drawing is made without replacement, and at each stage of both drawings it is assumed that the remaining items are equally available. This expected value can be represented by the formula

$$E(ITR) = \frac{(h-1)(k-1)}{W(W-1)} \qquad [9]$$

in which h is the number of words recalled on Trial n, and k is the number of words recalled on Trial $n + 1$. The difference between the number of *ITR*s obtained and the expected value is then taken as a measure of subjective organization.

Unfortunately, this measure is based on a faulty assumption. It has already been pointed out that in free recall items are not equally available. It is not entirely clear how important the violation of this assumption is. Nevertheless, one should keep it in mind when using the measure.

Limited availability of items. Tulving's (1962b) original measure was derived from information theory, and he quite appropriately termed it the *SO* measure. It is a ratio measure in which the organization obtained is expressed as a fraction of the maximum organization possible. The rows and the columns of the matrix described above are summed, and the following formula is used:

$$SO = \frac{\sum_{i,j} n_{ij} \log n_{ij}}{\sum_{i} n_i \log n_i} \qquad [8]$$

In the formula n_{ij} represents the numerical value of the cell in the ith row and the jth column, and n_i represents the marginal total of the ith row. The value of *SO* can vary from zero to one, although it can assume the value of one only when the same words are recalled

on all of the trials represented by the matrix. When new words are recalled or old words omitted on some of the trials, the maximum value of *SO* is some value less than one.

While this formula considers only sequences occurring in one direction, it can be easily modified to consider sequences occurring in either direction, that is, Word *A* followed by Word *B* as well as Word *B* followed by Word *A*, by summing the corresponding cells across the diagonal. The matrix can easily be modified to provide an estimate of higher order organization, for example, by letting the columns represent the $(n + 2)$th word rather than the $(n + 1)$th word. The method can also be extended to determine higher order dependencies, but the computations might be prohibitive.

Bousfield and Bousfield (1966) have recently revised the model used in calculating the expected number of *ITR*s. The model is the same one used in the measurement of predetermined organization. It will be remembered that a distinction is made between response recall and the sequential characteristics of the words recalled. They argue that the calculation of the chance value should be based on the assumption "that the given recall sequence is a random sample from among all possible orderings of the recalled items [p. 939]." The new formula for calculating the expected value is

$$E(ITR) = \frac{c(c-1)}{hk} \qquad [10]$$

where c is the number of items common to the two recalls. It can be readily seen that the value of $E(ITR)$ is always less than one.

One limitation of the deviation measure noted previously was that the maximum amount of clustering possible is not taken into account. Fagan (1968) has recently shown how the deviation measure can be converted into a ratio measure, and this converted measure can be used when it is desirable to use a proportion measure.

Comparison of the measures. Puff and Hyson (1967) have compared the deviation measure, using the limited-availability model, with the *SO* measure. Over 20 trials of free-recall learning they obtained virtually identical results with the two measures. Correlation coefficients between the two measures were calculated for each of the 19 trial pairs. They report that the lowest of these correlations was .90. They also calculated the correlation between the mean values (across subjects) of the two measures over the 19 trial pairs. The value of this coefficient was .94. The correlation between the observed number of *ITR*s and the value of *SO* was .92 across the 19 trial pairs. Unfortunately, they do not report correlations of the measures with the number of words recalled. Therefore, it is unclear whether or not the biases apparent for the ratio measure of clustering (*RR*) are also present in the ratio measures of subjective organization, for example, the *SO* measure.

Independent measures of organization. It would be desirable, of course, to obtain independent estimates of organization and recall. All of the measures discussed above estimate both variables from the same set of data, namely, the words which the subject recalled. A promising technique has recently been developed by Seibel [8] at the Pennsylvania State University which permits such independent estimates. With this method, a subject is given a matrix, that is, a blank sheet of paper with, for example, 10 rows and 10 columns, prior to the presentation of the stimulus list. The subject is instructed to write each word, as it appears, wherever he wants to in the matrix. After the last word has been presented and recorded, the matrix is removed, and the subject free recalls as many words as he can on another sheet of paper. At the beginning of the next trial, the recall sheet is removed, and the subject is given a new matrix. While this method allows the subject additional, unsystematic study time, it has the substantial advantage of providing an estimate of organization independent of the recall data. While any of the conventional measures of organization can be used, the data in the matrix also permit one to measure the consistency of organization independent of sequence and the degree to which each subject utilizes his own subjective organization in his recall tests.

A related procedure has been used by Mandler (1967; Mandler & Pearlstone, 1966),

[8] R. Seibel, personal communication, January 1968.

and it has the advantage of reducing the amount of unsystematic study time inherent in Seibel's procedure. Mandler has subjects sort lists of words into stacks so that only the top word of each stack is showing. While Mandler asks his subjects to recall the words only after they have reached a predetermined sorting criterion, it would be a simple matter to ask the subjects to recall the words after each sort. The present author has satisfactorily used this technique in an unpublished study.

Both of these techniques have one major advantage over the other measures of subjective organization. Both the *SO* and the deviation measures are based on pairwise contingencies, and in some instances these contingencies may provide only a crude index of the organization actually present. For example, it is possible that a unit of four words will be recalled together on every trial, but if the words are recalled in a different order on successive trials, both measures will underestimate the organization present. The two measures discussed in this section, however, permit the investigator to look at larger size subjective units.

Units of Analysis

The response unit selected for measuring the amount of material recalled can frequently influence the conclusions drawn from the data. The unit typically used in verbal learning studies is some experimentally defined single item such as a word, trigram, etc. While the use of such a unit is somewhat arbitrary, there is at least some face validity for selecting a word as the basic unit of measurement. However, a distinction can be made between such experimentally defined or nominal unit and the functional unit actually used by the subject. Tulving (1968) refers to the former as E-units and the latter as S-units. In any situation in which the subjects group the words together into some type of higher order units, for example, in free recall, there will be a discrepancy between E-units and S-units. While there are a number of situations in which this discrepancy is of little interest (cf. Tulving, 1968), there are many others in which important theoretical considerations depend on the unit selected. This is particularly true when one is concerned with the amount of material recalled or with the capacity of memory. For example, Tulving and Patkau (1962) were concerned with the effects of contextual constraint and word frequency on free recall. When the number of words recalled was used as the dependent variable, there were significant effects due to word frequency and to the interaction between frequency and order of approximation to English (contextual constraint). However, when the number of word sequences corresponding to uninterrupted sequences of words from the stimulus list ("adopted chunks") were used as the dependent variable, neither frequency nor the interaction was significant. Similar types of findings have been obtained by Cohen (1963b) and McNulty (1966).

Deese (1968), on the other hand, has suggested that units more basic than a word should be considered. Basically, he is suggesting a distinctive feature analysis in which the intersection of a set of distinctive source features enables the subject to retrieve or produce the single item, for example, the word. The unit which an investigator uses will depend upon both his theoretical predilection and the type of question he is asking.

The Study of Organization in Free Recall

In this section of the paper, consideration will be given to the variety of experimental studies on the organizational process. Tulving (1968) has distinguished between two types of organization in free recall. The first of these is referred to as primary organization. This type of organization is defined as the consistent discrepancies between input and output orders which are independent of the subject's prior familiarity with the input items. The serial position effect (e.g., Murdock, 1962) and the tendency for subjects to recall the terminal items first (Postman & Keppel, 1968; Shuell & Keppel, 1968) are examples of primary organization. The other type of organization is referred to as secondary organization. This type of organization is dependent upon the subject's prior acquaintance with the items in the stimulus list. Clustering on the basis of meaning would

be an example of secondary organization. Most of the studies to date have been concerned with this latter type of organization.

Variables Influencing Organization

Much of the research on organization in free recall has been concerned with the determination of the variables and conditions which influence the amount of clustering obtained. Recall performance has also been frequently used as a dependent variable. A large variety of independent variables have been used, and some of the more important ones are considered below.

Number of categories. A number of studies have investigated the effect of using varying numbers of categories for organizing the words in the stimulus list. The main purpose of these studies has been to determine if there is an optimum number of categories. Much of the more recent work has been concerned with the limits of memory and the process of chunking as suggested by Miller (1956). While the analogy between categorizing and chunking or recoding as discussed by Miller (1956) is not exact, especially when nonexaustive categories are used, it has been frequently assumed that the same process is being studied (cf. Mandler, 1967). Unfortunately, all of the studies investigating number of categories have held list length constant while varying the number of categories used. As a result, number of categories has been confounded with the number of words per category since the number of words per category must decrease as the number of categories is increased. Consequently, the results of these studies are somewhat equivocal.

The effect of varying numbers of categories appears to be dependent, at least in part, on the length of the list and on whether or not recall is cued (Dallett, 1964; Tulving & Pearlstone, 1966). In general, the relationship between recall and number of categories appears to be a direct one when cued recall is used and an inverse or curvilinear one when noncued recall is used. While it is not possible to presently separate the effects of number of categories and number of words per category, Earhard's (1967a) data indicate that at least for cued recall the use of categorized lists is effective only when the number of words per category is fewer than six or seven items.

With regard to measures of organization, however, the data are somewhat equivocal. Bousfield and Cohen (1956a) and Mandler (1967) found a direct relationship between clustering and number of categories used. Dallett (1964), however, found no consistent relationship for 12-item lists and an inverse relationship for 24-item lists. It is quite possible that these differences result from different measures of organization being used in the various studies. The Bousfield and Cohen and the Mandler studies used the *RR* measure while Dallett used the deviation measure (O − E). If each category is viewed as a sublist, then when the number of categories is increased, the length of each sublist is correspondingly decreased. The *RR* measure makes some provision for the maximum amount of clustering which can be obtained, although it has already been noted that this adjustment is somewhat limited. The deviation measure, however, does not consider the maximum amount of clustering possible. When only a small number of words per category is used, clustering or organization may be at or near maximum; however, when category size is increased, that is, number of categories decreased, the value of the deviation measure may increase merely because of the increase in the maximum number of repetitions which can occur.

Cohen and Bousfield (1956) used a two-level list in which each of the four main categories in a 40-item list could be divided into two subgroups (e.g., animals: felines and canines). The results were analyzed on the basis of both a four- and eight-category list and compared with the results of earlier experiments in which a four- or eight-category, single-level list had been used. For the eight-category analysis of the dual-level list the results were very similar to the results of the earlier experiment using an eight-category, single-level list. For the four-category analysis, however, recall was higher for the dual-level list, although *RR* was about the same in the two experiments.

Several studies have been concerned with the relationship between the number of categories used in a classification task and free-

recall performance. In her 1952 dissertation, Mathews (1954) had subjects classify names of 24 famous people into either two, three, or six categories such as poet, scientist, etc. Ten minutes after completing the classification task the subjects were given 5 minutes to recall as many names as they could. The mean number of words recalled was directly related to the number of categories used, although she mentions in the discussion that another group receiving a one-category list recalled just as many words as the six-category group. Later, she showed (Helson & Cover, 1956) that recall was facilitated if subjects used specific, rather than general, categories for classifying the names; each of the specific categories was a subcategory of one of the general ones.

Mandler (1967; Mandler & Pearlstone, 1966) has used a similar procedure with unrelated words. The subjects were asked to sort from 52 to 100 "unrelated" words into various classifications of their choosing; typically the subjects were asked to use from two to seven categories. The same words were sorted on successive trials until the subject achieved two identical sorts. Then, he was asked to free recall as many of words as he could remember. In general, both the mean number of words recalled and the amount of clustering (defined in terms of the subject-defined categories) were directly related to the number of categories used. Unfortunately, however, these data should probably be interpreted with some caution since only about 55% to 60% of the subjects provided useful data.

Exhaustive versus nonexhaustive categories. Exhaustive categories are those whose examples exhaust, or nearly exhaust, all of the items generally subsumed under the category label, for example, north, east, south, and west. Nonexhaustive categories, on the other hand, are those whose examples only partially exhaust the members included in the category, for example, dog, lion, horse. Cohen (1963a, 1963b) has shown that the mean number of categories recalled (i.e., represented by the words recalled) is the same for the two types of categories; however, significantly more words are recalled for the exhaustive than for the nonexhaustive categories. Interitem associate strength was significantly related to within-category recall (i.e., recall of other words in a category given that at least one of the words is recalled) but was not related to category recall.

Blocked versus random presentation. Blocked presentation refers to the experimental situation in which all members of a category are presented contiguously in the stimulus list, for example, all of the examples of one category are presented before those of another category are presented. The order of categories and the order of items within each category, however, can still be varied when more than one trial is used. All of the early studies of clustering used random presentation in that the words in the stimulus list were presented in a random order, and sometimes restrictions were placed on the number of examples of a given category that could appear together.

Blocked presentation is frequently considered to be more effective than random presentation for helping the subject perceive the categorized nature of the list. This is thought to be particularly true for lists comprised of low-frequency associates to the category name and for categories with only a few items. However, the results of various studies by Cofer and his associates (Cofer, 1967; Cofer, Bruce, & Reicher, 1966) raise some doubts as to the validity of this interpretation. These studies indicate that blocked presentation augments clustering to at least equivalent degrees in lists of high and low taxonomic frequency; there is some indication that this effect may be greater for lists of high taxonomic frequency. Also, blocked presentation facilitated recall only in the high-frequency list.

Dallett (1964) obtained both superior recall and superior clustering with blocked presentation, but the difference between blocked and random presentation interacted with the number of categories in the list such that the largest difference between the two methods was obtained for the condition in which there were three words per category. Puff (1966) used lists containing 10 words from each of three categories. The lists were constructed so that there were either 0, 9, 18, or 27 category repetitions in the order of presentation

which the subject received; the first of these conditions is the typical random presentation while the latter is the typical blocked method of presentation. With one exception, both recall and clustering were directly related to the number of repetitions in the stimulus list.

Thus, blocked presentation appears to facilitate both clustering and recall. However, the facilitation for clustering may be partly due to the fact that all members of certain categories appear in the most favorable positions, that is, the first and last serial positions with the terminal items tending to be recalled first (Postman & Keppel, 1968; Shuell & Keppel, 1968). In an unpublished study, Cohen [9] compared blocked and random presentation of a 70-item list comprised of 20 categories of three or four items each. While the mean number of words recalled was equivalent for the two methods of presentation, there was a significant tendency for the blocked presentation to increase the number of words recalled per category while decreasing the mean number of categories represented in recall. Thus, it is possible that at least under certain conditions, blocked presentation may facilitate the coding or organization of the predefined categories while decreasing the likelihood that stable intercategory associations will be developed.

Changes in organization and recall as a function of time. Changes in performance over time can occur either in the presence or absence of practice. In the presence of practice the stimulus list is presented for a series of trials. Recall or test trials may be interspersed between the study trials, or there may be a single recall trial after a given number of presentations. Both modes of presentation have been used in the study of clustering. Bousfield and Cohen (1953) presented a four-category, 40-item list for either one, two, three, four, or five presentations prior to a single 10-minute recall period. Both mean recall and mean *IR* were directly related to the number of presentations. The mean number of categorical intrusions, however, was inversely related to the number of presentations.

Several studies (Bousfield, Berkowitz, & Whitmarsh, 1959; Marshall, 1967b; Robinson, 1966; Shuell, 1968) employing the alternate study-recall procedure have shown that clustering (several different measures were used in the various studies), mean recall, and the mean number of categories recalled (as determined by the words recalled) increase progressively as a function of trials. Once again, there is a tendency for categorical intrusions to decrease over trials indicating that list differentiation is directly related to the number of trials.

Several studies (e.g., Bousfield, Puff, & Cowan, 1964; Tulving, 1962b) have shown that subjective organization increases as a function of trials. While the amount of organization obtained on the first few trials is at a chance level, the difference between the obtained and expected organization increases progressively over trials. The chance level for the *SO* measure (Tulving, 1962b) was calculated from data obtained from statistical subjects. For the *ITR* measure, of course, Bousfield, Puff, and Cowan (1964) used their model to compute the expected value. Tulving, McNulty, and Ozier (1965) obtained a positive relationship between *SO* and the number of lists which the subjects had learned. Mayhew (1967) had subjects learn two lists on each of 3 consecutive days and found a significant increase in *SO* both within and between experimental sessions. This increase in *SO* proceeded at a faster rate if the subjects were told to organize the words during recall than if they were given standard FR instructions.

The question can be raised as to whether the single-trial or the multitrial procedure is the better method for studying organization in free recall. Cofer (1967) has argued that the single-trial situation is more representative of free recall as it appears in daily life. On the other hand, Tulving (1968) argues in favor of the multitrial procedure. It must be realized, of course, that somewhat different types of questions are being asked in the two situations. The single-trial procedure usually deals with well-defined categories and provides a useful paradigm for studying the effects of a variety of independent variables. The multitrial procedure, especially when it is used to investigate subjective organization, is

[9] B. H. Cohen, personal communication, September 1968.

more concerned with the development of stable organization in the absence of any well-defined relationships. Both situations can provide us with useful information on organization and memory, although it is important to keep in mind the type of question which is being asked.

In the absence of practice, a retention situation exists. Brand and Woods (1958) compared the retention of a categorized list at 1-, 2-, and 3-week intervals. In one condition, a single group was tested at each of the retention intervals. In the other condition, independent groups of subjects were tested at just one of the three intervals. For the independent groups there was a tendency for both *RR* and recall to decrease as a function of time. For the group receiving repeated testing, there was an initial drop at the first retention interval, but there was no further decrease after that. It should probably be noted that subjects who did not demonstrate clustering above a chance level on an initial recall test were excluded.

Using "unrelated" words, Mandler (1967) obtained a negatively accelerated retention function over a 15-week period. The asymptote, approached about the fourth week, was between 20 and 30%. The *RR* measure (as determined from the *S*-defined categories) decreased from about .54 at one-half week to about .24 after 14 weeks; the value of .24 still exceeds the chance value of .16.

In a series of studies, Cofer and his associates (Cofer, 1967; Cofer et al., 1966; Gonzales & Cofer, 1959) have investigated changes in clustering and recall from an immediate-recall test to a second recall test 5 minutes later. In general, there was an increase in clustering and a decrease in recall. The clustering obtained on the second test is significantly greater than the clustering obtained in a control group which waited an equivalent amount of time but did not have the interpolated recall test (Cofer et al., 1966).

Several studies (Gonzalez & Cofer, 1959; Holroyd & Holroyd, 1961; Shuell, 1968; Winograd, 1968) have used categorized lists for studying retroactive inhibition (RI). In general, RI was obtained in all of the studies. Retroactive inhibition has also been demonstrated in free recall with unrelated words (Postman & Keppel, 1967; Shuell & Keppel, 1967; Tulving & Thornton, 1959). Perhaps the most interesting finding, however, is that significantly fewer words per category are recalled when the two lists contain members representing the same categories than when different categories are represented by the words in the two lists (Shuell, 1968; Winograd, 1968). At the present time, it is not clear whether this differential recall is the result of increasing the number of words per category or intracategory interference. Category recall, on the other hand, was significantly poorer when different categories were represented in the two lists.

Associated relatedness. The role of associative relationships in free recall is a complex and hotly debated one. A demonstration experiment by Jenkins and Russell (1952) showed that associative relationships, as well as categorical relationships, can influence clustering.[10] Twenty-four stimulus-response pairs of words from the Kent-Rosanoff word list were presented in a random order with the restriction that a response word could not follow its stimulus. Clustering was said to occur when the two words in a given pair occurred together in recall. The occurrence of arbitrary pairs, that is, the stimulus word followed not by its own response but by another, specified and randomly selected response from the list, was used as a chance base line against which the occurrence of both forward and backward associations was compared. There was a decided tendency to recall the pairs of words together in the stimulus-response order. The reverse associations, or recall of the pairs in the response-stimulus order, occurred significantly more frequently than the arbitrary pairings but significantly less frequently than the forward sequences.

The occurrence of associative clustering in both the forward and reverse direction has been shown to be a positive function of the

[10] While the publication date on this paper predates Bousfield's 1953 paper, the authors include a footnote referring to a paper which Bousfield presented at the American Psychological Association meeting in 1951. This 1951 paper apparently reports the results later presented in the frequently cited 1953 paper.

associative strength between the two words in the pair. With adult subjects (Jenkins et al., 1958), clustering was above the chance level in the pairs of low associative strength. However, with children (Wicklund et al., 1965), clustering occurred at a chance level in the pairs of low associative strength. In addition, Wicklund, et al. examined the degree of forward and reverse clustering for each individual pair and concluded that the tendency to cluster is strongly related to the presence of bidirectional associative strength.

Associative clustering also occurs when the two words of a pair are not directly related but are related to a third word not appearing in a stimulus list, that is, *A* elicits *B*, *B* elicits *C*, but *A* does not elicit *C*; *A* and *C* are the words presented (Cramer, 1965; Shapiro, & Palermo, 1967). Several studies (Bousfield, Whitmarsh, & Berkowitz, 1960; Marshall, 1967a, 1967b) have shown that associative clustering can be predicted by the extent to which two words elicit common associates, and the results of the study by Bousfield, Steward, and Cowan (1964) seem to indicate that clustering in a categorized list can be predicted better with an index of associative overlap, that is, the extent of which the words elicit common responses, than by means of an index of interitem associations, that is the extent to which the items in the list elicit one another. In general, there is a tendency for associative clustering to increase as a function of study-tests trials (Marshall, 1967a, 1967b). However, this relationship appears to be partly dependent upon the sex of the subject, the direction of association, and the normative strength of the associative pairs (Rosenberg, 1966). With mediated clustering, there appears to be an increase in clustering only when blocked presentation is used (Shapiro & Palermo, 1967).

In a series of frequently cited studies, Deese (1959a, 1959b, 1960, 1961) demonstrated a positive relationship between interitem associative strength (IIAS) and free-recall performance. This relationship existed for both intrusions and the probability of a word being correctly recalled. The same relationship has also been found with children (Hess & Simon, 1964; Simon & Hess, 1965). Also, the probability that a given word will be recalled is directly related to the number of other words in the list which elicit the given word as a response (Bodin, Crapsi, Deak, Morday, & Rust, 1965; Rothkopf & Coke, 1961).

Marshall and Cofer (1963) have considered various indexes of associative relatedness. They note that although these various indexes are able to predict performance in a variety of situations, there is presently no satisfactory theoretical interpretation of the way in which these associative relations function. Deese (1968) has recently argued that associations are only manifestations or indexes of underlying patterns of relation and therefore cannot cause organization as such. Nevertheless, it is clear that preexperimental relations influence performance on a free-recall task. A better understanding of the way in which these relationships operate must await future research.

Bases of Organization

Factors which mediate clustering. Most of the studies investigating clustering have used stimulus lists comprised of concrete nouns which are clearly members of some abstract class, for example, animals, countries, vegetables, etc. Cohen et al. (1957) collected cultural norms for use in clustering studies. Four hundred undergraduates at the University of Connecticut responded to 43 different category names with the first four specific associates they could think of for each class of items. These responses were then ranked on the basis of their frequency of occurrence to each of the category names. Several studies (Bousfield, Cohen, & Whitmarsh, 1958; Bousfield, Steward, & Cowan, 1964; Cofer et al., 1966; Holroyd & Holroyd, 1961) have demonstrated both superior recall and superior clustering in lists comprised of high-frequency responses from these norms than for lists comprised of low-frequency responses. Unfortunately, most of these studies have confounded normative frequency with frequency of usage as determined from the Thorndike-Lorge (1944) count. However, one study (Bousfield, Steward, & Cowan, 1964) indicates that clustering, but not recall, is apparently unaffected by a fairly wide variation in Thorndike-Lorge frequency as long as the members of the categories are equated on normative frequency.

A variety of attempts have been made to determine if clustering could be obtained on other dimensions. Studies which have varied response dominance (Bousfield & Puff, 1964), structural characteristics of geometric designs (Bousfield et al., 1959), and structural characteristics of words (Pellet, 1957) have met with limited success. Attempts to obtain clustering on the basis of form class (Cofer & Bruce, 1965; Gonzales & Cofer, 1959), synonyms (Cofer, 1959), and the semantic differential (Cowan, 1964) have generally been unsuccessful. Bousfield and Cohen (1956b) demonstrated that for both males and females there is a greater recall and greater clustering of words representing interests of their own sex than of words representing interests of the opposite sex.

Another way of determining various factors which can mediate clustering is to look at the various clusters of words which consistently occur in studies of subjective organization. Mandler (1967) and Tulving (1968) have presented some interesting examples of these types of clusters. The basis of organization in some of these clusters is quite obvious, while in others it is not at all obvious. Probably a whole variety of factors determine the basis of organization which the subject will use in studies of this nature. When an obvious form of organization is provided, for example, studies of clustering, the subject will usually use the organization provided. When no obvious form of organization is provided, for example, studies of subjective organization, the subject will find more subtle forms of organization to use. While subjective organization is primarily based upon intrasubject organization, there appears to be substantial commonality among subjects in the way the recalls are organized (Earhard, 1967b; Tulving, 1962b, 1965). This indicates that the subjects are using relationships present in the stimulus list even though the experimenter has assumed the words are "unrelated." While the nature of this organization is somewhat subtle, the subjects are nevertheless able to locate and make use of it in similar ways.

The role of context. Various contextual factors present during the presentation of the stimulus list can also influence the amount and type of organization which occurs. Gonzalez and Cofer (1959) did a series of experiments on the effect of context on clustering. Their technique was to present pairs of words rather than individual words, although the subjects were usually asked to recall only one member of the pair. Five types of effects are discussed.

1. Specificity effect. Forty unrelated adjectives were selected to modify 40 nouns representing four different categories. Clustering of the nouns was reduced to the chance level.

2. Mediational effect. When a list of uncategorized nouns, which did not cluster when presented alone, was modified by adjectives representing discreet categories, which did cluster when presented alone, the recall of the nouns alone revealed significant clustering when scored on the basis of the adjectives which modified them during presentation. The reverse was also found to be true; uncategorized adjectives would cluster on the basis of the nouns which they modified. Similar effects were obtained using either normal word order, adjective-noun, or reverse word order, noun-adjective.

3. Mediated facilitation effects. When clustering adjectives and clustering nouns were presented so that the adjective categories and the noun categories were congruent, clustering and recall of the nouns alone were facilitated. Similar effects were not obtained with adverb and verb combinations or with four word combinations, for example, adjective-noun-verb-adverb.

4. Mediated conflict effects. When clustering adjectives and clustering nouns were presented so that the adjective categories and noun categories were not congruent, that is, adjectives from a given category modified nouns from all of the categories, clustering and recall were impaired even though for the specific pairs the adjective appropriately modified the noun.

5. Inappropriate modification effects. When the adjective modifying a given noun is inappropriate, for example, leafy dog, clustering and recall were impaired, and the effect was more pronounced when the adjective-noun pair was recalled than when the noun only was recalled. Later, association data were collected for these various situations (Cofer,

1960). Except for the mediated facilitation condition, the results of the various conditions could be fairly well accounted for on the basis of changes in the associative relationships involved. In a more recent study, Cofer (1968) has shown that clustering can be disrupted by presenting nouns in the context of a sentence.

Thus, it is clear that context can influence organization. Whether this effect results from a modification of the associative relationships (Cofer, 1960) or from the fact that context determines the dimensions or distinctive features(Deese, 1968) which the subject selects as the bases of his organization is not entirely clear at present. It is quite possible that these two interpretations are in actuality referring to the same basic process.

Organization and Recall

Generally speaking, there is a positive correlation between measures of organization and the number of words recalled. However, there is evidence (cf. Cofer, 1967, 1968; Cofer et al., 1966; Dallett, 1964) that at least under certain conditions measures of organization and recall can vary independently. In addition, two studies (Laurence, 1966; Mandler & Stephens, 1967) have indicated that at least for subjective organization different results are obtained with children than are obtained with adult subjects. Laurence (1966) obtained a direct relationship between age and recall, but the values of *SO* were virtually the same for all of the age groups, although the values obtained by college students were significantly higher than those for the children. While these results are contrary to the results of two studies on clustering (Bousfield, Esterson, & Whitmarsh, 1968; Rossi, 1964) which found a direct relationship between clustering and age, they do raise an important question. The extent to which this variation may result from limitations of the measures or from other factors is not entirely clear at the present time. Also, it should be realized that the amount recalled will depend on the unit of analysis which is used.

A number of writers (e.g., Mandler, 1967; Tulving, 1968) have suggested that recall is dependent upon organization. Probably the best empirical evidence in support of such a position is the study on alphabetic organization (Tulving, 1962a). However, the situations discussed above in which organization and recall can vary independently raise some questions as to the validity of this interpretation. The explication of the relationship between organization and recall is still in need of further investigation.

While only one study (Earhard, 1967a) has made a direct comparison between categorized and unrelated lists, several studies (Dallett, 1964; Earhard, 1967a; Mathews, 1954) have used a condition in which all of the words represented the same category, and one study (Tulving & Pearlstone, 1966) used a condition in which each word in the list represented a different category. In general, there has been superior recall in the categorized lists, although there is some indication (Dallett, 1964; Earhard, 1967a; Mathews, 1954) that this facilitation may depend on whether recall is cued or noncued and on the number of words per category in the stimulus list.

Earhard, (1967a) included two free-recall (noncued) conditions in her study. One of these conditions consisted of words which all began with the same letter, and the other consisted of words which all began with different letters. Since alphabetic organization is not effective unless the subjects are told about the nature of the list (Tulving, 1962a), the later condition is equivalent to a list of unrelated words. There is virtually no difference between the two conditions over the 20 learning trials, although this finding cannot be generalized to other studies of categorical clustering since the subjects were not told about the nature of the list. The mean number of words recalled per trial in the cued-recall conditions (the subjects in these conditions were instructed as to the nature of the list) was superior to that in the free-recall conditions for 1, 2, 3, and 4 words per category, equivalent for 6 and 8 words per category, and inferior for 12 words per category.

Theoretical Issues

Various attempts have been made to develop a theoretical explanation for the clustering phenomenon. None of the present inter-

pretations is completely satisfactory, that is, none can explain all of the relevant data. Some of the major theoretical positions and issues which have developed from the study of organization in free recall will be discussed in this section of the paper.

Perception of the Superordinate

The early Bousfield papers (e.g., Bousfield, 1953; Bousfield & Cohen, 1953) interpreted clustering in terms of (*a*) the habit strength of an item which derived from the reinforcement it received before and during the experiment and (*b*) a relatedness increment which was taken as a hypothetical increment added to habit strength. It was hypothesized that when a word is recalled an increment is added to other words which are related to it. The strength of these two factors, subject to oscillation, determines the probability of whether the next item recalled will be from the same or from a different category and, therefore, determines the amount of clustering obtained. The finding that clustering changes as a function of the stage of recall (Bousfield, 1953; Bousfield & Cohen, 1953, 1955; Cohen & Bousfield, 1956) was explained by making the assumption that the latent period is an index of habit strength, although as Tulving (1968) has pointed out, these changes may be partly the result of such factors as the tendency to recall terminal items first.

The relatedness increment was assumed to be related to Hebb's (1949) conception of the development of superordinant perceptions. According to this conception, the repeated perception of a set of related items results in the formation of a new superordinant system. Eventually the activation of a single perception unit will suffice to arouse the superordinant system which, in turn, will facilitate its subordinant units. Phrased in terms of the experimental situation of the Bousfield studies, the perception or recall of a single word will tend to activate the superordinant systems which correspond to the category represented by the word. Once this superordinant system is activated, it will tend to facilitate the perception and recall of other words belonging to the same category.

In more recent papers (e.g., Bousfield & Bousfield, 1966; Bousfield & Puff, 1965; Bousfield, Steward, & Cowan, 1964), Bousfield has tended to interpret clustering more in terms of associative factors than in terms of the superordinate mechanism suggested earlier. Recently, Deese (1968) has suggested that while this supraordinate-subordinate structure is too elementary to describe adequately the complexities or the relationships involved, it is probably closer to the truth than the alternative interpretations presently available.

Coding Interpretation

Somewhat related to the foregoing position is the coding interpretation of clustering. This position is based primarily on Miller's (1956) concept of chunking. The basic notion is that the subject recodes the words in the stimulus list into the respective categories and then stores the category label (or some representation of it) in memory. When it comes time to recall the words he recalls the category as a whole. This interpretation is somewhat supported by the finding that if at least one word in a category is recalled, the average number of words recalled per category is remarkably consistent over a fairly wide range of conditions. (Cohen, 1966; Tulving & Pearlstone, 1966). Whether recall is actually a two-level search process in which the subject first searches for, or tries to recall, a category, for example, its name, and then tries to recall the words within that category as some have suggested (Cohen, 1966; Shepherd, 1966) is something on which little data are presently available. The best evidence to support such an interpretation comes from a study by Segal (1969). In this study a categorized list of words was used which included the category names. There was a distinct tendency for the category name to be recalled prior to the specific instances of that category.

One of the ways in which the coding interpretation differs from the superordinate interpretation is that the latter is made in terms of the habit strength of individual items. This habit strength can be augmented, apparently via some sort of generalization mechanism, by the recall or perception of related items. Thus, this augmentation can

occur either at the time of presentation or at the time of recall. The coding interpretation, on the other hand, assumes that the related words are encoded at the time of presentation into their respective categories, or chunks, and that during recall these chunks are recalled as a unit.

It is worthwhile to note one limitation to the notion that the category name as such is used as the coding cue. If instances of a nonexhaustive category, for example animals, are used in the stimulus list, it is rather difficult to explain why the subject recalls only those members of the category presented. If the instances were coded only as animals, there would be no way for the subject to differentiate between those animals which were presented and those which were not. Thus, it seems unlikely that in the strict sense the category name is used as the coding cue, although it is certainly possible that this may be one facet of the cue used by the subject. This is perhaps one explanation why exhaustive categories are better recalled than nonexhaustive categories.

Associative Interpretation

The role of associative relatedness in free recall was discussed in an earlier section. As might be expected, the question has been raised as to whether or not clustering can be explained solely in terms of associative relationships, that is, is it necessary to postulate an additional concept such as superordination or coding? Several studies (Bousfield & Puff, 1965; Field, 1969; Marshall, 1967b) have contrasted the associative and the coding interpretations. In general, the conclusion has been that both associational and coding processes are involved in most free-recall situations (cf. Cofer, 1965, 1966). Marshall (1967b) has shown that when pairs of words are matched for association relatedness, those pairs which are categorized cluster to a greater extent than noncategorized pairs at low and middle levels of associative relatedness but not at high levels.

It has been suggested (e.g., Deese, 1961) that subjects reconstruct the list at the time of recall. Cofer (1967) has found that when subjects are asked to generate additional items to match list length they apparently make use of both list-name and list-member associations. The fact that subjects do not match list length under typical free-recall procedures has been taken as evidence that subjects edit their recall (Cofer, 1967). The additional finding that differential performance as a function of associative strength is not obtained with a recognition test (Cofer, 1967; Field, 1969) suggests that perhaps associative relationships influence performance by means of varying the likelihood that a given word is generated by the subject. Thus, it is possible that during presentation the subject learns certain characteristics about the list in addition to some of the specific words and then uses this information during recall to generate words and to make decisions as to whether or not a particular word was on the list.

Deese (1968) has recently argued that associations cannot *cause* organization, although Postman (1968) has suggested that the term association can be used in at least six different ways. It is possible, of course, that both associative and categorical relationships can be explained in terms of some common underlying process.

Subjective Organization

Subjective organization does not actually offer a separate theoretical interpretation of clustering, although a number of important theoretical contributions have resulted from the study of subjective organization. It was suggested earlier in this paper that the same behavioral processes are being studied in all three paradigms. The main difference is in the way the problem is approached, although this difference may prove to be very important to the future understanding of the processes involved. Perhaps the most important contribution which the study of subjective organization has provided is the demonstration that there are discrepancies between the learning situation as defined by the experimenter and the learning situation as perceived by the subject.

Another issue which has been raised is the distinction between availability and accessibility (Tulving & Pearlstone, 1966). According to this interpretation, items can be available for recall, that is, in storage, but

unless sufficient retrieval cues are present at the time of recall the items will not be recalled. Several studies (Dong & Kintsch, 1968; Tulving & Osler, 1968; Tulving & Pearlstone, 1966) have demonstrated that recall can be increased if subjects are provided with relevant cues at the time of recall. However, Slamecka (1968), in a series of experiments, was unable to obtain facilitation by providing the subjects with varying numbers of items from the stimulus list as retrieval cues. Tulving (1967) has suggested a limited-capacity retrieval system in which the main limitation in memory is not in the storage of items but is determined by the number of accessible memory units. The contents or nature of the units is deemed to be of little importance.

The occurrence of intertrial repetitions above a chance level, as discussed in the section on measurement, provides evidence for the existence of higher order units in free recall. However, Tulving (1966; Tulving & Osler, 1967) obtained additional verification in a series of transfer studies. When the second list contains items from the first list, negative transfer is obtained relative to a control group which receives completely different items in the two lists. This effect is obtained both when subjects are transferred from a shorter list to a longer list (Tulving, 1966) and when they are transferred from a longer list to a shorter list (Tulving & Osler, 1967). Similarly, differential transfer has been obtained with categorized lists when the words in successive lists represent either the same or different categories (Roberts & Smith, 1966; Shuell, 1968). Both studies obtained superior recall and clustering when different categories are represented in the successive lists.

Throughout the present paper an attempt has been made to evaluate the research on organization in free recall. The quantification of organization has been achieved, and much has been learned about the relationship between organization and memory. However, many questions remain unanswered, and some of the directions which future research may take have been suggested at various points throughout the paper. A number of writers (e.g., Mandler, 1967; Tulving, 1968) have suggested that recall is dependent upon organization. While this suggestion appears reasonable, it must be admitted that the evidence which actually supports such a position is not overwhelming at present. Perhaps the use of measures of organization which are independent of recall will permit us to obtain additional data relevant to this important question.

It should be remembered that clustering is basically an output phenomena from which we infer some sort of organizational process on the part of the subject. Since organization is present in the stimulus list, and this is true for subjective organization as well as for the more traditional clustering studies, the inference is basically concerned with the way in which the subject uses this organization.

Finally, it should be noted that the locus of the organizational process has not been determined. The organization which is observed during recall could occur at the time of input, that is, presentation of the stimulus list, or at the time of retrieval, that is, when the subject is recalling the words. Slamecka (1968) has suggested that the items are stored independently along with a general representation of the list structure. At the time of recall this general representation forms a retrieval plan which then guides the subject's search for the specific items. However, there is increasing evidence (Anisfeld & Knapp, 1968; Rohwer, Shuell, & Levin, 1967; Tulving & Osler, 1968) that encoding or organization must occur at the time of presentation in order to be effective.

REFERENCES

ANISFELD, M., & KNAPP, M. Association, synonymity, and directionality in false recognition. *Journal of Experimental Psychology*, 1968, 77, 171–179.

BODIN, A. M., CRAPSI, L. A., DEAK, M. R., MORDAY, T. R., & RUST, L. D. Prediction of free recall from word-association measures: A replication. *Journal of Experimental Psychology*, 1965, 69, 103–105.

BOUSFIELD, A. K., & BOUSFIELD, W. A. Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports*, 1966, 19, 935–942.

BOUSFIELD, W. A. The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, 1953, 49, 229–240.

BOUSFIELD, W. A., BERKOWITZ, H., & WHITMARSH, G. A. Associative clustering in the recall of

minimally meaningful geometric designs. *Canadian Journal of Psychology,* 1959, **13**, 281–287.

BOUSFIELD, W. A., & COHEN, B. H. The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. *Journal of Psychology,* 1953, **36**, 67–81.

BOUSFIELD, W. A., & COHEN, B. H. The occurrence of clustering in the recall of randomly arranged words of different frequencies-of-usage. *Journal of General Psychology,* 1955, **52**, 83–95.

BOUSFIELD, W. A., & COHEN, B. H. Clustering in recall as a function of the number of word-categories in stimulus-word lists. *Journal of General Psychology,* 1956, **54**, 95–106. (a)

BOUSFIELD, W. A., & COHEN, B. H. Masculinity-femininity in the free recall of a categorized stimulus word list. *Perceptual and Motor Skills,* 1956, **6**, 159–165. (b)

BOUSFIELD, W. A., COHEN, B. H., & WHITMARSH, G. A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. *Psychological Reports,* 1958, **4**, 39–44.

BOUSFIELD, W. A., ESTERSON, J., & WHITMARSH, G. A. A study of developmental changes in conceptual and perceptual associative clustering. *Journal of Genetic Psychology,* 1958, **92**, 95–102.

BOUSFIELD, W. A., & PUFF, C. R. Clustering as a function of response dominance. *Journal of Experimental Psychology,* 1964, **67**, 76–79.

BOUSFIELD, W. A., & PUFF, C. R. Determinants of the clustering of taxonomically and associatively related word pairs. *Journal of General Psychology,* 1965, **73**, 211–221.

BOUSFIELD, W. A., PUFF, C. R., & COWAN, T. M. The development of constancies in sequential organization during repeated free recall. *Journal of Verbal learning and Verbal Behavior,* 1964, **3**, 489–495.

BOUSFIELD, W. A., & SEDGEWICK, C. H. W. An analysis of sequences of restricted associative responses. *Journal of General Psychology,* 1944, **30**, 149–165.

BOUSFIELD, W. A., STEWARD, J. R., & COWAN, T. M. The use of free associational norms for the prediction of clustering. *Journal of General Psychology,* 1964, **70**, 205–214.

BOUSFIELD, W. A., WHITMARSH, G. A., & BERKOWITZ, H. Partial response identities in associative clustering. *Journal of General Psychology,* 1960, **63**, 233–238.

BRAND, H., & WOODS, P. J. The organization of the retention of verbal material. *Journal of General Psychology,* 1958, **58**, 55–68.

COFER, C. N. A study of clustering in free recall based on synonyms. *Journal of General Psychology,* 1959, **60**, 3–10.

COFER, C. N. An experimental analysis of the role of context in verbal behavior. *Transactions of the New York Academy of Sciences,* 1960, **22**, 341–347.

COFER, C. N. On some factors in the organizational characteristics of free recall. *American Psychologist,* 1965, **20**, 261–272.

COFER, C. N. Some evidence for coding processes derived from clustering in free recall. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5**, 188–192.

COFER, C. N. Does conceptual organization influence the amount retained in immediate free recall. In B. Kleinmutz (Ed.), *Concepts and the structure of memory.* New York: Wiley, 1967.

COFER, C. N. Free recall of nouns after presentation in sentences. *Journal of Experimental Psychology,* 1968, **78**, 145–152.

COFER, C. N., & BRUCE, D. R. Form-class as the basis for clustering in the recall of nonassociated words. *Journal of Verbal Learning and Verbal Behavior,* 1965, **4**, 386–389.

COFER, C. N., BRUCE, D. R., & REICHER, G. M. Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology,* 1966, **71**, 858–866.

COHEN, B. H. An investigation of recoding in free recall, *Journal of Experimental Psychology,* 1963, **65**, 368–376. (a)

COHEN, B. H. Recall of categorized word lists. *Journal of Experimental Psychology,* 1963, **66**, 227–234. (b)

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

COHEN, B. H., & BOUSFIELD, W. A. The effects of a dual-level stimulus-word list on the occurrence of clustering in recall. *Journal of General Psychology,* 1956, **55**, 51–58.

COHEN, B. H., BOUSFIELD, W. A., & WHITMARSH, G. A. *Cultural norms for verbal items in 43 categories.* (ONR Tech. Rep. No. 22, Contract Nonr-631(00)) Washington, D. C.: United States Government Printing Office, 1957.

COHEN, B. H., SAKODA, J. M., & BOUSFIELD, W. A. *The statistical analysis of the incidence of clustering in the recall of randomly arranged associates.* (ONR Tech. Rep. No. 10, Contract Nonr-631(00) Washington, D. C. : United States Government Printing Office, 1954.

COWAN, T. M. The semantic differential as a basis for obtaining clustering. Unpublished doctoral dissertation, University of Connecticut, 1964. (*Dissertation Abstracts,* 1965, **25**, 4824–4825)

CRAMER, P. Mediated clustering and importation with implicit verbal chains. *Psychonomic Science,* 1965, **2**, 165–166.

DALLETT, K. M. Number of categories and category information in free recall. *Journal of Experimental Psychology,* 1964, **66**, 1–12.

DEESE, J. Influence of inter-item associative strength upon immediate free recall. *Psychological Reports,* 1959, **5**, 305–312. (a)

DEESE, J. On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology,* 1959, **58**, 17–22. (b)

DEESE, J. Frequency of usage and number of words in free recall: The role of association. *Psychological Reports,* 1960, **7**, 337–344.

DEESE, J. From the isolated verbal unit to connected discourse. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961.

DEESE, J. Association and memory. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

DONG, T. & KINTSCH, W. Subjective retrieval cues in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 813–816.

EARHARD, M. Cued recall and free recall as a function of the number of items per cue. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 257–263. (a)

EARHARD, M. Subjective organization and list organization as determinants of free-recall and serial-recall memorization. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 501–507. (b)

FAGAN, J. F. III Measuring verbal recall II: The ITR score expressed as a ratio. *Psychonomic Science*, 1968, 11, 205.

FIELD, W. H. Coding and association in the recall and recognition of categorized words. Unpublished doctoral dissertation, State University of New York at Buffalo, 1969.

GONZALES, R. C., & COFER, C. N. Exploratory studies of verbal context by means of clustering in free recall. *Journal of Genetic Psychology*, 1959, 95, 293–320.

HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.

HELSON, R. M., & COVER, A. Specificity-generality of classificatory categories as a variable in recall. *Perceptual and Motor Skills*, 1956, 6, 233–236.

HESS, J. L., & SIMON, S. Extra-list intrusions in immediate free recall as a function of associative strength in children. *Psychological Reports*, 1964, 14, 92.

HOLROYD, R. G., & HOLROYD, J. C. Associative clustering in a retroaction paradigm. *Journal of General Psychology*, 1961, 64, 101–104.

JENKINS, J. J., MINK, W. D., & RUSSELL, W. A. Associative clustering as a function of verbal association strength. *Psychological Reports*, 1958, 4, 127–136.

JENKINS, J. J., & RUSSELL, W. A. Associative clustering during recall. *Journal of Abnormal and Social Psychology*, 1952, 47, 818–821.

LAURENCE, M. W. Age differences in performance and subjective organization in the free-recall learning of pictorial material. *Canadian Journal of Psychology*, 1966, 20, 388–399.

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

MANDLER, G., & PEARLSTONE, Z. Free and constrained concept learning and subsequent recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 126–131.

MANDLER, G., & STEPHENS, D. The development of free and constrained conceptualization and subsequent verbal memory. *Journal of Experimental Child Psychology*, 1967, 5, 86–93.

MARSHALL, G. R. Effect of total association and conceptual cohesiveness among words on recall, clustering, and recognition association. *Psychological Reports*, 1967, 20, 29–44. (a)

MARSHALL, G. R. Stimulus characteristics contributing to organization in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 364–374.(b)

MARSHALL, G. R., & COFER, C. N. Associative indices as measures of word relatedness: A summary and comparison of ten methods. *Journal of Verbal Learning and Verbal Behavior*, 1963, 1, 408–421.

MATHEWS, R. Recall as a function of number of classificatory categories. *Journal of Experimental Psychology*, 1954, 47, 241–247.

MAYHEW, A. J. Interlist changes in subjective organization during free-recall learning. *Journal of Experimental Psychology*, 1967, 74, 425–430.

MCNULTY, J. A. The measurement of "adopted chunks" in free-recall learning. *Psychonomic Science*, 1966, 4, 71–72.

MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity to process information. *Psychological Review*, 1956, 63, 81–97.

MURDOCK, B. B., JR. The serial position effect of free recall. *Journal of Experimental Psychology*, 1962, 64, 482–488.

PELLET, B. The use of structural characteristics in the clustering of verbal associates. Unpublished dissertation, University of Connecticut, 1957. (*Dissertation Absracts*, 1957, 17, 2680).

POSTMAN, L. Association and performance in the analysis of verbal learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

POSTMAN, L., & KEPPEL, G. Retroactive inhibition in free recall. *Journal of Experimental Psychology*, 1967, 74, 203–211.

POSTMAN, L. & KEPPEL, G. Conditions determining the priority of new items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 260–263.

PUFF, C. R. Clustering as a function of the sequential organization of stimulus word lists. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 503–506.

PUFF, C. R., & HYSON, S. P. An empirical comparison of two measures of intertrial organization in free recall. *Psychonomic Science*, 1967, 9, 329–330.

ROBERTS, W. A., & SMITH, I. The effects of categorization and categorical repetition on successive recalls of word lists. *Psychonomic Science*, 1966, 5, 225–226.

ROBINSON, J. A. Category clustering in free recall. *Journal of Psychology*, 1966, 62, 279–285.

ROHWER, W. D., JR., SHUELL, T. J., & LEVIN, J. R. Context effects in the initial storage and retrieval of noun pairs. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 796–801.

ROSENBERG, S. Associative clustering and repeated trials. *Journal of General Psychology,* 1966, **74,** 89–96.

ROSSI, E. Development of classificatory behavior. *Child Development,* 1964, **35,** 137–142.

ROTHKOPF, E. Z., & COKE, E. U. The prediction of free recall from word association measures. *Journal of Experimental Psychology,* 1961, **62,** 433–438.

SEGAL, E. M. Hierarchical structure in free recall. *Journal of Experimental Psychology,* 1969, **80,** 59–63.

SHAPIRO, S. L., & PALERMO, D. S. Mediated clustering in free recall. *Journal of Experimental Psychology,* 1967, **75,** 365–371.

SHEPARD, R. N. Learning and recall as organization and search. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5,** 201–204.

SHUELL, T. J. Retroactive inhibition in free-recall learning of categorized lists. Unpublished doctoral dissertation, University of California, Berkeley, 1967.

SHUELL, T. J. Retroactive inhibition in free-recall learning of categorized lists. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7,** 797–805.

SHUELL, T. J., & KEPPEL, G. Retroactive inhibition as a function of learning method. *Journal of Experimental Psychology,* 1967, **75,** 457–463.

SHUELL, T. J., & KEPPEL, G. Item priority in free recall. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7,** 969–971.

SIMON, S., & HESS, J. L. Supplementary report: Influence of interitem associative strength upon immediate free recall in children. *Psychological Reports,* 1965, **16,** 451–455.

SLAMECKA, N. J. An examination of trace storage in free recall. *Journal of Experimental Psychology,* 1968, **76,** 504–513.

THORNDIKE, E. L., & LORGE, I. *The teachers word book of 30,000 words.* New York: Columbia University, Teachers College, 1944.

TULVING, E. The effect of alphabetical subjective organization on memorizing unrelated words. *Canadian Journal of Psychology,* 1962, **16,** 185–191. (a)

TULVING, E. Subjective organization in free recall of "unrelated" words. *Psychological Review,* 1962, **69,** 344–354. (b)

TULVING, E. The effect of order of presentation on learning of "unrelated" words. *Psychonomic Science,* 1965, **3,** 337–338.

TULVING, E. Subjective organization and effects of repetition in multitrial free-recall learning. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5,** 193–197.

TULVING, E. The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 175–184.

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

TULVING, E., MCNULTY, J. A., & OZIER, M. Vividness of words and learning to learn in free-recall learning. *Canadian Journal of Psychology,* 1965, **19,** 242–252.

TULVING, E., & OSLER, S. Transfer effects in whole/part free-recall learning. *Canadian Journal of Psychology,* 1967, **21,** 253–262.

TULVING, E., & OSLER, S. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology,* 1968, **77,** 593–601.

TULVING, E., & PATKAU, J. E. Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology,* 1962, **16,** 83–95.

TULVING, E., & PATTERSON, R. D. Functional units and retrieval processes in free recall. *Journal of Experimental Psychology,* 1968, **77,** 239–248.

TULVING, E., & PEARLSTONE, Z. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5,** 381–391.

TULVING, E., & THORTON, G. B. Interaction between proaction and retroaction in short-term retention. *Canadian Journal of Psychology,* 1959, **13,** 255–265.

WICKLUND, D. A., PALERMO, D. S., & JENKINS, J. J. Associative clustering in the recall of children as a function of verbal association strength. *Journal of Experimental Child Psychology,* 1965, **2,** 58–66.

WINOGRAD, E. List differentiation, recall, and category similarity. *Journal of Experimental Psychology,* 1968, **78,** 510–515.

(Received January 21, 1969)