

In Search of the M-Gram: The Structure of Organization in Free Recall

MICHAEL L. FRIENDLY

York University

This paper reviews three major classes of structural models of episodic and semantic memory which have been proposed to describe the organization of permanent memory. Although semantic memory workers have developed methods for obtaining representations of these models from empirical data, students of episodic memory generally have not. A method is presented for extracting representations of memory structure using output order information from a free recall learning task. It is shown that this method can be embedded within each of the three general models of memory structure to produce graphic representations of those structures. These representations may be obtained for individuals or for group data, and in some cases the nature and extent of individual differences in organizational structure can be characterized. A final section of the paper demonstrates the validity of the method using both real and simulated data, shows that the method is sensitive even to small differences in organizational structure, and illustrates the utility of this approach with an analysis of part-whole transfer.

Within theories of memory, two broad classes of components can be distinguished. Structural components relate to the format, arrangement, and interrelationships among item traces in the memory system. Functional components relate to the processes by which items become encoded, are stored, and subsequently are retrieved. In the language of information retrieval, structural components of memory models describe or represent the data base of the memory system, while functional components describe the routines which operate upon this data base.

Over the last two decades, the notion that the memory data base constitutes a highly organized system has played an increasingly dominant role in theory and research in human memory, particularly in the context of

The author is profoundly grateful to Peter A. Ornstein for his advice, guidance, and interest from the inception of this work through the writing of this paper. The author thanks Martha Friendly for assistance in data collection and editing, Peter Dean and Doug Vipond for thoughtful comments on the manuscript, and Susan Goldberg for cheerfully typing numerous revisions.

This research was begun while the author held a Psychometric Fellowship at Educational Testing Service and Princeton University. Initial research was supported by Educational Testing Service, by Office of Naval Research Contract NOOO 14-67-A-0151-0006, and by National Science Foundation Grant GB 8023X. More recently, this work has been supported by National Research Council of Canada Grant A8615.

Requests for reprints should be addressed to Michael Friendly, Department of Psychology, York University, Downsview, Ontario, Canada M3J 1P3.

multitrial free recall (MFR). At an operational level, organization is said to occur in multitrial free recall when the learner recalls list words in an order which becomes increasingly consistent or stereotyped as learning proceeds. Theoretically, organization in MFR is considered to reflect the grouping of nominally distinct list items into functional memory units in such a way that items linked or grouped together can be recalled more easily than unorganized items. The increasing consistency in output order over trials is assumed to reflect the development of these memory units, recalled as contiguous chunks.

The implicit view in the present paper is that the phenomenon of organization is a manifestation of both a structural component and a set of processes which determines how that structure is utilized to retrieve verbal information. A fully articulated theory of organization, I believe, must elaborate both components in detail. The aims of the present paper, however, are more limited, and are restricted to the structure of information in the memory store. This strategy reflects this writer's strong agreement with Anderson and Bower (1973) that

the most fundamental problem confronting cognitive psychology today is how to represent theoretically the knowledge that a person has: what are the primitive symbols or concepts, how are they related, how are they to be concatenated and constructed into larger knowledge-structures, and how is this "information file" to be accessed, searched, and utilized in solving the mundane problems of daily living. The choice of a representation is central, since how one handles this issue causes widespread effects throughout the remainder of his theoretical efforts. As computer scientists working on problem solving have known for years, a good structural representation of the problem already constitutes half of its solution (Anderson & Bower, 1973, p. 151).

The remainder of this paper is divided into four major sections. The first reviews three theoretical conceptions of the structure of organized memory and some of the empirical methods used to investigate these structures. The second section presents a method for quantifying the structure of organization and determining the contents of memory units, using order of recall information alone. The method provides a way of representing graphically the structure of organization for individual subjects, as well as for groups. Because of the reliance placed on graphical conceptualizations, I have come to call such a representation of the structural base of the memory system a *memory diagram*, or *M-gram* for short. The form of the M-gram is not atheoretical, however, but depends upon the conception of memory structure adopted. Nevertheless, the third section shows that the method can be embedded within, and used to investigate, each of the three conceptions reviewed in the first section. The final section presents some evidence relevant to the validity and usefulness of the method.

APPROACHES TO STRUCTURE

This section examines several general characterizations of the manner in which information is structured in the memory system. While the motivation here is directed toward developing an M-gram to represent organization in free recall, this review will draw on theories of semantic memory as well as theories of episodic memory. The distinction, suggested by Tulving (1972), is that traditional memory tasks, such as free recall, require the subject to remember the occurrence of temporally dated events or episodes, while semantic memory tasks call upon the subject to retrieve information from his base of knowledge of the world and of language.

Tulving (1972) argued that semantic and episodic memory should be regarded, at least conceptually, as two separate memory systems. Alternatively, it may prove more useful to conceive of episodic and semantic memory as reflecting the operations of different encoding, storage, and retrieval processes operating on a common structural data base. To the extent that the latter view is correct, the task of mapping the structure of memory organization assumes even greater importance.

Theories of memory organization differ in many respects, including their pretheoretical assumptions, the principal phenomena which they explain, and their detailed specifications of memory structures. However, they can be grouped into three major categories, dimensional models, tree-structure models, and network models, depending on the general format assumed for the structural component of memory.

Dimensional Models

One class of representations for the structure of organized memory is modeled on a multidimensional attribute space. The common feature of these approaches is the notion that verbal items are encoded or represented in memory in terms of their values on a set of features, attributes, dimensions, properties, or the like. These dimensions are sometimes conceived of as continuous, as in Euclidean space, and sometimes considered to take discrete values, e.g., semantic markers.

Bower (1967) presented a highly elaborated model for the encoding of item information in terms of a bundle or vector of attributes and provided evidence in support of this conception, mostly from recognition studies. The attributes of items encoded in this way were left unspecified, in general, but would presumably be determined by the cognitive context and demands of a particular task. Underwood (1969) also conceptualized the memory trace as a collection of attributes. He identified temporal, spatial, frequency, modality, orthographic, and associative attributes as playing operative roles in particular memory tasks.

In these papers Bower and Underwood are principally concerned with the manner in which individual items are represented in memory. Other

dimensional models of memory structure place more emphasis on the establishment and utilizations of interrelationships among items. Voss (1972) has suggested that organization occurs along four general structural dimensions, formal, associative, syntactic, and semantic, and he has formulated an analysis of associative learning in terms of the arousal or activation of the encodings of verbal items along these dimensions.

Smith, Shoben, and Rips (1974) proposed an attribute model of memory structure to account for the findings in tasks involving verification of semantic relations (such as *A robin is a bird*). In their model, lexical terms and concepts are represented by sets of features. The features for each item are assumed to be ordered in terms of the degree to which they serve to define the item. In addition, Smith et al. (1974) argued that *characteristic* features, attributes that are typical, but not necessary to define an item, must also be represented in memory. Presumably, items in a given conceptual category (e.g., *bird*) would share the same defining features, but would differ in their characteristic features. Rips, Shoben, and Smith (1973) used multidimensional scaling (MDS) to depict the structure of relationships within the conceptual categories *birds* and *mammals*. Some evidence for the validity of their dimensional model was provided by the fact that the distances between items in the MDS solutions proved to be a good predictor of reaction time in the sentence verification task.

Tree-Structure Models

A second general class of models is based on hierarchical structures. The principal assumption underlying this conception of the memory store is that the internal representation of items consists of sets of categories or groups of items arranged in a nested fashion. Within the context of free recall research, this idea was developed by Mandler (1967, 1968, 1970) as an extension of Miller's (1956) chunking hypothesis. Mandler (1967) described the operation of a hierarchically organized memory system as follows:

We assume first the basic limit of the organizing system at 5 ± 2 per set of items. . . . Given that limitation, categories will be formed and 3–7 items assigned to them. . . . Once these initial categories are filled up, new categories will be created to accommodate additional items. But in turn, there will be a limit of about 5 ± 2 categories at this first level of categorization. When all slots are taken up with first level categories, a second level of categories will be formed, each of which may contain up to about seven first-level categories, and so forth. In this manner, a hierarchical system of categories can be built up with an increasing level of complexity and an exponential growth in the size of the system (p. 366).

Wood (1972) and Bower, Clark, Lesgold, and Winzenz (1969) have discussed the nature of retrieval mechanisms which might operate on such a hierarchical structure. They suggest that hierarchical organization would

provide subjects with an efficient retrieval plan, so that access to a memory unit at any level of a hierarchy would provide a high probability of accessing the units at the next lower level.

The efficacy of hierarchical organization of recall has received considerable empirical support. Recall of a categorized list is better when the list categories form a hierarchical structure than when they are independent (Cohen & Bousfield, 1956). One also finds facilitation of a hierarchically categorized list when that list is presented in a way that makes its structure most apparent. For example, recall is improved by simultaneously presenting the items in a tree structure (Bower et al., 1969) or by presenting items sequentially, blocked according to the hierarchical structure (Friendly, Note 4). Similar results are found when subjects form their own subjective hierarchies by sorting list words into nested groupings (Friendly, Note 5): Recall and delayed recognition are better following hierarchical sorting than when items are sorted in a single level of categories.

Tree structures have also grown in semantic memory research. Miller (1967, 1969) used the method of sorting to study how semantic information is organized and stored in permanent memory. In this task the subject examines a series of words and places them into clusters on the basis of "similarity of meaning." Miller (1969) argued for a model in which words are encoded in terms of a list of features, but the features themselves are hierarchically ordered. Thus, *canary* might be represented by the features (*object, animal, vertebrate, bird, . . .*), where each feature presupposes those prior in the list. Miller suggested that subjects in the word sorting task must ignore certain distinguishing features, while paying attention to others. For example, to sort *canary* with *robin*, the subject must ignore color and domesticity features; to sort *canary* with *trout*, he must ignore class differences as well. The fewer the features on which two items differ, the greater the number of subjects who should sort the items together. Using the number of subjects sorting a pair of words together as an index of the proximity of those two words, Miller showed that a hierarchical organization of concepts produced definite constraints on the proximities. The hierarchy itself could be determined by applying hierarchical cluster analysis to the matrix of proximity values obtained from a group.

In tree-structure models a given node may only be connected to nodes one generation above or below. It is worthwhile to point out that the tree-structure models discussed above are *strictly hierarchical*, in that each node is connected to exactly one node at the next higher level. Lexical domains based on a class inclusion relationship (such as *Animal names*) do have a strictly hierarchical structure, as long as the classes do not overlap. However, as Kintsch (1970) points out, many interesting lexical domains (such as *Body parts*) have properties of a tree structure but are not strictly hierarchical. Methods for representing memory structure in terms of tree

structures with overlapping classes will therefore be considered in what follows.

Network Models

In a network the restrictions of tree structures are removed. Network models represent word concepts as nodes and relations among words as links connecting pairs of nodes. Any node may be linked to any other, and the links may be binary (representing a relation present or absent), real-valued (strength of relation), and/or labeled (type of relation).

The earliest network models arose as attempts to apply structural analysis to word association behavior (Guiliano, 1963; Kiss, 1967, 1969; Pollio, 1966). In the model proposed by Kiss (1967, 1969), items in the network are characterized at any instant by their level of activity or excitation. The links allow the flow of excitation throughout the network according to their transmittance values (strength of association between nodes).

Network models have been widely adopted in computer simulations of human memory. A computer model of free recall based on a simple associative network was developed by Anderson (1972). The model, called FRAN, consists of a semantic memory network of 262 nouns with links defined from dictionary definitions, together with routines to simulate study and retrieval processes. During a study trial, FRAN marks the node corresponding to the current word on the list with a tag, and searches its associative network for pathways leading from that node to previously studied (and hence marked) nodes. Successful pathways so found are themselves marked with a tag, which serves to provide efficient retrieval routes for searching the memory network during a recall trial. In recall, FRAN uses some starter items as entry points to its network, following marked pathways and retrieving nodes marked with a tag.

This is a relatively simple type of network structure, and the match between simulation results and comparable human data is therefore impressive: Averaged data from FRAN closely approximate those of real subjects in their learning curve, subjective organization, serial position curve, and study-time vs recall relationships.

FRAN fails, however, to duplicate some key MFR results. These misfits are instructive as to the boundaries of simple associative networks as models for memory structure. Human subjects recall quite well when semantic categories are present in a list, and do even better when the items are presented in a blocked fashion (Cofer, 1965; Dallett, 1964; Weist & Powell, 1972). On the whole, FRAN recalls less and shows less category clustering than humans when learning structured lists, and these discrepancies are even greater under blocked presentation. The reason for this, of course, is that FRAN's network does not distinguish among items

linked by different types of relationships and cannot recognize the commonalities among word sets such as *table*, *chair*, *sofa* or *fat*, *cat*, *hat*.

Anderson (1972) in fact, argued for the necessity of relation-labeled links in a network representation of memory structure, although the simulation program did not incorporate this feature. A number of memory models employing such relational networks have recently been proposed. These models are generally rather more ambitious in scope and therefore utilize more elaborate structures.

Quillian (1968) proposed a general model of semantic memory which is directly configured in terms of a labeled associative network, but has a number of the features of a hierarchical structure. The model was later extended (Quillian, 1969) into a theory and computer realization designed to deal with the understanding of continuous text.

Quillian's model consists of a network of nodes connected by different types of links.¹ Two types of nodes appear in the memory structure: *type* nodes which roughly correspond to dictionary word entries, and *token* nodes which represent the occurrence of word concepts in a particular context. The meaning of a type node resides in the configuration and types of links to its defining tokens. This configuration is called a *plane*. For example, the noun *retort*, meaning "a vessel in which substances are distilled or decomposed by heat," would be represented by a plane consisting of the type node RETORT with labeled links among token nodes for *vessel*, *substance*, *decompose*, etc. The token nodes have links which point out of this plane to their own respective planes. This arrangement makes it possible to build up progressively more elaborate concepts using other word meanings.

Class inclusion relations play a particularly dominant role in Quillian's (1969) model, in that every node contains a mandatory pointer to its immediate superset node. Further, property values related to a concept are assumed to be stored at the highest node for which the property value applies to all descendent nodes. Thus, it is assumed that people store only the general fact; e.g., they store the information that birds can fly and infer that a canary can fly by tracing their superset links from *canary* to *bird*. Reaction times in a sentence verification task have been shown to be consistent with this assumption (Collins & Quillian, 1969, 1970, 1972).

¹Similar models of memory structure have been devised by a number of psychologists and computer scientists concerned with language understanding (for example, Anderson & Bower, 1973; Kintsch, 1972, 1974; Norman & Rumelhart, 1975; Rumelhart, Lindsay & Norman, 1972; Schank, 1973; Simmons, 1973; Fridja, 1972 provides a thorough review of earlier work). Formally, these workers represent memory networks in various ways (for example, using list structures) and give greater or lesser prominence to various psychological and linguistic principles and relations according to their aims. For the present purposes, however, they do not introduce additional structural components into the memory representation and will not be considered further here.

Relations among the Models

It is not the purpose of this paper to evaluate these alternative models or contrast them in detail. However, there are several relations among these models which are worth noting here.

First, it should be clear that a hierarchical structure is a special form of network, which has restrictions on allowable links. In this light Quillian's network model is particularly interesting since the memory structure *locally* displays many features of a hierarchical system: Tracing all paths reachable from a given node generates a structure of progressively finer detail in a top-down manner; new concepts are linked to their superset nodes; and properties are stored hierarchically. The papers of Collins and Quillian in fact explicitly portray these local hierarchies, which are embedded in the overall memory network. Thus the memory system may be structured *both* hierarchically and as a network depending on the scope of view. Restricted lexical neighborhoods in permanent memory (a single-word concept, a taxonomic category) display strongly hierarchical features while larger, more general domains may appear to be networks.

In addition, there is a close connection between dimensional models which represent word concepts using lists of features and network models with relational links. For, given a feature-list representation for a set of word concepts, it is possible to construct a network representation in which the words appear as nodes and the features are used to label the links between nodes (Fridja, 1972).

The foregoing observations suggest that these models are actually more similar than has previously been realized and that the network model may be the most general. However, the conceptual differences among these three models are important since they affect how memory theorists think about the organization of memory and the experiments which they perform to unravel its mysteries.

Empirical Approaches to Organization in Free Recall

The free recall task has been the primary experimental vehicle used to study organization in episodic memory. However, the methods available for this purpose have been less strongly tied to the *structural* features of mnemonic organization than have tasks and methods developed by semantic memory researchers. There have, of course, been numerous proposals for measuring organization in free recall (Bousfield & Bousfield, 1966; Dalrymple-Alford, 1970; Ehrlich, 1965; Frankel & Cole, 1971; Hudson & Dunn, 1969; Pellegrino, 1971; Roenker, Thompson, & Brown, 1971; Tulving, 1962a). However, these procedures are all concerned with measuring the *amount* of organization in free recall output, rather than with the explicit *structure* of this organization.

It is only with the use of categorized lists that measures of the amount of

organization (category clustering) can be related to organizational structure. It has been argued, however, that category clustering provides an indication of organizational format only to the extent that the subject actually *discovers* and *makes use* of the list structure defined by the experimenter (Mandler, 1967). A pervasive interpretive problem in these studies, therefore, is that the appearance of weak or moderate amounts of clustering is theoretically ambiguous. One often cannot tell whether the conditions of learning made it difficult for the subjects to organize, or whether they were merely organizing their recall in some unanticipated fashion.

The use of experimenter-defined organizational schemes relies upon the assumption that language users share a sizable core of semantic information which can provide a dominant basis for organized recall. In contrast, Mandler (1967), Postman (1972), and Tulving (1968) have stressed the importance of focusing attention on the structure of subjective organization of unrelated lists: "In order to fully evaluate the relation between the type of subjective organization and recall, it is desirable to make the entire structure generated by the learner accessible to inspection" (Postman, 1972, pp. 24-25). Several investigators have attempted to provide this type of assessment.

Seibel (Notes 11, 12) introduced a type of word-sorting task involving a modification of the typical presentation phase. With this procedure, the subject was given a study sheet with a large grid at the beginning of each trial. The subject was instructed to write each word as it was presented in any cell of the grid. This procedure allowed the subject to establish a subjective categorization during input and to rehearse these categories as presentation proceeded. At the end of each presentation, the subject wrote the words he could remember on a new blank sheet of paper. Seibel found that items written together on the study sheet also appeared as output sequences during the subjects' recall. A control group, instructed to write the items on the study sheet in the order of presentation, recalled less well than the group allowed to form subjective categories.

In a comprehensive series of experiments, Mandler (1967, 1970); Mandler & Pearlstone, 1966) used a word-sorting task both to induce a stable, subject-determined organization and to make this organization directly observable. In these studies, the subject was typically required to sort 50 to 100 words into anywhere from two to seven subjective groups. After several sorting trials, memory for the items was tested, usually in a single free recall trial. In these studies, Mandler was primarily concerned with the number of categories used in sorting as a predictor of subsequent recall performance and found a linear increase in recall as a function of this variable (up to approximately seven categories). He did not, however, focus on other structural features of the subjective categorizations.

Allen (Notes 1, 2) developed similar memory unit identification tasks.

These tasks were employed after a series of FR trials had been completed and required the subject to indicate which groups of words went together in his memory for the list. Allen showed how the concepts and techniques of graph theory could be applied to the data from these tasks to generate an empirical directed graph representing the subjective structure of memory items. He also demonstrated that various aspects of recall, such as amount of subjective organization and number correct, could be postdicted from the graph representation of the individual subject's organization. Since a directed graph is just a special case of a network (one having directed links either present or absent between nodes), Allen's work provides a basis for developing an M-gram for free recall which can be coordinated with the network models of memory described earlier. This point will be elaborated later in this paper.

PROXIMITY ANALYSIS

The procedures used by Seibel, Mandler, and Allen for identifying memory units are based on the use of supplementary information or modifications of the FR task. In contrast, the technique discussed below attempts to identify subjective groupings from order-of-recall information directly. The method subsumes the measurement of categorical clustering and subjective organization within a single unified framework in that it assumes no prior knowledge by the experimenter of the bases of organization. To the contrary, it offers an objective way to determine these bases. It should be emphasized that the concern here is with quantifying the *structure* of organization, rather than the *amount* of organization, for which there are already many measures available (Shuell, 1969). As noted above, organizational phenomena can be discussed in various languages, based on different structural analogs. In the discussion that follows, I will use the terminology of a hierarchical memory model, subjective memory units, higher-order units, etc., primarily for the sake of convenience, but also because organization in free recall has largely been discussed in these terms.

Briefly, this procedure locates subjective memory units on the basis of ordinal separation or *proximity* between pairs of words in recall protocols. A proximity measure is developed from the assumption that items which are coded together in subjective units will tend to be recalled contiguously at the time of test. This assumption is entailed by all measures of clustering and subjective organization. An item-by-item proximity matrix can then be constructed with numerical entries representing the degree to which each pair tends to occur in contiguous (though not necessarily adjacent) output positions over some set of trials. It is shown that for each class of models of organizational structure there is a corresponding natural method for

analyzing the proximities obtained from individual or group data to provide an M-gram representation.

It is useful to proceed heuristically at first to develop the logic and assumptions of the technique. Following that, the proximity measure is presented formally and then illustrated with sample data.

Consider a hypothetical subject presented with a categorized word list who recalls the following items on a given trial: pants, *shirt*, *shoe*, doctor, shrub, *bush*, *tree*, lawyer, *dentist* in that order. In the category clustering approach, organization would be assessed by counting the number of sequential repetitions of items from the same category (the five italicized words above).

Note that this approach only considers adjacent output pairs as evidence for grouping in memory. But all pairs of items in a memory unit cannot be immediately adjacent in output which is structured sequentially. So, the degree of organization is probably underestimated (cf. Postman, 1971). "As a network of multiplex interitem relations develops, there is a corresponding increase in the number of equivalent pathways leading from one response to another" (Postman, 1972, p. 23). However, what these alternative pathways have in common is that members of a memory unit occur in close proximity. Therefore, the underestimation of organization can be partly avoided by allowing for varying degrees of proximity between items. Thus *pants* and *shirt*, for example, are maximally close while *pants* and *shoe* are less proximal, and so on for the other categories. The assumption made here is that the proximity in recall between items which are not adjacent also carries information about memory structure.

Going a step further, it is possible to look at the proximities between *all* pairs of words in the protocols, not just those within the given categories. For example, *bush* and *tree* are more proximal than are *shoe* and *tree*, though the reverse could have occurred if the subject has thought of the compound noun, *shoetree*, and clustered on that basis. The actual outcome can be expressed quantitatively by giving the pair *bush* and *tree* a higher proximity score for that trial than the pair *shoe* and *tree*, and so on for all pairs of items, basing the proximity score on their ordinal separation in the protocol.

Considering the proximities between all pairs of items makes this way of looking at the subject's organization of a list independent of any knowledge of a priori categories. The use of the number of sequential repetitions as an index of organization requires, by definition, a knowledge of which groups of items belong together. Through the use of proximities, however, it is possible to "discover" the grouping that the subject is using by defining the subjective units to be those groups of items that have mutually high interitem proximities.

Stated alternatively, one can ask what grouping of the items into memory

units would be most likely to result in an individual's observed response protocols. In the analysis suggested here, the aspects of order of recall information most relevant to the study of memory structure in MFR may be represented by the proximities between all pairs of items. Questions concerning the organization of list items in memory can therefore be translated to corresponding questions concerning the proximities among them in recall (cf. Martin, 1970).

Measure of Interitem Proximity

It remains to specify a way to quantify proximity, or its inverse, distance. Several workers have used measures based on the frequency of immediately adjacent recalls (Kintsch, 1970; Koh, Vernon, & Bailey, 1971) to investigate organizational structure. However, in order to get stable estimates of interitem proximity it is important to utilize more of the information contained in recall order. One simple way to do this is to measure the distance between two items in terms of the number of other items which separate them in recall. Consider a list of L items presented to a group of S subjects on each of T trials under typical conditions. For a given subject the data consist of T sequences of items, each of length r_t , where r_t is the number of words recalled on trial t .

Denote by l_{it} the position of item i in the subject's output on trial t . Then the intraserial distance between two items, i and j , both recalled on a given trial will be $|l_{it} - l_{jt}|$. The proximity measure which has been found most useful is based on the average intraserial distance, D_{ij} for a pair, over those trials on which *both members of the pair are recalled*. D_{ij} is defined as

$$D_{ij} = \frac{\sum_{t=1}^T \phi_{ijt} |l_{it} - l_{jt}|}{\sum_{t=1}^T \phi_{ijt}}, \quad (1)$$

where ϕ_{ijt} is an indicator variable which equals 1 if words i and j are both recalled on trial t and equals 0 otherwise. The denominator of Eq. (1) is just the number of trials on which words i and j were recalled. The distance values defined in Eq. (1) have a minimum value of 1, obtained when a pair of words is recalled adjacently on all trials. The maximum value is \bar{r} , the average number of words recalled; it is obtained when a pair is consistently recalled at opposite ends of the recall output.

It is often convenient to think in terms of proximity between items, rather than distance, since the former is directly related to tightness of organization. For the present purposes, the distance measure can be transformed to an index of proximity by subtracting it from a positive

constant, so that large numbers represent more proximal, interconnected items. Average interitem proximity is thus defined as list-length minus average distance,

$$P_{ij} = L - D_{ij}. \quad (2)$$

In comparing proximities among items across lists which vary in length, it is often convenient to normalize the proximities to proportions of their maximum value, i.e., to use $P'_{ij} = P_{ij}/(L - 1)$. Several aspects of these definitions are noted below; further elaborations and some technical details are discussed in Friendly (Note 4).

(a1) When either or both members of a pair are not recalled on a given trial, it is difficult to decide how a distance may be rationally assigned. A value could be assigned ad hoc, but it is safer to assume that this event gives no information regarding the organizational relatedness of the pair. Equation (1) essentially averages over only those trials on which both members of a pair are recalled, ignoring trials on which one or both items are omitted.

(a2) Although Eq. (1) indicates averaging over all free recall trials, it is certainly possible, and in some cases desirable, to restrict attention to some subset of all trials. With unrelated lists, for example, it is sometimes useful to drop the first trial or two to remove initial serial position effects.

(a3) The definitions above were stated for a single subject. More generally, the proximities obtained from a set of $S \cdot T$ subject-trials may be averaged over both subjects and trials to give a single group matrix, or over subjects for each trial separately. The former procedure is convenient for comparing entire groups, while the latter approach, adopted by Caramazza, Hersh, and Torgerson (1976, Expt IV) in their application of proximity analysis, allows one to study trial-to-trial changes in organizational structure. Averaging helps ensure that the estimated proximities are sufficiently stable. Pooling over subjects, however, entails the assumption that the subjects so pooled share a common organization of the list. This assumption is most likely to be satisfied when the list consists of normatively related words, rather than unrelated words. However, it is also possible to assess the degree to which individual differences in organizational structure exist and to isolate subgroups of subjects who share a common organizational scheme. This will be discussed and illustrated later.

The proximity measure given above appears to be a reasonable index of interitem relatedness in recall output. It is not claimed, however, that it is the best such measure, since what is "best" can only be defined in the context of a detailed theory of organization (Colle, 1972). A number of possible refinements can be suggested:

(b1) Examination of Eq. (1) indicates that D_{ij} varies linearly with the

separation between items i and j in recall. To give greater weight to pairs recalled in close contiguity, D_{ij} could be defined as a function of the squared separation between items, e.g.,

$$D_{ij}^* = \left[\frac{\sum \phi_{ijt} |l_{it} - l_{jt}|^2}{\sum \phi_{ijt}} \right]^{1/2} \quad (3)$$

(b2) Proximities could be weighted differentially across trials, e.g., by giving increasing weight to later trials, where presumably, the subject's organization would be more stable.

(b3) When subjects recall orally, it is often noticed that subjects recall items in bursts, with short interresponse times (IRTs). The bursts typically correspond to strongly interconnected words in terms of word associations (Pollio, 1966; Pollio, Kasschau, & DeNise, (1968) or members of taxonomic categories (Patterson, Meltzer, & Mandler, 1971). Also, the IRTs increase as less strongly connected items are retrieved (Murdock & Okada, 1971). It would be reasonable, therefore, to use the total interresponse time between a pair of words in recall as an inverse measure of interitem proximity.

Although these or other modifications may improve the proximity measure to some degree, it should be noted that the proximity values will be utilized in a manner which depends only on their ordinal properties. Consequently, different methods for calculating proximities will lead to the same results as long as the rank order of the item pairs is maintained.

Illustrative Data

To make things more concrete, consider the data in Fig. 1. This shows the actual recall protocols from one subject on the last six trials of an eight-trial free recall session.² On each trial $L = 12$ unrelated words were presented visually in a different random order, and the subject's task was to recall as many words as possible.

Consider Trial 5. Item pairs which are immediately adjacent, such as (*highway, structure*) and (*inventor, professor*), differ in ordinal position by one, so their proximity on that trial is $L - 1$ or 11. On the other hand, words widely separated in the protocol have a lower proximity on that trial; for example, *mast* and *assault* which are 5 positions apart, have a proximity of 7.

Table 1 shows the proximities of selected pairs of items for the six trials in Fig. 1. For each selected pair, the average proximity over all trials on which both items were recalled and the normalized value, P' , are also shown. Thus, *quarrel* and *assault* were immediately adjacent on all six

²These data come from a study by Ornstein (1970, Expt. 1). They have been used here by permission of the author.

Trial 3:	urge	quarrel	assault	captive	decree	execution	mast	professor	inventor	north			
Trial 4:	urge	north	quarrel	assault	captive	execution	highway	professor	inventor				
Trial 5:	highway	structure	decree	mast	north	urge	execution	captive	assault	quarrel	inventor	professor	
	$p \approx 11$		$p \approx 9$				$p \approx 10$			$p \approx 7$		$p \approx 11$	
Trial 6:	quarrel	assault	decree	execution	captive	urge	north	structure	highway	inventor	professor		
Trial 7:	quarrel	assault	captive	execution	decree	north	structure	mast	highway	inventor	professor	urge	
Trial 8:	quarrel	assault	captive	execution	decree	urge	structure	north	inventor	professor	highway	mast	

FIG. 1. Recall protocols for one subject in a free recall task.

trials and have an average proximity of 11, the maximum possible for a list of 12 words. *Captive* and *highway*, on the other hand, were consistently quite far apart with an average proximity of 6.6. Hence there would be little reason to believe that these two items belonged to the same functional memory unit for this subject.

Pairs of items also differ in the frequency with which both members of the pair are recalled. Thus *captive* and *highway* were both recalled on all six trials. *Mast* and *highway*, on the other hand, were both present in output on only three of the trials shown. When they were both recalled, however, they were quite proximal.³

The proximities are calculated for all pairs of words and arranged in a square matrix as shown in Table 2. The matrix is necessarily symmetric, so only the lower half is shown. The principal diagonal has also been omitted, since it conveys no information: $D_{ii} = 0$ for all items.

This matrix shows that there are several groups of words which have mutually high proximities within each group and relatively low proximities with items outside the group. *Inventor* and *professor*, for example, seem to constitute a fairly distinct grouping for this subject since their normalized proximity is 1.0 and each of these words has relatively low proximities with all the other items (columns 1 and 2 of the matrix). Similarly, the items *assault*, *quarrel*, *captive*, *execution*, and *decree* are all highly proximal to one another in this subject's recall. A third highly organized group consists of *highway*, *mast*, *north*, and *structure*. The word *urge* appears to be a singleton; it is recalled on all trials by this subject, but it does not appear consistently near any other items. These four sets of words constitute a reasonable approximation to the subjective groups displayed in this subject's recall.

Usually, however, the items will not be arranged in the proximity matrix so that their structure is so apparent. Indeed, in preparing the table the rows and columns were reordered so that the groups of co-organized items would be together, giving rise to the triangular blocks of high proximities.

³If the frequency of recall for a pair of items, $\Sigma \phi_{ijk}$, is small, the estimate of proximity for that pair will not be very stable. In an extreme case, two items which are recalled concurrently only once, but in adjacent positions, would receive the same proximity value as a pair recalled adjacently on all trials. One way to avoid this possibility is to set a threshold value, so that pairs recalled less often than this value have their proximity set equal to zero or reduced by some constant fraction.

TABLE 1
PROXIMITIES FOR SELECTED ITEM PAIRS FROM RECALL PROTOCOLS OF FIG. 1

Stimulus pair	Trial						Average ^a proximity	Normalized ^a proximity
	3	4	5	6	7	8		
quarrel-assault	11	11	11	11	11	11	11.0	1.000
captive-assault	11	11	11	9	11	11	10.7	.970
captive-highway	0	10	5	8	6	4	6.6	.600
mast-highway	0	0	9	0	11	11	10.3	.939
mast-quarrel	7	0	6	0	5	1	4.7	.432
north-structure	0	0	9	11	11	11	10.5	.955
north-urge	3	11	11	11	6	10	8.7	.788
professor-structure	0	0	2	9	8	9	7.0	.636
professor-inventor	11	11	11	11	11	11	11.0	1.000
decree-inventor	8	0	4	5	7	8	6.4	.582

^a Entries rounded off to three significant digits.

In general, the proximities will need to be subjected to further analytical scrutiny to reveal the underlying organization reflected in the order of recall. How this is done depends on the underlying model assumed for the structure of organization.

SPATIAL REPRESENTATIONS OF ORGANIZATIONAL STRUCTURE

To each of the three general models of memory structure, there corresponds a natural way of representing the relations among items in recall on the basis of output order proximities. The term "natural" in this context means that the representation is a useful graphical realization of a plausible memory structure and that the characteristics and parameters of these representations bear reasonable interpretations in the context of the corresponding structural models. In each case the representation involves a mapping of the proximities into distance-like values, say, d_{ij} , with the property that

$$\mathbf{P} \stackrel{m}{\approx} \mathbf{d}, \quad (4)$$

where \mathbf{P} is the symmetric $L \times L$ matrix of proximities, and \mathbf{d} is the corresponding matrix of distances under the model. The notation $\stackrel{m}{\approx}$ indicates that the proximities are approximately monotonic with the d_{ij} , i.e., if $P_{ij} > P_{kl}$ then $d_{ij} \leq d_{kl}$. In essence, the representations are obtained by applying relatively well-known scaling methods to the proximities. The cases are distinguished by the additional restrictions on the d_{ij} which are imposed by the scaling analog of the particular memory model. The correspondences are shown in Table 3 and explained in the remainder of this section.

Since the mathematical bases of these techniques have been treated extensively elsewhere, the descriptions which follow will be incomplete in some respects. The comments below focus on issues particular to their use for representing memory structures. Several of the methods are illustrated using sample data. However, no claim is made that the particular representations are perfectly suitable for all types of semantic domains.

Dimensional Representations

Some dimensional models assume that each item has a unique location in a continuous multidimensional Euclidean space; others are concerned with discrete or binary features. In the former case the proximities may be regarded as monotonically related to distances among points in such a Euclidean space. That is, we seek a representation of the list items as points in a Euclidean space such that the spatial distance between two points represents the average recall proximity between the corresponding items. The closer together the items are recalled, the closer should be their points in the spatial representation.

Using a nonmetric multidimensional scaling (MDS) procedure (e.g., Kruskal, 1964a,b; Young & Torgerson, 1967), it is possible to determine

TABLE 2
MATRIX OF AVERAGE PROXIMITIES

	1	2	3	4	5	6	7	8	9	10	11	12
1. Inventor												
2. Professor	100*											
3. Highway	79*	78										
4. Mast	77	77	94*									
5. North	76	70	79	86								
6. Structure	73	64	91*	85	95*							
7. Assault	55	52	47	52	67	57						
8. Quarrel	48	45	38	43	61	48	100*					
9. Captive	64	61	60	61	76	70	97*	88				
10. Execution	68	65	65	80	80	75	89	80	98*			
11. Decree	58	53	68	80	82	86*	82	73	87	95*		
12. Urge	64	64	69	68	79	82*	70	67	76	76	76	

Note. Entries are expressed as a percentage of the maximum possible value, and have been rounded (decimal points omitted). Entries marked with an asterisk (*) are explained later in the text.

TABLE 3

MEMORY MODELS AND THE CORRESPONDING SCALING METHODS PRODUCING AN M-GRAM REPRESENTATION FROM RECALL PROXIMITIES

Memory model	Representation	Scaling methods
Dimensional	Euclidean space	Multidimensional scaling; INDSCAL
	Overlapping attribute sets	Arabie & Shepard (Note 3)
Tree structure	Strict hierarchy	Complete-link HCS
	Hierarchy with overlapping clusters	B_k clusters
Network	Undirected graph	Single-link HCS Minimum spanning tree Maximal connected subgraphs B_k graph

the location of items in a Euclidean space as well as the number of dimensions necessary to fit the proximities adequately. The item locations are specified by a matrix of coordinates, $\mathbf{X} = \{x_{ik}\}$ $i = 1, \dots, L$; $k = 1, \dots, m$, where m is the number of dimensions. In the MDS solution the distances, \mathbf{d} , referred to in Eq. (4), are related to the coordinates by

$$d_{ij} = \left[\sum_{k=1}^m (x_{ik} - x_{jk})^2 \right]^{1/2}. \quad (5)$$

Although L items can always be represented in a space of $L - 1$ dimensions (Kruskal, 1964a), a Euclidean representation would not be very useful unless the proximities could be fit into a space of considerably smaller dimension. In the MDS procedure due to Kruskal, the adequacy of a dimensional representation is indexed by a normalized residual sum of squares coefficient, called stress,⁴ (S_1)

$$S_1 = \left[\frac{\sum (d_{ij} - \hat{d}_{ij})^2}{\sum d_{ij}^2} \right]^{1/2}, \quad (6)$$

⁴ An alternative formula for stress, S_2 , uses $\sum (d_{ij} - \bar{d})^2$ as the normalizing factor [the denominator in Eq. (6)], where \bar{d} is the mean of the distances. It has been found that S_1 sometimes produces degenerate solutions, especially when the items cluster strongly into groups such that all proximities within a group are greater than any of the proximities between groups (see Shepard, 1974). In analyzing recall proximities, such degeneracies are particularly likely when the items belong to salient taxonomic categories. In such cases S_2 will often help avoid degenerate solutions. However, with words of this sort a representation in terms of continuous underlying dimensions will not typically be appropriate to display a categorical structure. Instead, one of the hierarchical or nonhierarchical clustering methods discussed in the following subsection would be more appropriate.

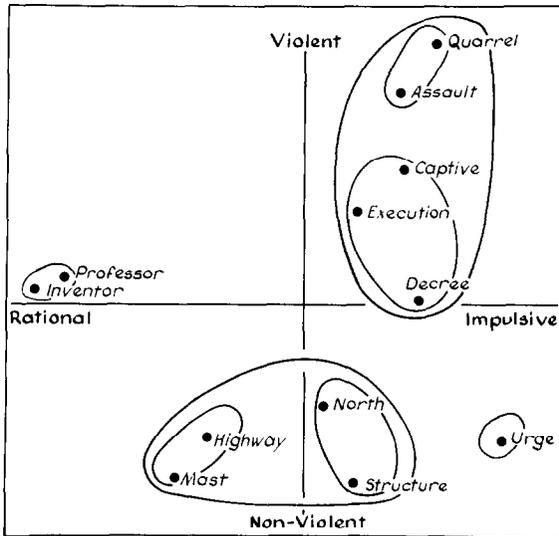


FIG. 2. Multidimensional scaling solution for the proximities in Table 2.

where \hat{d}_{ij} are the best-fitting distances which strictly preserve the order of the interitem proximities. Approximate significance tests of fit are available from the work of Wagenaar and Padmos (1971), Stenson and Knoll (1969), and Spence and Ogilvie (1973).

The interpretation of the M-gram which results from applying MDS to the recall-order proximities is simplified to the extent that an adequate fit in two or three dimensions can be obtained, so that the representation is readily visualized. The interpretation of the spatial configuration is then largely a matter of searching for meaningful clusters or orderings of the item points in the space. In addition, several objective methods for locating reference axes and rotating the space (Carroll & Chang, 1964; Miller, Shepard, & Chang, 1964; Shepard, 1972) can be used as aids in interpreting the space. For categorized lists, Degerman's (1970) rotation method is particularly useful for separating qualitative variation due to class membership from quantitative variation due to semantic dimensions.

For illustration, the MDS procedure was applied to the proximity matrix in Table 2. The computer program KYST (Kruskal, Young, & Seery, Note 7) was used to obtain the solution.⁵ A two-dimensional solution was obtained having a stress value $S_1 = .064$, which may be regarded as a reasonably good fit. The resulting configuration is shown in Fig. 2. The

⁵Since KYST allows the proximities to be weighted differentially, the pair recall frequencies, $\Sigma\phi_{ijk}$, were entered as weights. For this set of data, the frequencies range from three to six and have little effect on the solution; when the proximities are obtained by averaging over more subject-trials, the use of weights gives emphasis to item-pairs in direct proportion to the frequency with which they are both recalled on the same trials.

dimensions in this M-gram representation might be interpreted as *Violent-Nonviolent* and *Impulsive-Rational*. However these dimensional labels do not appear to be strongly compelling, since the words were selected as "unrelated" words. (Since *highway*, *mast*, *north*, and *structure* are inanimate, the label *Nonviolent* is somewhat strained.) The structural relations among these items are captured better, in this instance, by the clusters circled in Fig. 2. These correspond to the subjective clusters noted by inspection in Table 2; but in addition, several subgroups within the two larger clusters may be noted. It would not be unreasonable to define subjective memory units as such relatively compact isolated clusters which appear in the MDS solution. More objective criteria for identifying memory groupings are considered later.

Discrete features. When the memory structure is assumed to be represented in terms of discrete features, rather than continuous dimensions, the model of Euclidean space can no longer be used. One possibility for representing the structure of recall proximities within such a conception is embodied in a procedure developed by Arabie and Shepard (Note 3). In their method the set of items are assumed to be characterized by a total of m distinct properties which vary in psychological salience. Any pair of items will share some features and the proximity between two items is assumed to arise as a sum of the salience values of all shared features, i.e.,

$$P_{ij} = \sum_{k=1}^m w_k v_{ik} v_{jk}. \quad (7)$$

where $v_{ik} = 1$ if item i has property k and equals zero otherwise; w_k is a nonnegative weight representing the psychological salience of property k . The method determines m overlapping groups of items which possess each feature and each feature may be interpreted by identifying the common property which all items in the corresponding group share. Although the representation of memory structure embodied in this method is not inherently a graphical one, the results may be displayed graphically by circling the overlapping groups in the manner of Fig. 5 below. Initial applications reported by Arabie and Shepard (Note 3) suggest that their method may provide a useful representation of memory structures for models such as that proposed by Smith et al. (1974).

Tree-Structure Representations

On the view that memory is structured in terms of a hierarchical classification of items into memory units, it is natural to seek a representation of the structure underlying proximities in terms of compact, nested clusters. Such a representation may be obtained

using methods of hierarchical cluster analysis. In this representation the items appear as terminal branches of a tree structure rather than as points in space. The branches representing items which are recalled contiguously merge sooner in the hierarchy than the branches of items recalled less proximally. The relations of items within memory units and higher-order units are represented by the sequence of mergings of the branches.

Several types of hierarchical clustering procedures have been developed (Hartigan, 1967; Lance & Williams, 1967; McQuitty, 1960, 1964; Ward, 1963). Among these, two closely related techniques have received wide attention in psychological applications concerned with memory and verbal learning (e.g., Anglin, 1970; Fillenbaum & Rapoport, 1971; Kintsch, 1970; Martin, 1970; Miller, 1967, 1969). These hierarchical clustering schemes are the single-link (SL) method (also called connectedness method and nearest neighbor clustering) and the complete link (CL) method (also called diameter method and furthest neighbor). The discussion below follows Johnson (1967) and Miller (1969), whose work appears to have had the greatest impact on psychologists.

A hierarchical clustering scheme (HCS) consists of a tree structure with numerical values at the branches. The tree structure is composed of a sequence of partitions C_0, C_1, \dots, C_{L-1} of the items such that each successive partition in the series is formed by merging clusters from the immediately preceding partition. The numerical levels, $\{\alpha_0, \alpha_1, \dots, \alpha_{L-1}\}$, represent the compactness of the clusters at each stage.

The method begins with the finest partition, C_0 (the disjoint or "weak" clustering), in which all clusters consist of single items. The first nontrivial clustering is found by placing together the most proximal items in recall. The merged items are subsequently treated as a single element, and proximities between this new cluster and all other items are determined. Suppose that the two most proximal items are w_i and w_j with a distance of $D_{ij} = L - P_{ij}$. These items are therefore merged to form the cluster (w_i, w_j) and distance $D_{(ij)k}$ is assigned between the cluster (w_i, w_j) and any other item, w_k . For example, in Fig. 1, *inventor* and *professor* were recalled adjacently on all trials and have the highest possible proximity. When these are joined to form a cluster, the procedure assigns a proximity between this cluster and any other item, e.g., *urge*. Again, the most similar items/clusters are joined, and so forth until all items have been merged into a single cluster (the conjoint or "strong" clustering).

In a sense the complete-link and single-link methods provide opposite extremes of hierarchical clusterings which could be derived from the data. In the CL method when items w_i and w_j are joined to form a cluster, the distance between this cluster and any other item w_k is assigned $\max(D_{ik}, D_{jk})$, i.e., the distance between w_k and the *farthest* element in the cluster. Since it is the most proximal elements which are merged at any step, this algorithm produces clusters with minimum *diameters*, where the diameter

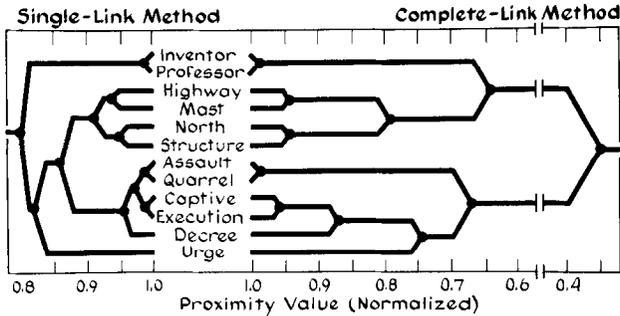


FIG. 3. Single-link and complete-link clustering solutions for the data of Table 2.

of any cluster is the largest distance within the cluster. Equivalently, for a given criterial distance, the CL method does not admit an item to a cluster unless it is at least that close to *all* items in the cluster. The method therefore produces clusters which are *completely* connected.

In contrast, in the SL method the distance between an item and a cluster is set equal to the distance between w_k and its *nearest* neighbor in the cluster, i.e., the minimum of D_{ik} and D_{jk} . This scheme will add an item to a cluster as soon as it is at a given distance from *any* item in the cluster, and the method produces clusters with minimum *chain-distances* (the smallest path of links connecting two items) which are only locally connected.

To illustrate how these methods work, they have been applied to the matrix for the 12 words in Table 2. The results are shown in Fig. 3. (The cluster analyses were performed using Gruvaeus and Wainer's (1972) algorithm, which orders the clusters so that items on the edge of each cluster are adjacent to that item *outside* the cluster which is most proximal). The first clusters formed contain those items which were recalled by this subject in immediately adjacent output positions on all trials (*inventor*, *professor*) and (*assault*, *quarrel*). The next highest proximity is between *captive* and *execution*, so these items are merged next, and so on, until all items have been merged into one cluster.

In this case, there is reasonably good agreement between the two methods. Both solutions indicate *assault*, *quarrel*, *captive*, *execution*, and *decree* as a higher-order unit, although they disagree on the order with which the smaller units (*assault*, *quarrel*), (*captive*, *execution*), and (*decree*) merged together. *Highway*, *mast*, *north*, and *structure* are clustered by both methods. The methods disagree most on the order in which these clusters and *urge* (seemingly a loner) merge subsequently.

This result is fairly typical of data from experiments using unrelated lists. A moderate degree of subjective clustering is observed, but sometimes these clusters do not appear to be tightly organized; occasionally no apparent structure above the level of two- or three-item clusters can be

discerned. When subjects learn lists of related sets of items, on the other hand, subjective groupings of the items are more obvious, more consensual, and output orders reflect more strongly connected *S*-units (cf. Cofer, 1965).

A small demonstration experiment was conducted to provide an illustration of the organization of categorized lists. Subjects in this experiment learned a list consisting of 18 items grouped into three categories, *Animals*, *Vegetables*, and *Body parts*. Six paid subjects received oral, randomized presentation for eight study-test trials and wrote their recalls. The hierarchical M-gram obtained by the CL method for a typical subject, with proximities pooled over all trials, is shown in Fig. 4.

The tree structure displays the strong grouping of the items into tightly knit, isolated clusters corresponding to the list categories. The items within the list categories merge together at high levels of interitem proximity, while separate categories do not merge until relatively low levels of proximity are reached. The merging of these categories with each other, ((*Animals*, *Vegetables*), *Body parts*), appears sensible, but is not statistically reliable.

Interpretable subgroupings can also be identified within the major categories. The items in the *Vegetable* and *Animal* categories were selected to have subclass relations with, respectively, two and three subcategories; the *Body parts* items have whole-part relations (e.g., *face-mouth*) with three subcategories. The M-gram in Fig. 4 displays this detailed structure clearly for the *Vegetables* and *Animals*, while the *Body parts* items only partially preserve the whole-part relation. This pattern of subgrouping, which generally appeared for all subjects tested, is interesting in itself and provides graphic support for Wood's (1972) suggestion that

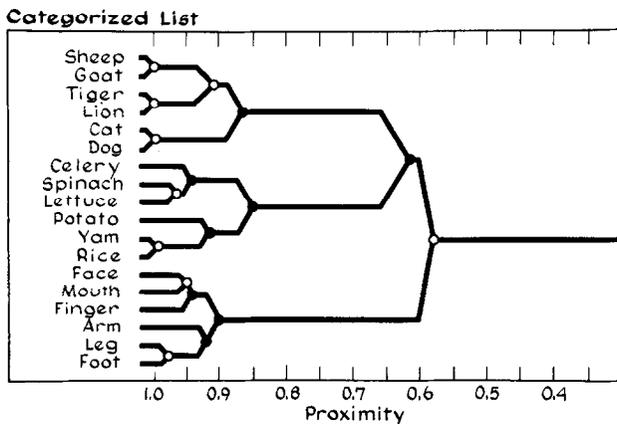


FIG. 4. Complete-link hierarchical clustering for subject A. G., obtained from free recall proximities, pooled over eight trials. The filled nodes represent "significant" clusters indicated by Johnson's (Note 6) cluster statistic.

subjects will utilize structure within memory units to divide them into smaller ones.

Goodness of fit. Since the SL and CL methods will produce tree structures whatever the nature of the data, some means are needed for evaluating the degree to which a hierarchical representation actually is consistent with the data. The question of goodness of fit involves a numerical comparison between the data values and a set of fitted distances implied by the tree structure. In either method, the fitted *tree-distance*, u_{ij} , between items w_i and w_j is defined as the level value, α_l , at which that pair first belongs to the same cluster.

$$u_{ij} = \min_l \{ \alpha_l | w_i, w_j \text{ belong to the same cluster in } C_l \}. \quad (8)$$

Note that u_{ij} and α_l are defined as distances rather than proximities. In examples, where normalized proximities, P' , are used, corresponding distances and cluster values, α_l , will be defined as $1 - P'$. The tree-distances are monotonically related to the proximities as in Eq. (4) and also satisfy the ultrametric inequality (UMI),

$$u_{ij} \leq \max\{u_{ik}, u_{jk}\}, \quad (9)$$

for all triples, w_i, w_j, w_k . The UMI is more restrictive than the triangle inequality, $D_{ij} \leq D_{ik} + D_{jk}$, which must hold for *any* set of metric distances. Furthermore, the ultrametric tree-distances derived from the single-link (u^S) and complete-link (u^C) methods are related to the input distances by

$$u_{ij}^S \leq D_{ij} \leq u_{ij}^C. \quad (10)$$

The importance of this is that when the UMI holds for an empirical distance matrix, the SL and CL methods give identical results and there is an exact equivalence between the distance matrix and a hierarchical clustering (Hartigan, 1967; Johnson, 1967; Miller, 1969). Thus, although the SL and CL methods are opposite extremes in the sense elaborated earlier, the solutions they produce for a given set of proximity data will agree to the extent that the UMI is satisfied. It is important to note, however, that the UMI is a sufficient condition, but not a necessary one for the existence of a "perfect hierarchy," i.e., one in which the item pairs within a cluster at any level in the hierarchy are at least as proximal as item pairs belonging to different clusters (Hubert, 1974a).

Descriptive indices of goodness of fit of a proximity matrix to the UMI were proposed by Hartigan (1967) and by Jardine and Sibson (1971). Approximate overall significance tests for the nonrandomness of an obtained hierarchy were developed by Cunningham and Ogilvie (1972) and Hubert (1974a). Hubert's test applies the γ statistic developed by Goodman and Kruskal (1954) to measure the association between the ranks of item

pairs in the proximity matrix and their corresponding ranks in the tree structure.

The γ statistic may also be used descriptively to assess the similarity of two tree structures on the same set of items. Alternatively, a product moment correlation may be computed between the ultrametric tree-distances resulting in the two tree structures, although this requires stronger assumptions about the metric properties of the proximities. These indices are useful for comparing the organizational structures of different subjects or groups or for comparing the results of different clustering methods applied to the same data.

Although both the SL and CL methods have been applied to memory and learning data, the CL method, on logical grounds, is more appropriate to reveal the type of hierarchical structures considered by organizational theorists (Tulving, 1962a; Mandler, 1967). The complete-link procedure constructs partitions that contain optimally compact clusters of minimum diameter, with all items within a cluster interlinked. This criterion is consistent with an intuitive idea of what subjective memory units should be,⁶ as well as with experimental evidence. That is, in recall a memory unit functions as a redintegrative ensemble such that retrieval of one item within the unit provides, with high probability, access to the remaining items in the unit (Cohen, 1966; Horowitz & Prytulak, 1969; Tulving & Patterson, 1968). Lists with an experimenter-imposed (Bower et al., 1969) or subject-defined (Friendly, Note 4) hierarchical structure lead to high recall because they provide multiple retrieval links for the subject to "move" within higher-order memory units and from one unit to the next at recall (Wood, 1972). Thus, it is not surprising that investigators who have applied clustering methods to episodic memory (Kintsch, 1970; Koh, Vernon, & Bailey, 1971) and semantic memory studies (Anglin, 1970) have tended to favor the CL method, usually on the grounds that it gave more reasonable or interpretable results. Nevertheless, it is shown later that the single-link method corresponds to a natural and reasonable representation of structural organization within the context of a network model, corresponding to the notion of a minimum cost retrieval route on a network.

S-units and clusters. The cluster analyses provide a family of clusterings, rather than just a single partition of the items. In order to identify subjective memory units, it is useful to be able to determine the strengths of clusters formed at different levels in the tree structure. Two features of clusters can

⁶The CL criterion actually seems more severe than is necessary, since it requires clusters to be completely interlinked. If the representation errs, however, it will be in the direction of failing to recognize memory units, rather than producing spurious ones. For some possible ways in which the CL criterion can be relaxed, while retaining the spirit of homogeneity, see Hubert (1974b).

serve as guides in interpreting S-units: *compactness* and *isolation*. These notions may be defined precisely in terms of the cluster analysis.

For the complete-link method, the cluster diameter (the largest intracluster distance, or smallest proximity) provides a natural measure of compactness. The diameter of any specific cluster (w_i, w_j, w_k, \dots) is the cluster value, α_l associated with the first level in which the items w_i, w_j, w_k, \dots are all in the same cluster. In the complete-link solution of Fig. 3, for instance, the diameter of the cluster (*highway, mast, north, structure*) is $1 - .79 = .16$; .79 is the smallest proximity among those items of Table 2.

The cluster diameter, however, gives no information about the relationship between clusters. Two clusters may both be compact, but be so close to each other that one would not want to consider them as distinct memory units. The notion of cluster isolation can be used to distinguish among units at different levels of a hierarchy. The *isolation* of a cluster X expresses the diameter of X relative to the diameter of the first clustering in which X is merged with another cluster. If the cluster X is formed at level g in the hierarchy and the smallest cluster which properly includes X is formed at level h , the isolation of cluster X is $I(X) = |\alpha_g - \alpha_h|$. The isolation of a cluster can be thought of as a measure of the "empty space" or "moat" (Estabrook, 1966) around it. In Fig. 4 the isolation of the animal cluster is $|(1 - .849) - (1 - .629)| = .22$. Ling (1973) developed a similar measure, using proximities converted to ranks, and derived the distribution theory necessary to provide a statistical test.

Using these notions, an S-unit may be defined as a cluster which is sufficiently compact and isolated. Although the criteria for "sufficiently" are left unspecified here, reasonable choices would likely identify reliable S-units. Alternatively, a more objective criterion may be obtained using Johnson's (Note 6) test for reliable clusters, which incorporates both the compactness and isolation criteria. The test statistic, λ , is based on the difference between the average within-cluster proximity and the average proximity between an item in the cluster and an item outside the cluster. Fillenbaum and Rapoport (1971) discuss the use of this statistic to provide a conservative significance test for individual clusters. This test was used to determine the significant clusters in Fig. 4.

Overlapping clusters. While the CL method yields hierarchical representations that conform closely with theories of organization in free recall and are quite reasonable for many types of lists, the requirement that the clusters be strictly nested seems too restrictive for some purposes. As indicated by Shepard (1974), "although hierarchical systems can represent some of the discrete or categorical structure underlying a set of similarity data, it cannot represent psychological properties, however salient, that correspond to overlapping subsets" (p. 413).

A number of clustering procedures which lead to overlapping subsets at each level have recently been proposed. Of these, the B_k procedures

TABLE 4
CLUSTERS OBTAINED BY B_2 METHOD FOR THE DATA OF TABLE 1

Proximity values	B_2 clusters								
1.000	(1,2)	(7,8)							
.984	(1,2)	(7,8)	(9,10)						
.970	(1,2)	(7,8)	(7,9)	(9,10)					
.955 ^a	(1,2)	(5,6)	(7,8)	(7,9)	(9,10)				
.945	(1,2)	(5,6)	(7,8)	(7,9)	(9,10)	(10,11)			
.939	(1,2)	(3,4)	(5,6)	(7,8)	(7,9)	(9,10)	(10,11)		
.909 ^a	(1,2)	(3,4)	(3,6)	(5,6)	(7,8)	(7,9)	(9,10)	(10,11)	
.894	(1,2)	(3,4)	(3,6)	(5,6)	(7,8)	(7,9,10)	(10,11)		
.879	(1,2)	(3,4)	(3,6)	(5,6)	(7,8,9,10)	(10,11)			
.873 ^a	(1,2)	(3,4)	(3,6)	(5,6)	(7,8,9,10,11)				
.864	(1,2)	(3,4)	(3,6)	(4,5)	(5,6)	(6,11)	(7,8,9,10,11)		
.848 ^a	(1,2)	(3,4,5,6)	(6,11)	(7,8,9,10,11)					
.818	(1,2)	(3,4,5,6,11)	(6,12)	(7,8,9,10,11)					
.802	(1,2)	(3,4,5,6,7,8,9,10,11)	(6,12)						
.791	(1,2)	(1,3)	(3,4,5,6,7,8,9,10,11)	(6,12)					
.788	(1,2)	(1,3)	(3,4,5,6,7,8,9,10,11,12)						
.782	(1,2,3)	(3,4,5,6,7,8,9,10,11,12)							
.773	(1,2,3,4,5,6,7,8,9,10,11,12)								

Note. The items are numbered as in Table 1, i.e., 1, Inventor; 2, Professor; 3, Highway; 4, Mast; 5, North; 6, Structure; 7, Assault; 8, Quarrel; 9, Captive; 10, Execution; 11, Decree; 12, Urge.

^a These clusters are shown graphically in Fig. 5.

developed by Jardine and Sibson (1968, 1971) are relatively simple to understand, have been studied widely (Hubert, 1974b; Rohlf, 1975), and have been implemented in several computer algorithms (Cole & Wishart, 1970; Jardine & Sibson, 1971, Appendix 3; Rohlf, 1974; Shafto, 1974). The B_k methods constitute a generalization of the single-link method in which any pair of clusters at a given level is allowed to overlap to at most $k - 1$ items in common. B_1 produces disjoint clusters and is identical to the SL method; B_2 yields clusters which may share one item, etc. As the amount of overlap allowed increases, successively better fit to the original proximities is obtained, though at the cost of increased complexity in the representation.

Since the B_k methods are rather close in spirit to single-link clustering, they do not adequately capture the homogeneity criterion which is required to identify subjective memory units. The intent of this discussion is to indicate the possibility for representations of memory structures in terms of overlapping subsets, rather than to suggest that the B_k methods provide a fully satisfactory solution. Alternative methods for overlapping cluster representations, which correspond more closely to current notions of

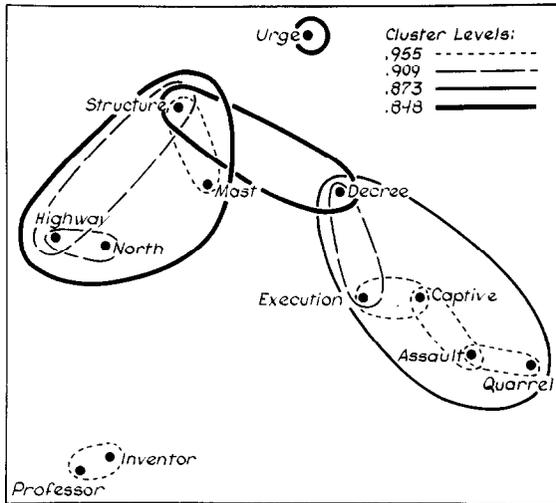


FIG. 5. Clusters found by the B_2 method, mapped on the MDS space of Fig. 2.

subjective memory units, may be found in Peay (Note 9, Note 10; also described in Fillenbaum & Rapoport, 1971), Constantinescu (1966), and Hubert (1974b).

To illustrate the type of structures which result when memory units are allowed to overlap, the B_2 clusters derived from the unrelated list proximity matrix (Table 2) are listed in Table 4. Since the overlap allowed between clusters, even for $k = 2$, makes this type of representation somewhat more difficult to apprehend, it is often convenient to display the clusters formed at a small number of levels. The result is shown in Fig. 5, where the clusters at four levels have been mapped onto the two-dimensional MDS space of Fig. 2.

Network Representations

In the network models considered earlier, it is generally assumed that the subjects' retrieval processes search through the memory network, following tagged or activated links. In this context, the proximities among items in free recall output order may be considered to reflect the pattern of links among item nodes in this structure. Representations of the network structure may be obtained by applying the methods and results of graph theory to the proximities. Indeed, the types of network structures which theorists have proposed as models for a memory data base correspond closely to various types of abstract graph-theoretic systems developed in combinatorial mathematics. The literature of graph theory is extremely extensive, however, and a thorough discussion of even those aspects which are relevant to the present purposes would require more space than

is available, as well as the introduction of considerable notation and terminology. The discussion that follows is therefore highly selective. In particular, no discussion is given to the area of random graph theory (e.g., Erdős & Rényi, 1960) or its application (Rapoport, Rapoport, Livant, & Boyd, 1966; Fillenbaum & Rapoport, 1971) to graphs constructed by subjects in semantic memory tasks. Nor is much attention per se given to directed graphs (Harary, Norman, & Cartwright, 1965) which has been used by Allen (Note 2) to develop empirical techniques for identifying memory units in free recall and by Hollan (1974) to study the "epistemological structure" of concepts in mathematics. For a general, elementary discussion of graph theory, see Ore (1963); a more advanced treatment is given in Harary (1969), Busacker and Saaty (1965), and Marshall (1971). These latter texts should be referred to for a rigorous discussion of what is presented here informally.

Graph theory deals with structural properties and relations between the elements of two sets, *points* and *lines* (alternatively, *nodes* and *edges*). In this discussion a node will generally correspond to the internal representation ("trace") of a word concept in a person's memory, while an edge may be interpreted as the mental "link" or "connection" between traces. Various types of graphs are distinguished by the properties ascribed to the points and lines. An *undirected graph* consists of a set of nodes, W , together with a collection of edges, E , denoted as unordered pairs (w_i, w_j) of elements of W . If the lines of a graph have a direction assigned to them, then the graph is a *directed graph* or *digraph*. If numerical values are also associated with directed lines, the graph is called a *network*. (Use of the term "network" to characterize the general class of structural memory models under discussion will be distinguished from its graph theory usage by context or by modification, e.g., "memory network" or "graph network.")

In the simplest type of network memory model (e.g., Pollio, 1966), items in memory are assumed to be connected or associated in an all-or-none fashion. This structural model is a realization of a *simple undirected graph*. In a *simple graph* no node may be linked to itself, and multiple edges are not allowed (i.e., at most one link between a given pair of nodes).

A proximity matrix, \mathbf{P} , on a set of items, w_1, w_2, \dots, w_L can be represented by simple graphs in several ways. First, for any given proximity value, s , a corresponding *threshold graph* G_s may be defined where w_i and w_j are linked if and only if $P_{ij} \geq s$. This definition reduces the proximity matrix to a graph whose links correspond to proximities at least as great as the threshold, s . Choosing a high value for s will yield a graph with few connections and many *isolated* nodes, i.e., items with no links; a low value results in a graph with a great many links. Waern (1971) has considered ways of representing the structure in a proximity matrix by

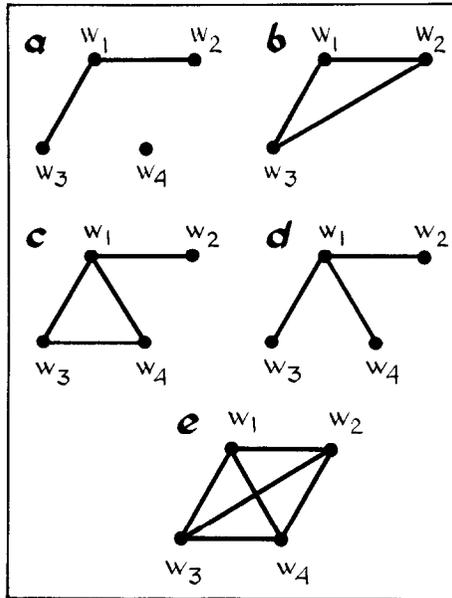


FIG. 6. Graphs illustrating some graph theory concepts.

such graphs, using either a single, well-selected threshold value, or several ordered thresholds simultaneously.

Further, the set of graphs $\{G_s\}$ generated by considering *all* distinct proximity values in \mathbf{P} bear strong relations with the clusters identified by the single-link and complete-link methods (Hubert, 1974b). In order to describe these relations and other structural features of graphs efficiently, a number of graph theory terms are defined here informally and illustrated with the graphs in Fig. 6.

A *subgraph* of a graph G consists of selected nodes and edges of G . A *maximal* subgraph consists of a subset of the nodes of G together with all links of G which join pairs of nodes included in the subset. A *spanning* subgraph of G contains all of the nodes of G together with a subset of the links of G . In Fig. 6, graphs a , b , c , and d are all subgraphs of e , while d is a subgraph of c . Graph b is a maximal subgraph of e defined on the nodes w_1 , w_2 , w_3 ; graphs a , c , and d are spanning subgraphs of e .

Several graph theory concepts deal with groups of interconnected items. A graph or subgraph is *connected* if every distinct pair of nodes is joined by at least one *path*, i.e., a series of intervening linked nodes. A (sub)graph is *complete* if and only if every pair of nodes is directly linked. In Fig. 6, graphs b and e are complete, while b is a maximal complete subgraph of e . Graphs b , c , d , and e are all connected, and b is a maximal connected subgraph of e . Completeness is thus a stronger condition than connectedness, since all complete (sub)graphs are connected, but not conversely.

Finally, a subgraph of a connected graph G is called a *spanning tree* if it contains no *closed* paths (i.e., a series of linked nodes which form a closed curve) and includes all nodes of G . Graph d in Fig. 6 is a tree spanning graph e .

Using these concepts, the clusters formed in the single-link algorithm can be characterized in a way which leads to representations of memory structure which seem natural within the context of network models of memory. Reference to the description given earlier shows that the SL clusters at a given level, α_l , consist of the maximal connected subgraphs of the threshold graph at that proximity level, G_{α_l} .⁷ The graphs at levels α_0 (the disjoint clustering) and α_L (the conjoint clustering) are noninformative. At the $L - 1$ intermediate levels, items w_i and w_j belong to the same cluster at level α_l if they are connected by a path in G_{α_l} . In a network memory model, such a connected path corresponds to the notion of a retrieval route, i.e., a succession of memory nodes joined by links which have been activated or strengthened during prior study and retrieval. Sets of stimulus words which form serial chains could be expected to be organized in these connected paths by all or most subjects. For example, in mediated associative chains (e.g., Shapiro & Palermo, 1967) such as *billiards, pool, water, tap, dance*, each item has strong, normative associations to its immediate neighbors, but is not directly linked to nonadjacent items. This type of chaining would also be expected when individuals organize a list alphabetically (Tulving, 1962b). For unrelated lists, it is also plausible for subjects to develop consistent retrieval plans involving search through a network, although the particular configurations of linked nodes will vary from one person to the next.

Minimum spanning tree. The SL clusters in the set of threshold graphs $G_{\alpha_1}, \dots, G_{\alpha_{L-1}}$ contain rather complete information regarding the connectedness of a memory network derived from interitem proximity in free recall. The information in these graphs, however, is difficult to apprehend for lists of reasonable length. A more compact representation of memory structure may be obtained in terms of minimum spanning tree (Gower & Ross, 1969; Rohlf, 1974) defined on the matrix of single-link tree distances.

In the graphs $G_{\alpha_1}, \dots, G_{\alpha_{L-1}}$ items are linked at level α_l if and only if $u_{ij} \leq \alpha_l$, but all links which appear are equivalent in simple undirected graphs. In contrast, links in the *minimum spanning tree* (MST) are assigned values equal to the cluster levels, α , at which components first become connected, and only a single such graph with real-valued links is necessary to represent the structure contained in the proximity matrix. The minimum spanning tree is then defined as the spanning tree with minimum total length

⁷ Analogously, CL clusters may be described as maximal complete subgraphs in G_{α_l} . However, as noted earlier, the CL clusters are closer in spirit to S -units discussed in the context of the tree structure models.

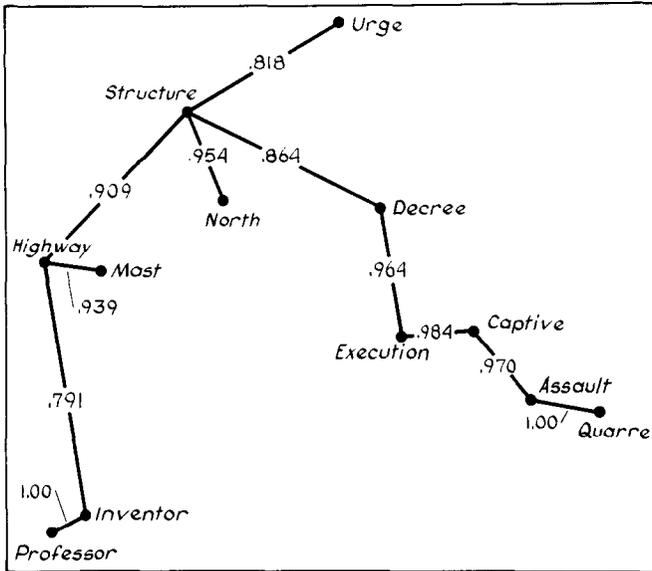


FIG. 7. Minimum spanning tree representation of the proximities in Table 2. The values assigned to the links correspond to the entries marked with asterisks (*) in Table 2.

of all links. (For a clustering performed on proximity values, the MST would have maximum sum of proximity values for any spanning tree.)

If connected graph components are interpreted in a network memory model as retrieval routes in this structure, and the proximity between items w_i and w_j is considered to index (inversely) the retrieval cost of accessing item w_i from w_j (or vice versa), then the MST is readily interpreted as a retrieval path of minimum overall cost or effort. The MST may be derived directly from the proximity matrix (e.g., Prim, 1957; Gower & Ross, 1969), or it may be found from the ultrametric tree distances (u_{ij}^S) resulting from the single-link method (Rohlf, 1974). When the proximities are distinct,⁸ there will be exactly $L - 1$ elements among the u_{ij} for which $u_{ij} = D_{ij}$, and for these elements equality holds in Eq. (10). These values are also the cluster values $\alpha_1, \dots, \alpha_{L-1}$ and the item pairs for which this equality holds correspond to the links in the minimum spanning tree; the (cost) value assigned to the link is u_{ij} .

The MST derived from the data in Table 2 is shown in Fig. 7. For reasons explained below, the locations of points in this figure were taken from the MDS configuration presented earlier (Fig. 2). The reader can verify that the links in this figure connect item pairs whose proximity value is marked with an * in Table 2, and that the values assigned to the links are the cluster

⁸When there are ties among the proximities, the MST will not be unique, as there will be more than one configuration of links with minimum sum, and more than $L - 1$ equalities will occur. In such cases, one can select one link from each set of tied values.

values in the single-link solution (Fig. 3, left) at which members of the linked pair first belong to the same cluster.

The single-link hierarchy in Fig. 3 and the MST in Fig. 7 are isomorphic, in the sense that each contains all the information needed to construct the other. Nevertheless, the MST representation has two virtues. First, since the nodes in a graph may be positioned arbitrarily, the MST may be embedded freely in a spatial configuration to give added information about the structural relations among items. In the SL method it is the small distances (large proximities) which have the greatest influence in determining the tree. MDS on the other hand, apparently gives all proximities equal weight. Thus, an MDS spatial representation may contain some information which is not preserved in the single-link hierarchical clustering from which the minimum spanning tree is derived. Second, the MST shows the specific interitem links in a representation which may be interpreted as having overall minimum retrieval cost. This format is useful in that it may stimulate research questions different from those suggested by the hierarchical clustering scheme.

These graph theory methods can be used to provide meaningful representations of memory structure which are compatible with a number of network models reviewed earlier. Some mention should be made, however, of the types of memorial information which are beyond the scope of these methods. First, these methods are not quite powerful enough to *completely* represent the type of activated structural network discussed by Kiss (1967, 1969) and Fridja (1972). In particular, these models make the reasonable assumption that links between items in the memory structure vary in strength. Although the value assigned to a link in the MST may be considered to be an index of the strength of the relationship between the items in memory, the MST is best interpreted as a *path* through the memory network, rather than as the complete network itself. Kiss (1968, 1969) has shown how the more general methods of graph networks and signal-flow graphs may be applied to proximity data obtained in word association tasks. These methods assume, however, that the proximities have at least interval scale properties. As the interest in nonmetric (i.e., ordinal) scaling attests, this assumption is often difficult to justify. One possible solution to this problem would be to subject the proximities to a preliminary transformation to bolster their metric properties and then perform a network analysis on the transformed values. Such a procedure has in fact been suggested by Cunningham and Shepard (1974). They describe a procedure of "maximum variance nondimensional scaling" (MVNDS) for transforming ordinal proximity measures into distances in a general metric space. The transformed values are required to satisfy the metric axioms (e.g., the triangle inequality), to be as nearly monotonic with the original proximities as possible, and to have maximum variance. Cunningham and Shepard (1974) note that they

are currently exploring the possibility of using the maximally additive distances determined by MVNDS to construct a graph or network with n nodes, representing the n objects, and a minimum set of weighted internode links such that the distances determined by MVNDS are, as nearly as possible, approximated by the sum of weights over the minimum connected path between appropriate pairs of nodes in the network. Such an adjunct to MVNDS appears quite straight-forward and, as we have suggested, might provide useful representations of the structure underlying similarity data from various semantic and cognitive domains (p. 361).

Second, the methods of graph theory, including the more general ones just mentioned, cannot represent the type of relational network models developed by semantic memory theorists (Anderson & Bower, 1973; Kintsch, 1974; Norman & Rummelhart, 1975; Quillian, 1968). In these models interitem links in the memory structure are labeled according to the type of relation which connects them. Structurally, these models can be described as *directed edge-labeled graphs*, but this appellation does not provide a means of determining such structures on the basis of empirical data from free recall or from semantic memory tasks. In the present context, this means that although the methods of graph theory can be used to determine which items are linked together in an individual's memory in free recall, they do not directly show the qualitative ways in which links differ. The type of relations represented by the links must be determined by the investigator as part of the interpretation of the graph.

Individual Differences

It was noted earlier that the averaging of proximities over subjects in a group is only appropriate to the extent that the subjects share a common organization of the list. In practice this condition tends to be fulfilled when subjects learn categorized lists. However, subject differences in organization are more prevalent if unrelated lists are used. When analysis of group proximities must be ruled out due to heterogeneity of organizational structure, the alternative of a separate analysis for each individual may be unattractive because it is unwieldy. In such cases it is desirable to have a technique to determine the nature and extent of individual differences in organization.

In the context of the dimensional models, a number of methods are available for analyzing individual differences in MDS (Carroll & Chang, 1970; McGee, 1968; Tucker & Messick, 1963). The Carroll-Chang INDSCAL method assumes that a single set of dimensions is common to all individuals but that these dimensions are differentially salient or important to different individuals. The weighting of dimensions has the effect of differentially stretching or shrinking the common group configuration in directions corresponding to the coordinate axes. This method is based on the assumption that the interitem proximities are linearly related to distances in space, rather than being merely monotonically related as in

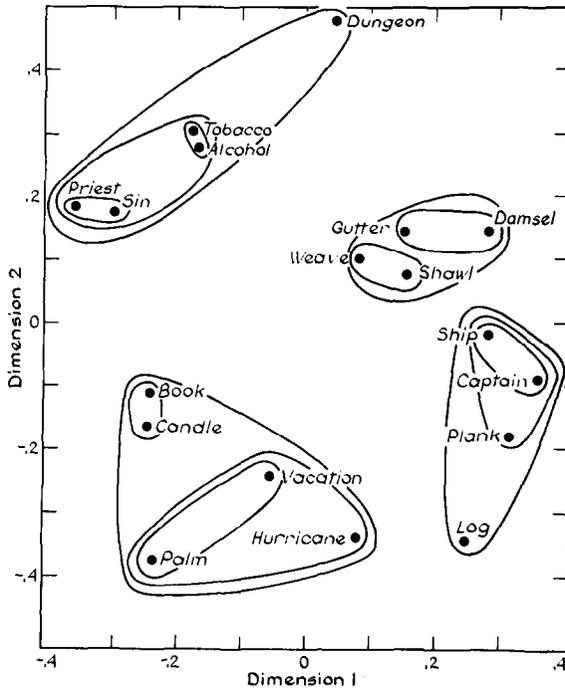


FIG. 8. Two-dimensional group configuration of list items resulting from individual differences analysis of recall proximities.

nonmetric MDS. Nevertheless, to the extent that this stronger assumption is fulfilled,⁹ the Carroll-Chang method leads to a representation which provides valuable information about the extent of individual differences within a group of learners and helps to characterize the nature of these differences.

In the demonstration experiment described previously, the subjects also learned a second list. This list consisted of 18 concrete words which were neither related nor completely unrelated, but rather were chosen to be relatable in different ways. Eight presentation-recall trials were given to each of the six subjects. Proximities over all trials were computed for each subject, and the Carroll-Chang INDSCAL program was run on these data. In addition, the group average proximity matrix was included in the analysis to serve as a reference point. The results indicated that a

⁹A strong suggestion that the proximities approximately satisfy the linearity condition was provided by the MDS solutions for the data presented in this paper, as well as by several other sets of data which have been run. In these cases, the plots of the original proximities from individual and group data against the recovered distances in MDS solutions showed no systematic deviations from linearity. The exceptions to this statement occurred for lists with salient taxonomic categories. For categorized lists, these plots have the appearance of a step-function.

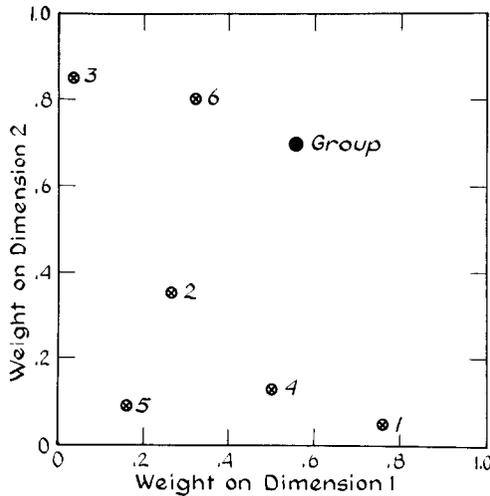


FIG. 9. Subjects' weights on the dimensions of Fig. 8 derived from INDSCAL analysis.

three-dimensional solution provided an excellent fit to the data, accounting for 96% of the variance in the proximities. A two-dimensional configuration did not fit as well (variance accounted for = 62%), but is simpler and is adequate for the purpose of illustration here. The two-dimensional group stimulus space, together with the complete-link clusters obtained from the group average proximity data, is shown in Fig. 8. The subject weights on these two dimensions are plotted in Fig. 9. This plot shows considerable difference among subjects in their weighting of the dimensions. It is apparent that the group average data do not represent the data of any of the subjects particularly well. Nor is it the case that the subjects fall into distinct subgroups corresponding to different organizational schemes. Subject 2 and the group average data weight the two dimensions of Fig. 8 about equally; the results of separate MDS analyses of these two sets of data are very similar to Fig. 8. Subjects 4 and 1 have moderate or high weights on dimension 1, but negligible weights on dimension 2. The consistent recall units for these subjects include (*priest, palm, candle*) and (*dungeon, plank, log*) which cut across dimension 2. Subject 3, on the other hand, had units like (*tobacco, alcohol, gutter*), (*palm, hurricane, log*), and (*sin, damsel, captain*), which result in a negligible weight on dimension 1.

Aside from the relative weighting of dimensions, the positions of subject points in Fig. 9 also indicate the goodness of fit of the Carroll–Chang model to an individual's data: The greater the distance from the origin, the higher the proportion of variance accounted for. Interestingly, the group average proximities are fit best in the analysis, which reflects the reduction of noise achieved by averaging over individuals. In addition, Tulving's *SO* measure (Tulving, 1962a), an index of amount of subjective organization, was found

to correlate strongly with the variance accounted for in the INDSCAL analysis, $r = .78$. This result serves to underline the importance of examining the structure of organization rather than just its amount: Subjects 3 and 4 had nearly the same *SO* scores, but as Fig. 9 shows, their organizational schemes were quite different.

Individual differences in organizational structure can also be explored with hierarchical representations. The basis for one simple method is provided by the measures of similarity between two HCS representations mentioned previously. Since an HCS is equivalent to the matrix of ultrametric tree distances, \mathbf{u} (Johnson, 1967), the similarity of two hierarchies can be assessed by computing a measure of similarity between the entries in the corresponding matrices. Gruvaeus and Wainer (1972) have suggested using the Spearman rank-order correlation, ρ , or Pearson's r when the proximities have interval scale properties. The Goodman-Kruskal γ suggested by Hubert (1974a) is superior to ρ in that it is not affected by ties, although it is somewhat more difficult to compute.

Using one of these measures a second-order clustering of the subjects can be performed as follows: From the individual cluster analyses for S subjects, obtain $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_S$. Then compute the similarity index for all pairs of subjects and array these in an $S \times S$ matrix, \mathbf{R} . A cluster analysis of \mathbf{R} will then indicate subgroups of the subjects whose organizational groupings produce similar hierarchical representations. This approach is illustrated later in this paper.

VALIDATION

This section presents some data obtained to test the validity, sensitivity, and utility of the approach to recall organization described here. No attempt is made to determine which of the three representational formats is most valid, sensitive, or useful. As indicated previously, each may have its place in the context of different theories and with different types of word lists. It is primarily the nature of the proximities among items in recall and the extent to which they can be said to capture the structure of organization which is at issue here. Tree-structure representations are employed where the format is immaterial to the discussion.

Recovery of Known Organizational Schemes

The minimal test which the method must pass is that it should perform appropriately when the organizational scheme used in recalling a list may be specified a priori. To provide such a test, artificial recall protocols were created for two types of conceptual subjects. For both groups the list consisted of 20 items, grouped into four categories (labeled *A, B, C, D*) with five exemplars each (e.g., *A1, A2, . . . , A5*). Four trials were constructed for each subject according to a "some-or-none" model

(Cohen, 1966): On trial t , a subject recalled a given number of items (say, r) from each of t categories, and none from the remaining categories. Schematically, the pattern of recalls can be indicated as,

Trial 1: A^r

Trial 2: $A^r B^r$

Trial 3: $A^r B^r C^r$

Trial 4: $A^r B^r C^r D^r$

where for example, A^r means r items recalled from category A . Thus these artificial subjects "learned" by remembering one additional category on each trial. The number of items per category varied across subjects from $r = 2$ to $r = 5$. For half of the subjects in each group the pattern of category recall was as shown above; for the remaining subjects, the assignment of categories to trials was reversed, i.e., the categories recalled on the four trials were D^r ; $D^r C^r$; $D^r C^r B^r$; $D^r C^r B^r A^r$. This provided some balance in the frequencies of recall of the individual categories across subjects.

The organization for one group consisted of independent clusters corresponding to the categories of the list. These protocols were constructed by randomly ordering the categories recalled on each trial and randomly ordering the positions of items within categories. As indicated previously, the complete-link HCS is most appropriate for this type of organization and should display the categories as compact, isolated clusters.

The second group of subjects recalled according to a linear organization, such as might be expected with alphabetical organization or mediated associative chains. In these protocols the categories recalled on a given trial appeared either in alphabetic order ($ABCD$) or the reverse ($DCBA$). The items appearing in a given category came either from the beginning of the category (e.g., $A1$, $A2$) or the end ($A4$, $A5$) and appeared either in numerical order or the reverse ($A2$, $A1$ and $A5$, $A4$). For this group either the single-link clusters or the minimum spanning tree representations are appropriate. The MST representation would be expected to show four regular chains (one for each category) with one link between each successive pair of categories.

There were 16 artificial subjects in each group. Collectively, the two groups recalled at exactly the same level since the protocols for the cluster group were constructed by randomizing those of the linear group on a subject-by-subject basis. Any differences in the organizational structure of these groups must be a consequence of different order of recall, and cannot be due to differential recall.

In analyzing these data, the proximity measure described here was compared to a measure of proximity in recall which has been used by others

TABLE 5
 BADNESS-OF-FIT MEASURES FOR HIERARCHICAL CLUSTERING (HCS) AND
 MULTIDIMENSIONAL SCALING (MDS) REPRESENTATIONS
 OF ARTIFICIAL DATA

Measure	Group	HCS badness-of-fit (Δ)		MDS stress
		Single-link	Complete link	
Adjacency	Linear	.518	.315	.118
	Cluster	.401	.220	.286
Proximity	Linear	.047	.064	.034
	Cluster	.101	.043	.280

(Anglin, 1970; Kintsch, 1970; Koh et al., 1971; Schwartz & Humphreys, 1973). These investigators have essentially used the proportion of times a pair was recalled in *adjacent* output positions as their index of proximity. As discussed earlier, the criterion of adjacency in output discards much of the information about organizational structure in a subject's order of recall. The proximity measure, which takes into account varying degrees of contiguity, should therefore do better. To provide an empirical test of this statement, both measures were computed for each group, in all cases averaging across subjects.

Both the single-link and complete-link hierarchical clustering solutions were obtained for each of the four matrices (two groups \times two measures). Goodness of fit of the data to the tree structures was assessed using the normalized root-mean-squared error suggested by Jardine and Sibson (1971):

$$\Delta = \left[\frac{\sum (D_{ij} - u_{ij})^2}{\sum D_{ij}^2} \right]^{1/2},$$

where u_{ij} is the tree distance defined in Eq. (8). Although significance levels for this statistic are unknown, values below .10 may be regarded as indicating an adequate fit. In addition, two-dimensional MDS representations were obtained for the four matrices. The stress value for each solution, as well as the HCS badness-of-fit measure, Δ , is shown in Table 5. The solutions for the adjacency measure are shown in Fig. 10; solutions obtained using the proximity measure appear in Fig. 11. It is clear from Table 5 and Fig. 10 that the adjacency measure in general is too weak to pick up the recall organizations constructed in the artificial data. In all four cases the fit of the data to the tree structure is intolerably poor. The complete-link solutions show rather diffuse clusters which do not adequately capture the strong categorical structure of the list. The

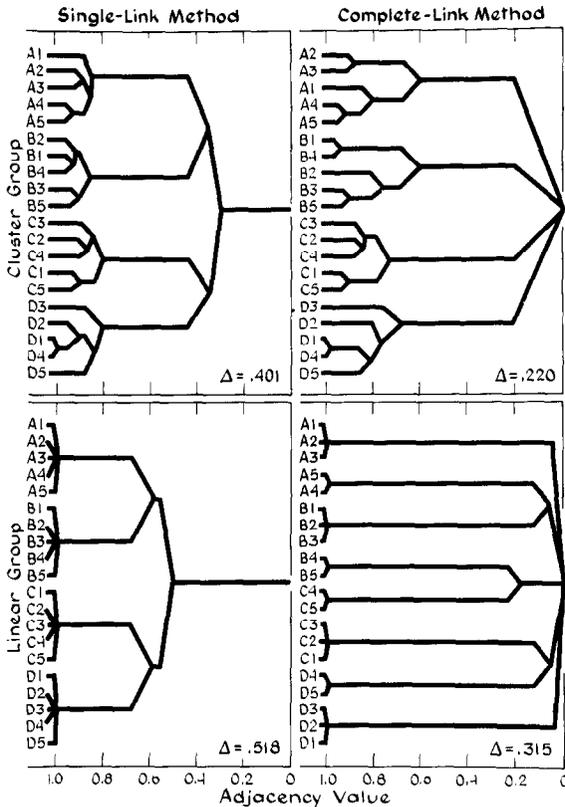


FIG. 10. Hierarchical clustering solutions obtained using adjacency measure.

single-link solutions fare better in this regard, because they depend on only the local interitem relations to which the adjacency measure is sensitive.

The values in Table 5 indicate that all solutions for the proximity measure are at least adequate fits to the data. For the data of the Cluster group (top half of Fig. 11) the complete-link solution has the best fit. Furthermore, it shows the expected grouping of the items into four clusters which are tightly integrated internally (the least within-cluster proximity is .9), but isolated from each other. In the data of the linear organization group, the single-link solution fits best, as expected. The merging of all items within each category at the *same* level ($P' = 1.0$) is characteristic of linear schemes of organization. The minimum spanning tree representation in Fig. 12 shows the linear structure more clearly. Again, the locations of the points in the MST graph were found from a two-dimensional multidimensional scaling solution. It is worth noting that of the four matrices, only the one based on the proximity measure for the linear group resulted in an adequate stress in the MDS solutions (see Table 5). The horizontal spacing

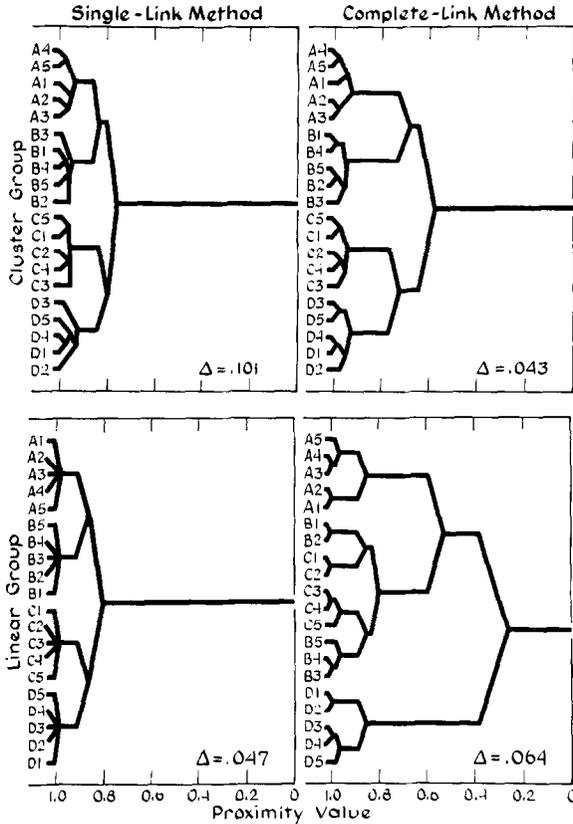


FIG. 11. Hierarchical clustering solutions obtained using proximity measure.

of the categories and the vertical spacing of the items within categories arises from the linear ordering of recalls and the counter balancing design on the categories and items. All links within a category in the MST have a proximity value of 1.0, reflecting the fact that the linked pairs were always adjacent in this organizational scheme.

Thus, the analyses based on the proximity measure showed that the organizational structures built into the artificial data, independent clusters and linear chains, are revealed by the appropriate representations. On the other hand, the results obtained from the adjacency measure were not nearly as convincing, even though the organization of the artificial protocols was strong and completely consistent among subjects in each group. The difference between these measures is likely to be magnified for experimental data, which are typically less strongly organized and more variable from subject to subject. In the light of these results it is not surprising that Kintsch (1970) seemed less than enthusiastic about his

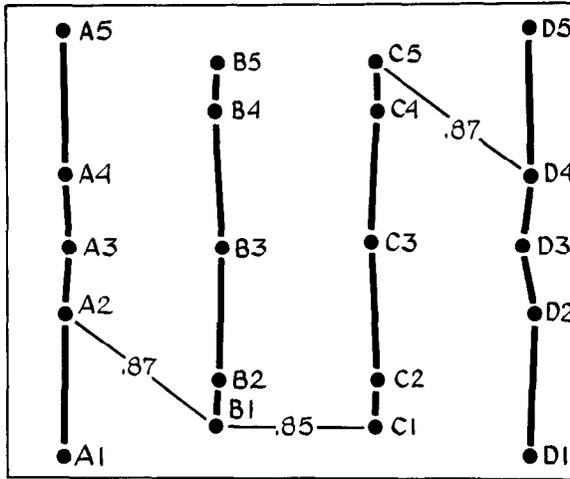


FIG. 12. Minimum spanning tree representation for the linear group artificial data (proximity measure). All within-category links have the value 1.0; values for the three intercategory links are as shown. The positions of points were determined from a two-dimensional MDS solution.

results based on the adjacency measure and found no consistent organization in recall of an unrelated list.

Sensitivity to Structural Differences

These results show that the proximity method can recover built-in organization and distinguish between grossly different structures. The next question investigated is whether the method can distinguish among small or subtle variations in organizational structure. For this purpose, some results from an experiment involving hierarchical organization in free recall (Friendly, Note 4) are presented here.

The experiment was designed to determine whether different subjects could be induced to organize a single list differentially by manipulating presentation order alone. The list contained 42 items structured in taxonomic categories nested at three levels. All of the items were edible substances (level 1). At the second level, the items belonged to three 14-item categories (*Seafood, Farm produce, Animal foods*). Each of these major categories contained two 7-item subcategories (e.g., *Farm produce* contained *Fruit* and *Vegetables* at level 3).

Four groups of subjects received presentation blocked at different levels of the hierarchy. In blocked presentation the items within a "block" are always presented together, but the order of the blocks and the order of items within a block are randomized from trial to trial. Three experimental groups differed according to whether the items were blocked into major

categories at level 2 of the hierarchy (Group B2), blocked according to the six minor categories at level 3 (Group B3), or blocked according to both level 2 and level 3 categories (Group B4). A control group (Group B1) received a different completely random ordering of the stimulus list on each trial. Each subject received 12 presentation-recall trials. The number of subjects per group ranged from 35 to 42.

In recall, Group B4 recalled the greatest amount and Group B1 recalled the least. The difference between Group B4 and the two groups which received presentation blocked at just one level of the hierarchy (B2 and B3) was not large, however. For the present purposes, the interesting results concern the representations of organizational structure obtained from proximity analysis. Average interitem proximities were computed for each group, averaging over subjects and trials. The proximities were analyzed by the hierarchical clustering procedure. The complete-link solutions for Group B2 and B3 are shown in Figs. 13 and 14. These figures, as well as

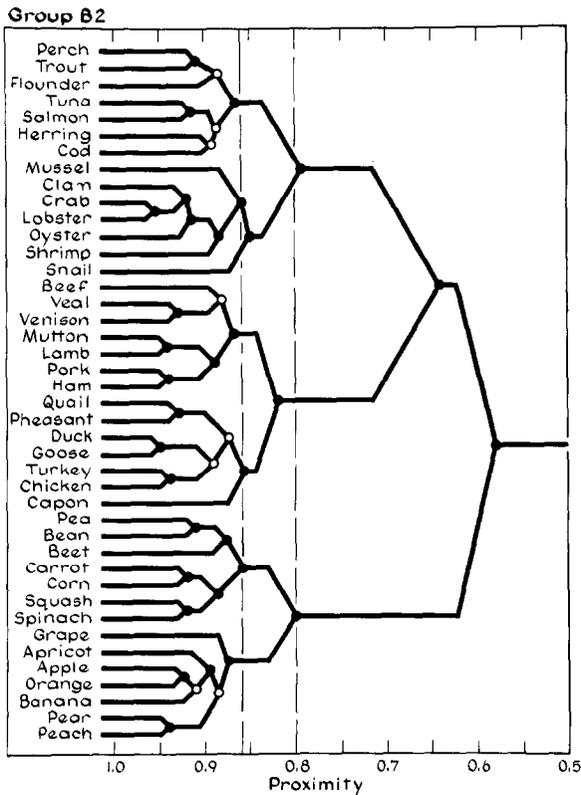


FIG. 13. Organizational structure (complete-link clustering) for Group B2. Filled nodes indicate clusters which appeared identically in the complete-link and single-link methods. Data pooled over subjects and trials.

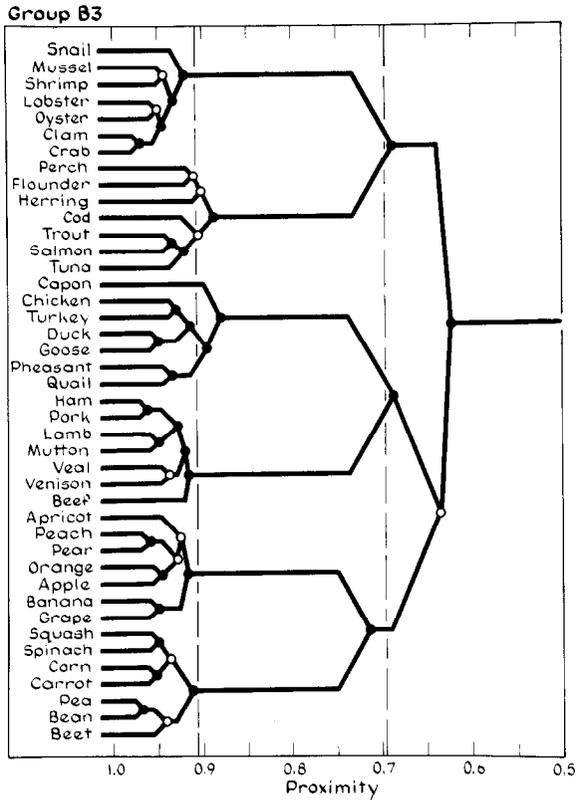


FIG. 14. Organizational structure (complete-link clustering) for Group B3. Filled nodes indicate clusters which appeared identically in the complete-link and single-link methods. Data pooled over subjects and trials.

corresponding solutions for Groups B1 and B4 (not shown here) look qualitatively similar. The tree structures for all four groups match the hierarchical structure of the list. This is not surprising since (a) the clustering method always fits a hierarchy, and (b) subjects tended to discover and utilize the additional levels of list structure not made obvious by presentation order.

The differences in recall structure among the four groups are reflected quantitatively in the tree diagrams in the relative strengths of the category systems at level 2 and level 3 of the hierarchy. The dashed lines in Figs. 13 and 14 show the average diameters of the level 2 and level 3 categories. The separation between these lines represents the (average) isolation of the level 3 categories. For Group B2, the clusters which are both compact and isolated are those at level 2, corresponding to the blocking condition which that group received. In contrast, the compact, isolated clusters in the

Group B3 M-gram are those at level 3. Again this corresponds exactly to the presentation condition of that group.

These differences in cohesiveness of the organizational structures at the two levels of the list hierarchy may be seen more readily in Fig. 15, which displays the diameter and isolation measures for the four groups. It can be seen that the strength of organization of the level 3 categories increases steadily (diameters decrease) from Group B1 to Group B4. A different picture is presented by the diameters at level 3 and the degree of separation between the two organizational schemes. Subjects whose presentation was blocked at level 2 (B2) have the most cohesive organization at this level and their clusters at level 3 are least isolated. The reverse is true of Group B3: These subjects have the weakest organization at level 2 and the greatest isolation between the two systems of categories.

The diameters of clusters obtained from the hierarchical clustering evidently convey fairly detailed information about organizational structure. In fact, even if the nature of the list organization and of the blocking conditions were not known, the sequence of cluster diameters, $\alpha_0, \alpha_1, \dots, \alpha_L$, could by themselves provide strong indications of a dual-level list structure in which the strengths of the levels varied across groups. Figure 16 shows a plot of the successive differences of diameter values, $\alpha_i - \alpha_{i+1}$, against the number of clusters at level i in the complete-link hierarchies. The "spikes" appearing at 3 and 6 clusters show

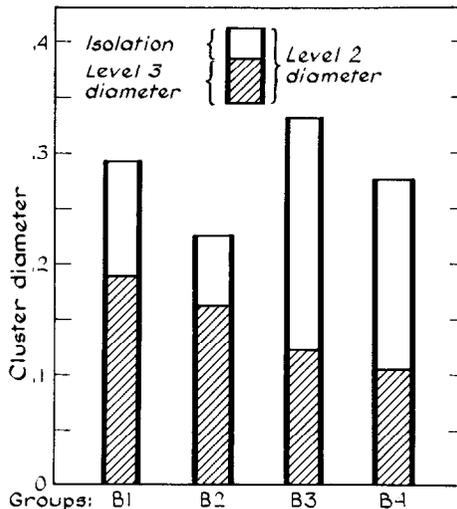


FIG. 15. Cluster isolation and diameter from complete-link HCS. For each group, the total height of the bar represents the mean diameter of the major (level 2) categories. The shorter the bar, the more tightly knit is organization at this level. The average diameters of the minor categories (level 3) are indicated by the filled portions of the bars, while the lengths of the unfilled portions indicate the isolation of these two modes of organization.

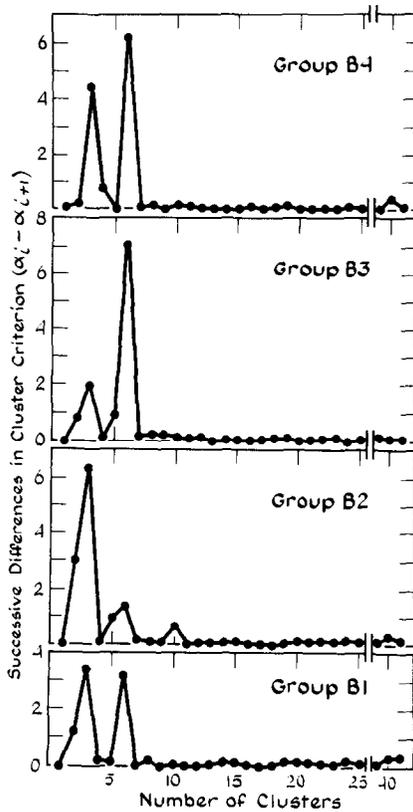


FIG. 16. Successive differences in cluster diameters vs number of clusters at that level. The peaks at the 3 and 6 clusters correspond to the major and minor categories, respectively.

the same pattern as that displayed in Fig. 15. Experience has shown that successive difference plots such as Fig. 16 provide a valuable means of identifying real clusters.

Sensitivity to within-group differences. The analyses of organizational structure described above were based on the average proximities for each group, and therefore reflect the aspects of organization common to each group as a whole. Although these results are relatively clear-cut, an even stronger demonstration of the sensitivity of proximity analysis would be provided by showing that the method can distinguish among identically treated subjects on the basis of their output organizations. Therefore the Gruvaeus-Wainer (1972) individual differences procedure was used to obtain rank-order correlations between the complete-link solutions for all pairs of subjects in each group. The average proximity matrix for each group was included in this analysis as a point of reference. In general, the correlations were quite high; the median intersubject correlations were .65,

.74, .80, and .83 for Groups B1 to B4. Thus, as the degree of structure present in the input order increased, so too did the agreement among subjects of their output structure.

A cluster analysis was then performed on the intersubject correlation matrix for each group. The results of this analysis showed that subjects within a group tended to vary in the degree to which their organization resembled the modal organization for the group, rather than forming homogeneous subgroups. Each group was divided into roughly equal halves: those whose organization was most like ("central") and least like ("remote") the average for the group. Pooling the proximities within each subgroup separately, it was found that the remote subjects differed mainly in that their organization was less cohesive at the level of the minor categories of the hierarchy (see Fig. 17). Some qualitative differences between remote and central subjects in the pattern of organization were

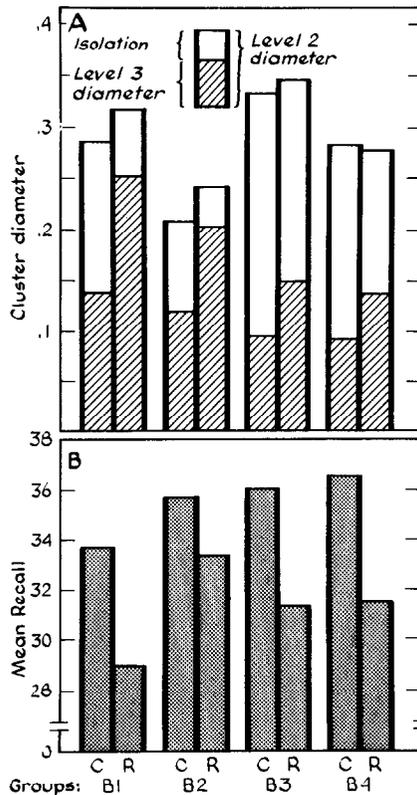


FIG. 17. Category diameters and mean recall for empirically isolated subgroups. C, central subgroup; R, remote subgroup. Panel A shows mean diameters of minor categories (shaded portion) and major categories (total height). Panel B shows average recall over all trials for the subgroups.

also apparent. For example, most of the remote subjects in Group B2 organized the items according to the three major categories with little subgrouping according to the minor categories, while the central subjects in this group tended to organize in terms of both levels. Many of the remote B3 subjects also organized primarily at one level: that of the minor categories.

The category diameters determined for these subgroups appear in Fig. 17 which also shows performance in recall, averaged over trials for each subgroup. Comparison of the shaded portions of the two panels shows that recall varies directly with the cohesiveness (inversely with diameters) of the level 3 categories. The recall results are quite surprising. They indicate that the difference in recall between subgroups determined empirically *within* a given experimental group is at least as large as the range of mean recall scores *across* all groups in this experiment. Since all individuals in a given experimental group were treated identically, and since the use of categorized words usually tends to reduce intersubject variability (Marshall, 1967), it may be that the magnitude of individual differences in free recall has been vastly underestimated.

Thus the between-group and within-group analyses of the hierarchical list experiment both demonstrate that the method of proximity analysis can distinguish among relatively subtle differences in organizational structure. It may be concluded that the method is sufficiently sensitive for the types of free recall experiments one might contemplate.

Freedom from Bias

One possible objection to proximity analysis is that the method may be confounding, to an unknown degree, amount of recall with order of recall information. If so, the results described earlier would have considerably less impact. Although the proximity measure was constructed so that it would not be influenced by sheer frequency of recall (i.e., by dividing by $\sum_i \phi_{it}$), better evidence comes from comparisons in which differences between subjects can be attributed to organization, but not recall. To do this, statistical subjects were generated, yoked to real subjects from the hierarchical list experiment under two possible models of random organization. Under an *independent trace* (IT) model, a statistical subject was matched to each real subject only in terms of number of items recalled. The specific items recalled by the statistical subject and their sequential order was chosen at random with uniform probability. According to a *dependent trace* (DT) model, a yoked subject was matched item-for-item to each real subject and only recall order was randomized. Repetitions and intrusions were eliminated from the protocols in both cases.

Essentially, these two models consider the information contained in a real subject's protocol as consisting of three parts: (a) the number of items

recalled; (b) the co-occurrence of pairs of items in recall in terms of conditional probabilities of recall, $P(i|j)$, $i, j = 1, 2, \dots, L$; and (c) the sequential order of recall. Artificial subjects generated under the independent trace model are equated with real subjects in number recalled only. If the proximity method is indeed independent of recall performance per se, no resemblance to the real subject's organization should appear in the IT data. Any differences between real and IT organization should depend only on recall order and the probabilities that some items are recalled, given that other items appear in output.

On the other hand, artificial data generated under the dependent trace model match real subjects in all but the last of the three components. A comparison of the proximity results of real subjects with their DT yoked counterparts should depend only on the order of recall. The notion of "item properties," which Bousfield and Bousfield (1966) felt should be excluded from measures of organization, encompasses both total recall and conditional recall probabilities. Their measures of category clustering (SCR) and subjective organization (ITR) are therefore based upon a comparison of observed values with chance expectation under the dependent trace model.

Finally, the extent to which the mere co-occurrence of particular sets of items in recall influences the proximity results can be judged by comparing the results for IT and DT data, since they differ only in that the conditional probabilities of item recall are included in the latter. The concept of a higher-order memory unit implies that recall of a single item from such a unit should increase the probability that other items from that unit are also recalled. Therefore, the conditional probabilities might be expected to provide some information regarding organization.

Interitem proximities were computed from the protocols of IT and DT statistical subjects in an analysis parallel to that described for Groups B1 to B4. To summarize these results, two measures of organization were derived from the proximity matrices. To the extent that subjects consistently organize groups of items, some proximities will be high and others will be low. Thus the range of proximity values is one indicant of the degree of subjective organization. Also, if subjects organize according to some predetermined set of categories, the average value of proximities for pairs belonging to the same category should exceed the average value for pairs belonging to different categories. The difference between these two average values can be taken as a simple index of categorical organization.

The results in terms of these statistics were quite simple. Artificial data generated under both models displayed no semblance of organization in the proximities among items. Table 6 presents the summary statistics from the analyses carried out for subjects yoked to Groups B1 and B4. The difference of within-category proximities and between-category proximities determined from real data exceeded the corresponding values for

TABLE 6

PARTIAL SUMMARY OF PROXIMITY ANALYSES FOR REAL AND STATISTICAL SUBJECTS

Group	Data	Subjective organization (range of proximities)	Category clustering			Badness of fit to hierarchy (Δ)
			Within	Between	Difference	
B1	Real	13.88	35.74	28.05	7.69	.0140
	IT	3.12	30.12	30.24	-0.12	.0145
	DT	3.99	30.19	30.18	0.01	.0152
B4	Real	16.17	38.30	27.29	11.01	.0112
	IT	3.40	29.32	29.35	-0.03	.0136
	DT	4.60	29.36	29.34	0.02	.0157

both types of statistical subjects by several orders of magnitude. Similarly, the range of proximity scores for real data was about four times that of the statistical data. However, Table 6 also shows small differences between the DT and IT models. The dependent trace subjects matched in terms of the actual items recalled display slightly more organization by these measures than their independent trace counterparts.

It may be concluded that proximity analysis (a) is dependent almost entirely on the order in which items are recalled, (b) is influenced to a slight extent by the conditional probabilities among items in recall, but (c) is virtually independent of the overall level of recall. One further point should be noted regarding the use of the badness-of-fit measure to evaluate cluster analysis results (see Table 6, last column). While the tree structures determined for statistical data were not meaningful in any sense, they did fit a hierarchical clustering scheme as well as the solutions derived from real data. Thus, although a good fit to a hierarchy is a necessary condition for interpreting an organizational structure, it is by no means a sufficient condition for useful results.

Utility of Proximity Analysis

The results above have concentrated mainly on the measurement properties of the proximity analysis. This section attempts to illustrate the utility of the method in a situation where standard methods (e.g., number recalled, or amount of organization) alone provide insufficient evidence for strong conclusions. These results will be described more fully in a forthcoming paper.

The application concerns the effects of organization on part-whole transfer in free recall. In such studies the subject learns a list for several trials and then learns a second list which contains the first-list items. In

MFR, the transfer paradigm provides a means for determining the functional significance in a subsequent task of higher-order memory units which have been developed in prior learning. If memory units are more than a momentary product of recall, the relation between units formed in the two tasks should be an important determinant of performance on the second list. In fact, if one assumes that the number of memory units which can be retrieved on a given trial is limited, then learning the whole list should require the subject to reorganize the part-list memory units in order to accommodate the new items on the whole list (Tulving, 1966). This necessity to reorganize should, according to this view, interfere with or delay learning on the whole list.

This counterintuitive result was demonstrated by Tulving (1966). Two groups learned different initial 18-item lists, and transferred to a common second list of 36 words. The first-list items for a part-whole group (Group PW-R) all reappeared on the second list; a control group (C) learned an entirely different set of items initially. On the second list the part-whole subjects did no better than the controls. In fact, the part-whole group seemed to show a slower rate of whole-list learning.

Tulving's account is quite plausible and predictions from organization theory have been confirmed in a number of other transfer studies (Birnbaum, 1968; Bower & Lesgold, 1969; Novinski, 1969; Ornstein, 1970). Tulving's (1966) data, however, do not compel an explanation based on inappropriate *S*-units. In fact, there are other explanations which are equally compatible with the data.

One possibility is that PW subjects employ an input strategy of selectively attending and rehearsing the new items in List 2 at the expense of old items. In addition, new items tend to be recalled earlier in output than old items (Roberts, 1969). Such a strategy would make new items less susceptible to intratrial forgetting (Tulving, 1964) during the recall period. However, the combined effects of input strategy (selectively attend to new items) and recall strategy (recall new items first) would cause old, previously learned items to undergo interference. Essentially, the recall of old items would be attempted after greater intervening time and interpolated recall. This explanation of negative transfer has also been suggested by Postman (1971) and is supported by the finding that prior part-list learning produces a greater negative effect on the recall of old words than of new words (Bower & Lesgold, 1969).

On the basis of Tulving's transfer studies (Tulving, 1966; Tulving & Osler, 1967), nothing more can be said to decide between these two explanations. However, the organizational account can be tested directly by using the method of proximity analysis to determine the contents of *S*-units at the end of List 1 learning and their composition at various stages in List 2 learning. In addition, it would be useful to compare organizational structure on List 2 for conditions where prior-list organization was

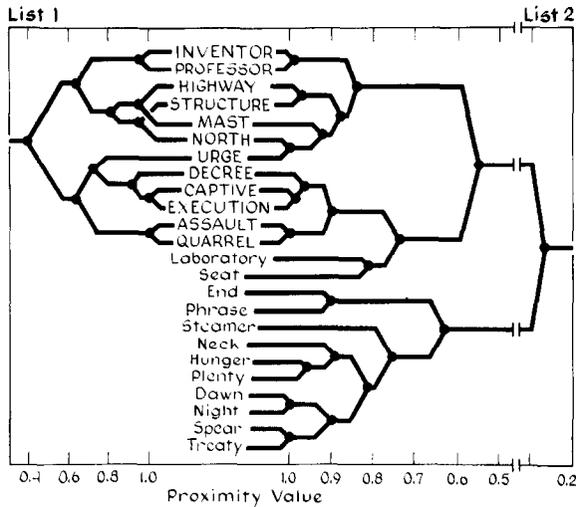


FIG. 18. Comparison of List 1 and List 2 organization for a subject receiving blocked presentation (O/N) of List 2. Data from Ornstein (1970, Expt I).

appropriate for transfer. Several studies by Ornstein (1970) have manipulated prior-list organization in ways which provide data for this test.

One of Ornstein's experiments (1970, Expt I) attempted to maintain prior-list subjective organization by presenting List 2 in blocks of old and new items, in contrast to the Tulving study in which the two sets of words were randomly intermixed on the final list. In addition to groups replicating Tulving's study, Ornstein's design included a group (Group PW-O/N) which received the final list in a blocked fashion. Blocked presentation should serve to facilitate discrimination of old and new subsets and allow subjects to develop a separate parallel organization for the new items, while preserving List 1 groupings of the old words. Transfer was from a 12-word list to a 24-word list, all unrelated words, and eight trials were given on both lists.

As in Tulving's study, Group PW-R, which received the List 2 items randomly arranged, did no better than the control on the final recall task. Group PW-O/N, however, showed large positive transfer.

This result is consistent with the organizational interpretation, but a strong, direct test of this view can be made by analyzing the proximities among items in List 2 recall for subjects in these groups. In order for Tulving's hypothesis to be supported, Group PW-O/N subjects should maintain the organizational pattern developed in List 1, while subjects in the random presentation, part-whole group should show structures for which the organization of old items is fragmented with respect to List 1 organization.

The data first used to illustrate the method of proximity analysis (Fig. 1)

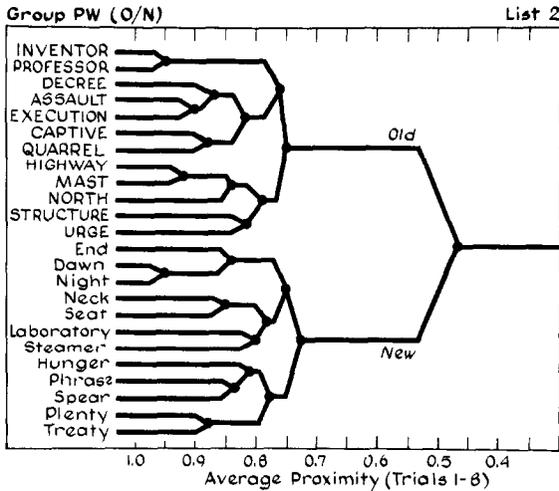


FIG. 19. List 2 organization for Group PW-(O/N).

were taken from List 1 recall protocols of one of the PW-O/N subjects. The cluster analysis performed on the proximities from this person's data (Fig. 3) indicated a hierarchical organization which could be described by three *S*-units. Figure 18 presents the organizational structure for this subject derived from the List 2 protocols (Trials 1–8). The corresponding List 1 M-gram for old words has been redrawn at the left of Fig. 18 for ease of comparison. The most striking feature of the List 2 organization is the separation of the tree structure into “old” and “new” components. The separation is not perfect—*laboratory* and *seat* merge with the old rather than new items—but these two words are only weakly associated with the old items. The groupings of the new items (shown in lower case) seem to make sense semantically: (*end, phrase*), (*hunger, plenty*), (*dawn, night*), and (*spear, treaty*). Also, comparing the organization of the old items with this subject's structure in first-list learning, it can be seen that the major subjective units uncovered earlier have remained intact: (*inventor, professor*), (*highway, structure, mast, north*), (*decree, captive, execution, assault, quarrel*).¹⁰

Although this analysis was in terms of a single subject, the most general results also obtain at the group level, i.e., segregation into old and new components, and maintenance of higher-order units of List 1 items. Figure

¹⁰Perhaps the greatest difference between the two M-grams is in the position of URGE. This word did not appear consistently near any other word during List 1 learning, but is merged with NORTH at the highest proximity level in List 2. The reason for this is not entirely clear, but the proximity of NORTH and URGE may have been underestimated in List 1. Table 1 shows the trial-to-trial proximities of these two items. On trial 3 these items appeared at opposite ends of the protocol ($P = 3$), but on four of the five remaining trials, they were recalled in adjacent or nearly adjacent positions.

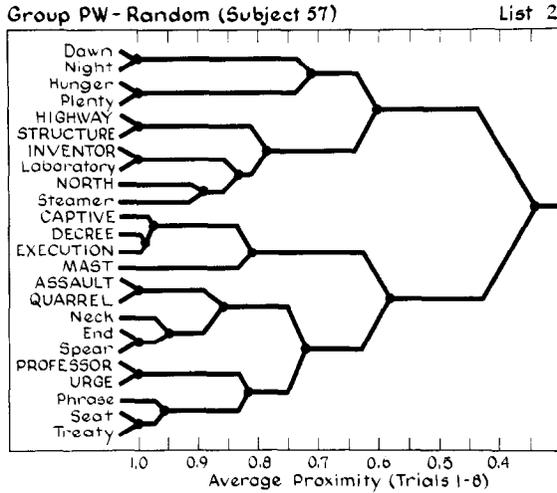


FIG. 20. List 2 organization PW-R for a subject from group PW-R. Data from Ornstein (1970).

19 shows the clustering results for the pooled data of the seven subjects in this group.

With random presentation, Group PW-R showed no positive transfer. The structure of List 2 organization for a fairly typical subject from this group is shown in Fig. 20. The old and new items in the M-gram for this subject are completely mingled. This mixing of items from the two subsets in the organization of List 2 occurred for every subject in Group PW-R.

First-list organization for this subject appears in Figure 21. The groupings differed so markedly that the first and second list hierarchies could not be juxtaposed without considerable crossing of lines. Comparison of the two M-grams indicates that several of the highly proximal pairs of old items do carry over when the whole list is learned: (*decree, execution*), (*highway, structure*), and (*professor, urge*). However, unlike the situation in Group PW-O/N, subjects in the random group seem to have either lost or discarded the higher level *S*-units in whole list learning. Again, this general pattern of intermixing of old and new in List 2 structure, with maintenance of only the strongest primary *S*-units, appears for almost all subjects in this group.

As a result of these analyses, it seems that for both groups, the highly organized, primary *S*-units acquired in learning the part list are maintained and used by the subject in recalling the whole list. What differentiates the groups is the degree to which they use the higher-order units of List 1 to aid recall of List 2. If higher-order units can be thought of as access routes which guide the retrieval system from one primary *S*-unit to the next, then, anything which interferes with these informationally rich units should have a disruptive effect on the overall success of recall. This appears to be

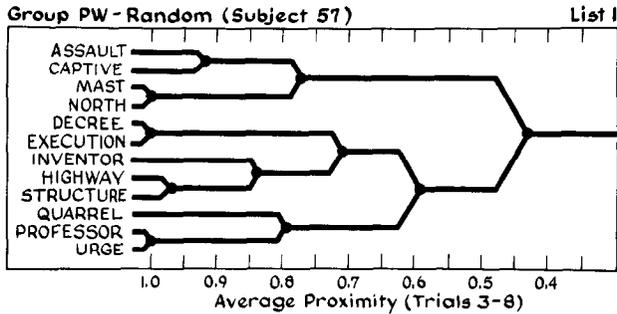


FIG. 21. Organization of "old" words in List 1 learning for the subject whose List 2 M-Gram is shown in Fig. 20. Data from Ornstein (1970).

precisely what has occurred in Group PW-R. Subjects receiving blocked input on the whole list, however, maintain the higher-order units of List 1. For the most part they develop a separate and parallel organization for the new item.

CONCLUDING REMARKS

The last two decades have witnessed successively more detailed attempts to characterize the nature of the information handled by the human memory system. The early steps in this direction occurred with the application of information theory to study of the structure of stimulus and response events (Garner, 1962); and it is of interest that Tulving's (1962a) formulation of the subjective organization measure was based on information theory. While both approaches have led to much useful research, they share a common limitation: Both compress an informationally rich ensemble of responses into a single number. What is gained thereby in precision, is paid for by loss of scope. L. A. Steen, a mathematician concerned with catastrophe theory recently observed that "when the mind must work with numbers, it works linearly, on one track. But if a particular problem can be transformed into a picture, then the mind grasps the problem as a whole and can think creatively about solutions" (*Newsweek*, January 19, 1976, p. 54). I believe his comments apply equally to study of the organization of memory.

The initial success of the information processing approach and the emergence of free recall as a paradigm for the study of mnemonic organization led theorists to seek broader scope. The structuring of information in memory was characterized in such terms as chunks, *S*-units, tree structures, semantic markers, feature lists, semantic networks, and so forth. Most recently, such theories have become greatly more comprehensive. Anderson and Bower (1973), Norman and Rumelhart (1975), and Kintsch (1974) among others have developed detailed specifications for the

internal representation of the knowledge of the world, of language and of experienced events which we store in memory. Obviously, it is important to confront these theories with empirical data. This paper has developed a methodology for extracting structural information from free recall output and fitting these data to three general models into which these theories may be classed.

It should be noted that the techniques described herein apply to situations other than free recall. At one level, the proximity measure may be used in any case where the order of a set of responses is observed over a series of occasions. For example, proximity analysis could be used to study the organization of chess (Chase & Simon, 1973) or Go pieces from data obtained when subjects attempt to reproduce a given board configuration from memory. At a second level, the scaling methods described earlier may be used to extract representations of organizational structure in any situation where measures of interitem similarity can be obtained by other means.

Given proximity analysis, together with analogous procedures for semantic memory tasks (Miller, 1969), a number of interesting questions may be posed and perhaps resolved. Is it reasonable to suppose, for example, that semantic similarity judgments and output order organization in free recall reflect a single underlying semantic configuration? Schwartz and Humphreys (1973) report moderately large correlations between these two types of data which are consistent with this hypothesis. It would be useful, however, to compare the actual organizational structures, and also to determine whether such correspondences hold with other semantic memory tasks. Further, developmental studies have indicated that free recall performance can increase with age without corresponding increases in amount of subjective organization (Ornstein, Hale, & Morgan, Note 8). It is hoped that the approach developed here can suggest explanations of this and other puzzles.

REFERENCES

- Anderson, J. R. FRAN: a simulation model of free recall In G. H. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press, 1972. Vol. 5.
- Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D. C.: V. H. Winston & Sons, 1973.
- Anglin, J. M. *The growth of word meaning*. Cambridge, MA: MIT Press, 1970.
- Birnbaum, I. M. Free-recall learning as a function of prior-list organization. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 1037-1042.
- Bousfield, A. K., & Bousfield, W. A. Measurement of clustering and sequential constancies in repeated free recall. *Psychological Reports*, 1966, 19, 935-942.
- Bousfield, W. A., & Sedgewick, C. H. W. An analysis of sequences of restricted associative responses. *Journal of General Psychology*, 1954, 30, 149-165.
- Bower, G. H. A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*. New York: Academic Press, 1967. Vol. 1.

- Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 323–343.
- Bower, G. H., & Lesgold, A. M. Organization as a determinant of part-to-whole transfer in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 501–506.
- Busacker, R. G., & Saaty, T. L. *Finite graphs and networks*. New York: McGraw–Hill, 1965.
- Caramazza, A., Hersh, H., & Torgerson, W. S. Subjective structures and operations in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 103–117.
- Carroll, J. D., & Chang, J.-J. A general index of nonlinear correlation and its application to the interpretation of multidimensional scaling solutions. *American Psychologist*, 1964, **19**, 540. (Abstract)
- Carroll, J. D., & Chang, J.-J. Analysis of individual differences in multidimensional scaling via an N -way generalization of “Eckart-Young” decomposition. *Psychometrika*, 1970, **35**, 283–319.
- Chase, W. G., & Simon, H. A. Perception in chess. *Cognitive Psychology*, 1973, **4**, 55–81.
- Cofer, C. N. On some factors in the organizational characteristics of free recall. *American Psychologist*, 1965, **20**, 261–272.
- 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 effect of a dual-level stimulus word list on the occurrence of clustering in recall. *Journal of General Psychology*, 1956, **55**, 51–58.
- Cole, A. J., & Wishart, D. An improved algorithm for the Jardine–Sibson method of generating overlapping clusters. *The Computer Journal*, 1970, **13**, 156–163.
- Colle, H. A. The reification of clustering. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 624–633.
- Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 240–247.
- Collins, A. M., & Quillian, M. R. Facilitating retrieval from semantic memory: The effect of repeating part of an inference. In A. F. Sanders (Ed.), *Attention and performance, III*. *Acta Psychologica*, 1970, **33**, 304–314.
- Collins, A. M., & Quillian, M. R. Experiments on semantic memory and language comprehension. In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
- Constantinescu, P. The classification of a set of elements with respect to a set of properties. *The Computer Journal*, 1966, **8**, 352–357.
- Cunningham, K. M., & Ogilvie, J. C. Evaluation of hierarchical grouping techniques: A preliminary study. *The Computer Journal*, 1972, **15**, 209–213.
- Cunningham, J. P., & Shepard, R. N. Monotone mapping of similarities into a general metric space. *Journal of Mathematical Psychology*, 1974, **11**, 335–363.
- Dallett, K. M. Number of categories and category information in free recall. *Journal of Experimental Psychology*, 1964, **68**, 1–12.
- Dalrymple-Alford, E. C. The measurement of clustering in free recall. *Psychological Bulletin*, 1970, **74**, 32–34.
- Degerman, R. L. Multidimensional scaling of complex structure: Mixtures of class and quantitative variation. *Psychometrika*, 1970, **35**, 475–491.
- Ehrlich, S. Le rôle de la structuration dans l'apprentissage verbal. *Psychologies Française*, 1965, **10**, 119–146.
- Erdős, P., & Rényi, A. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 1960, **5**, 17–61.
- Estabrooke, G. J. A mathematical model in graph theory for biological classification. *Journal of Theoretical Biology*, 1966, **12**, 297–310.

- Fillenbaum, S., & Rapoport, A. *Structures in the subjective lexicon*. New York: Academic Press, 1971.
- Frankel, F., & Cole, M. Measures of category clustering. *Psychological Bulletin*, 1971, **76**, 39–44.
- Fridja, N. H. Simulation of long-term memory. *Psychological Bulletin*, 1972, **77**, 1–31.
- Garner, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
- Goodman, L. A., & Kruskal, W. H. Measures of association for cross-classifications. *Journal of the American Statistical Association*, 1954, **49**, 732–764.
- Gower, J. C., & Ross, G. J. S. Minimum spanning trees and single linkage cluster analysis. *Applied statistics*, 1969, **18**, 54–64.
- Gruvaeus, G. T., & Wainer, H. C. Some notes on hierarchical clustering procedures. *British Journal of Mathematical and Statistical Psychology*, 1972, **25**, 200–206.
- Guiliano, V. E. Analog networks for word association. *IEEE Transactions on Military Electronics*, 1963, **MIL-7**, 221–225.
- Harary, F. *Graph theory*. Reading, MA: Addison-Wesley, 1969.
- Harary, F., Norman, R. Z., & Cartwright, D. *Structural models: An introduction to the theory of directed graphs*. New York: Wiley, 1965.
- Hartigan, J. A. Representation of similarity matrices by trees. *Journal of the American Statistical Association*, 1967, **62**, 1140–1158.
- Hollan, J. D. Some effects of epistemological structure on memory. *Memory and Cognition*, 1974, **2**, 670–676.
- Horowitz, L. M., & Prytulak, L. S. Redintegrative memory. *Psychological Review*, 1969, **76**, 519–531.
- Hubert, L. Approximate evaluation techniques for the single-link and complete-link hierarchical clustering procedures. *Journal of the American Statistical Association*, 1974, **69**, 698–704. (a)
- Hubert, L. Some applications of graph theory to clustering. *Psychometrika*, 1974, **39**, 283–309. (b)
- Hudson, R. L., & Dunn, J. E. A major modification of the Bousfield (1966) measure of category clustering. *Behavioral Research Methods and Instrumentation*, 1969, **1**, 110–111.
- Jardine, N., & Sibson, R. The construction of hierarchic and non-hierarchic classifications. *The Computer Journal*, 1968, **11**, 177–184.
- Jardine, N., & Sibson, R. *Mathematical taxonomy*. New York: Wiley, 1971.
- Johnson, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, **32**, 241–254.
- Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Wiley, 1970.
- Kintsch, W. Notes on the structure of semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- Kintsch, W. *The representation of meaning in memory*. Hillsdale, NJ: L. Erlbaum Associates, 1974.
- Kiss, G. R. Networks as models of word storage. In N. L. Collins & D. Mitchie (Eds.), *Machine intelligence, 1*. Edinburgh: Oliver & Boyd, 1967.
- Kiss, G. R. Words, associations, and networks. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 707–713.
- Kiss, G. R. Steps toward a model of word selection. In B. Meltzer & D. Mitchie (Eds.), *Machine intelligence, 4*. Edinburgh: The University Press, 1969.
- Koh, S. D., Vernon, M., & Bailey, W. Free-recall learning of word lists by prelingual deaf subjects. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 542–547.
- Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, **29**, 1–27. (a)
- Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, **29**, 115–129. (b)

- Lance, G. N., & Williams, W. T. A general theory of classificatory sorting strategies I. Hierarchical systems. *Computer Journal*, 1967, **9**, 373–380.
- Ling, R. F. A probability theory of clustering. *Journal of the American Statistical Association*, 1973, **68**, 159–164.
- Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*. New York: Academic Press, 1967. Vol. 1, pp. 327–372.
- Mandler, G. Association and organization: Facts, fancies, and theories. In T. R. Horton & D. L. Dixon (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, NJ: Prentice-Hall, 1968.
- Mandler, G. Words, lists, and categories: An experimental view of organized memory. In J. L. Cowan (Ed.), *Thought and language*. Tuscon: University of Arizona Press, 1970.
- Mandler, G., & Pearlstone, Z. Free and constrained concept learning and subsequent recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 126–131.
- Marshall, C. W. *Applied graph theory*. New York: Wiley, 1971.
- Marshall, G. R. Stimulus characteristics contributing to organization in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 364–374.
- Martin, E. Toward an analysis of subjective phrase structure. *Psychological Bulletin*, 1970, **74**, 153–166.
- McGee, V. M. Multidimensional scaling of N sets of similarity measures: A nonmetric individual differences approach. *Multivariate Behavioral Research*, 1968, **3**, 233–248.
- McQuitty, L. L. Hierarchical linkage analysis for the isolation of types. *Educational and Psychological Measurement*, 1960, **20**, 293–304.
- McQuitty, L. L. Capabilities and improvements of linkage analysis as a clustering method. *Education and Psychological Measurement*, 1964, **24**, 441–456.
- Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81–96.
- Miller, G. A. Psycholinguistic approaches to the study of communication. In D. L. Arm (Ed.), *Journeys in Science*. Albuquerque: University New Mexico Press, 1967. Pp. 22–73.
- Miller, G. A. A psychological method to investigate verbal concepts. *Journal of mathematical psychology*, 1969, **6**, 169–191.
- Miller, J. E., Shepard, R. N., & Chang, J.-J. An analytical approach to the interpretation of multidimensional scaling solutions. *American Psychologist*, 1964, **19**, 579–580. (Abstract).
- Murdock, B. B., Jr., & Okada, R. Interresponse times in single-trial free recall. *Journal of Experimental Psychology*, 1971, **86**, 263–267.
- Norman, D. A., & Rumelhart, D. E. *Explorations in cognition*. San Francisco: W. H. Freeman, 1975.
- Novinski, L. Part-whole and whole-part free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 152–154.
- Ore, O. *Graphs and their uses*. New York: Random House, 1963.
- Ornstein, P. A. The role of prior-list organization in a free-recall transfer task. *Journal of Experimental Psychology*, 1970, **86**, 32–37.
- Patterson, K. E., Meltzer, R. H., & Mandler, G. H. Inter-response times in categorized free recall. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 417–426.
- Pellegrino, J. W. A general measure of organization in free recall for variable unit size and internal sequential consistency. *Behavioral Research Methods and Instrumentation*, 1971, **3**(5), 241–246.
- Pollio, H. R. *The structural basis of word association behavior*. The Hague: Mouton, 1966.
- Pollio, H. R., Kasschau, R. A., & DeNise, H. E. Associative structure and the temporal characteristics of free recall. *Journal of Experimental Psychology*, 1968, **76**, 190–197.
- Postman, L. Organization and interference. *Psychological Review*, 1971, **78**, 290–302.

- Postman, L. A pragmatic view of organization theory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Prim, R. C. Shortest connection networks and some generalizations. *Bell System Technical Journal*, 1957, **36**, 1389–1401.
- Quillian, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing*. Cambridge, MA: MIT Press, 1968.
- Quillian, M. R. The teachable language comprehender. *Communications of the Association for Computing Machinery*, 1969, **12**, 459–475.
- Rapoport, A., Rapoport, A., Livant, W. P., & Boyd, J. A study of lexical graphs. *Foundations of language*, 1966, **2**, 238–376.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 1–20.
- Roberts, W. A. The priority of recall of new items in transfer from part-list learning to whole-list learning. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 645–652.
- Roenker, D. L., Thompson, C. P., & Brown, S. C. Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, 1971, **76**, 45–48.
- Rohlf, J. F. Graphs implied by the Jardine–Sibson overlapping clustering methods, B_k . *Journal of the American Statistical Association*, 1974, **69**, 705–710.
- Rohlf, J. F. A new approach to the computation of the Jardine–Sibson B_k clusters. *The Computer Journal*, 1975, **18**, 164–168.
- Schank, R. C. Identification of conceptualizations underlying natural language. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: W. H. Freeman, 1973.
- Schwartz, R. M., & Humphreys, M. S. Similarity judgments and free recall of unrelated words. *Journal of Experimental Psychology*, 1973, **101**, 10–15.
- Shafto, M. Clique: A Fortran IV program for the Needham–Moody–Hollis Cluster-listing algorithm. *Behavioral Research Methods and Instrumentation*, 1974, **6**, 58–59.
- Shapiro, S. I., & Palermo, D. S. Mediated clustering in free recall. *Journal of Experimental Psychology*, 1967, **75**, 365–371.
- Shepard, R. N. A taxonomy of some principal types of data and multidimensional methods for their analysis. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences*, New York: Seminar Press, 1972. Vol. 1.
- Shepard, R. N. Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 1974, **3**, 373–421.
- Shuell, T. J. Clustering and organization in free recall. *Psychological Bulletin*, 1969, **72**, 353–374.
- Simmons, R. F. Semantic networks: Their computation and use in understanding English sentences. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language*. San Francisco: W. H. Freeman, 1973.
- Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A feature model for semantic decisions. *Psychological Review*, 1974, **81**, 214–241.
- Spence, I., & Ogilvie, J. F. A table of expected stress values for random rankings in non-metric multidimensional scaling. *Multivariate Behavioral Research*, 1973, **8**, 511–517.
- Stenson, H. H., & Knoll, R. L. Goodness of fit for random rankings in Kruskal's nonmetric scaling procedure. *Psychological Bulletin*, 1969, **71**, 122–126.
- Tucker, L. R., & Messick, S. J. An individual differences model for multidimensional scaling. *Psychometrika*, 1963, **28**, 333–367.
- Tulving, E. Subjective organization in free recall of “unrelated” words. *Psychological Review*, 1962, **69**, 344–354. (a)
- Tulving, E. The effect of alphabetical subjective organization on memorizing unrelated words. *Canadian Journal of Psychology*. 1962. **16**. 185–191. (b)

- Tulving, E. Intratrial and intertrial retention: Notes toward a theory of free recall verbal learning. *Psychological Review*, 1964, **71**, 219–237.
- Tulving, E. Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 193–197.
- Tulving, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, NJ: Prentice Hall, 1968.
- Tulving, E. Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- Tulving, E., & Osler, S. Transfer effects in whole/part free-recall learning. *Canadian Journal of Psychology*, 1967, **21**, 253–262.
- Tulving, E., & Patterson, R. D. Functional units and retrieval processes in free recall. *Journal of Experimental Psychology*, 1968, **77**, 239–248.
- Underwood, B. J. Attributes of memory. *Psychological Review*, 1969, **76**, 559–573.
- Voss, J. F. On the relationship of associative and organizational processes. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- Waern, Y. Structure in similarity matrices—a graphic approach. *Scandinavian Journal of Psychology*, 1971, **12**, 1–12.
- Wagenaar, W. A., & Padmos, P. Quantitative interpretation of stress in Kruskal's multidimensional scaling technique. *British Journal of Mathematical and Statistical Psychology*, 1971, **24**, 101–110.
- Ward, J. H., Jr. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 1963, **58**, 236–244.
- Weist, R. M., & Powell, A. Blocked presentation in multitrial free recall. *Journal of Experimental Psychology*, 1972, **93**, 398–403.
- Wood, G. Organizational processes and free recall. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Young, F. W., & Torgerson, W. S. TORSCA, a Fortran IV program for Shepard–Kruskal multidimensional scaling analysis. *Behavioral Science*, 1967, **12**, 498.

REFERENCE NOTES

- Allen, M. Graph theory, organization and memory. In C. N. Cofer (Chair-person), *Organization and memory: The influence of Weston A. Bousfield*. Symposium presented at the meeting of the Eastern Psychological Association, New York, 1971.
- Allen, M. *Graph theory, organization and memory*. Unpublished manuscript, University of Connecticut, 1972.
- Arabie, P., & Shepard, R. N. Representation of similarities as additive combinations of discrete overlapping properties. Presented at the Mathematical Psychology Meeting, Montreal, 1973.
- Friendly, M. L. *Proximity analysis and the structure of organization in free recall*. Princeton, NJ: Educational Testing Service Research Bulletin RB-72-3, 1972.
- Friendly, M. L. Hierarchical organization in free recall and recognition. Paper presented at the meetings of the Eastern Psychological Association, Philadelphia, 1974.
- Johnson, S. C. *A simple cluster statistic*. Unpublished manuscript, Bell Telephone Laboratories, 1968.
- Kruskal, J. B., Young, F. W., & Seery, J. B. *How to use KYST, a very flexible program to do multidimensional scaling and unfolding*. Murray Hill, NJ: Bell Telephone Laboratories, undated.
- Ornstein, P. A., Hale, G. A., & Morgan, J. S. *Developmental differences in recall and output organization*. Princeton, NJ: Educational Testing Service Research Bulletin RB-75-36, 1975.
- Peay, E. R. *An interactive clique detection procedure*. Michigan Mathematical Psychology Program, Report 70-74, Ann Arbor, MI, 1970.

10. Peay, E. R. *Non-metric grouping: Clusters and cliques*. Michigan Mathematical Psychology Program, Report 70-75, Ann Arbor, MI, 1970.
11. Seibel, R. An experimental paradigm for studying the organization and strategies utilized by individual Ss in human learning and an experimental evaluation of it. Paper presented at meetings of The Psychonomic Society. Niagara Falls, Ontario, October 1964.
12. Seibel, R. Organization in human learning: The study sheet paradigm and experiments one and two. Paper presented at meetings of The Psychonomic Society, Chicago, October, 1965.

(Accepted October 26, 1976)