THE ORGANIZATION OF "CHUNKS" IN FREE RECALL

A Thesis

Presented to the

Faculty of Graduate Studies

University of Manitoba

In Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

by

Richard Allen Stojak

February, 1971

ACKNOWLEDGEMENTS

The author is indebted to Dr. J.S. McIntyre and Dr. T. Duffy for their assistance and constructive criticism throughout all phases of this research.

This research was supported by Grant Number APA-317 from the National Research Council.

ABSTRACT

The present investigation was undertaken to test predictions derived from a model of memory proposed by Mandler (1967, 1968, 1970). The model was presented as a hierarchical arrangement of subjectively organized categories of words limited to a value of 5±2 at any given level. Word recall is considered to be a linear function of the number of subjective categories with a strong tendency for words within a category to be generated in clusters at recall. The present study sought primarily to extend the model by testing the hypothesis that "chunks" were subjectively organized into "superchunks" in a process similar to the manner in which words were organized into "chunks". To test this hypothesis, Ss sorted experimenter-defined "chunks" into subjective categories (i.e., superchunks) and were then required to recall the items.

The results did not lend themselves to an unequivocal summary statement regarding their support for the model of organized memory. The relationship between number of categories and chunk recall offered some support to the model with two of the four experimental groups showing the expected relationship. The measures of the organization of chunks showed a distinct tendency for "chunks" within a "superchunk" (or subjective category) to be recalled in clusters beyond that expected by chance. The limiting value of 5±2 categories was demonstrated, however the mean value was slightly higher than other related studies in the area. Consistent with the model of organized memory which views the Ss' degree of organization as the determining variable, there was no correlation between the amount of exposure to the test material and the amount of recalled material.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

In recent years, an increasing amount of research and theoretical speculation has been directed to the free-recall area of verbal learning. Review articles by Shuell (1969), Tulving (1968) and Wallace (1970) have emphasized that many diversified techniques and theoretical models have been incorporated within the free-recall paradigm. One of the major reasons for the increased attention has been the use of the free-recall paradigm to demonstrate the presence of a subject's (S's) pre-experimental cognitive habits - both associative and conceptual - which are activated when the S encounters a verbal learning task. A number of investigators (e.g., Bousfield, 1953; Bousfield and Cohen, 1955; Cofer, 1955; Cohen, 1963a; 1963b; Jenkins and Russell, 1952) have reliably demonstrated that the recall of lists of verbal material organized in terms of conceptual or associative relationships exceeds the recall of control lists which lack these relationships. Further studies (e.g., Mandler, 1967; 1968; Tulving, 1962; Tulving and Patterson, 1968) have shown that subjectively organized verbal material facilitates the recall of the material to a greater extent than the amount of exposure to the material. The basic position taken by these researchers is that organization is a sufficient - and some feel a necessary - condition for memory.

It is the nature of these organizational processes and the concomitant superior recall that serves as the subject matter for this investigation. This thesis will attempt to further test Mandler's (1967) model of permanent or long term memory. Mandler, who holds that organization is a necessary condition for memory, describes his model in terms of

subjectively determined categories arranged in a hierarchical fashion. Each level of the hierarchy is limited to 5 ± 2 categories. The model proposes that a S given a lengthy list of words to remember classified the words into relevant categories, usually three to seven in number. This method of organization has been described as "chunking" or "unitization" after Miller (1956a; 1956b). Mandler (1967) in a series of studies demonstrated that words are subjectively organized into chunks and speculated that these chunks may in turn be organized into higher order chunks or "superchunks" (Miller, 1956b). This investigation will attempt to test Mandler's speculations by answering the question: "Are chunks organized into superchunks in a manner analogous to the process in which words are organized into chunks?"

Before specifying the experimental hypotheses exactly or reviewing the literature relevant to the problem, certain qualifications which limit the scope of the investigation should be made at the outset. These qualifications are made necessary by the abundance of research in the area of free recall.

The first qualification deals with the basic unit of analysis in experiments dealing with the capacity or structure of organized, long-term memory. Typically, the unit used is some operationally defined item such as a digit, letter, word or phrase. Tulving (1968) suggests that the item is such that it can be readily identified in terms of a highly integrated response or, according to Mandler (1967), the item is an organized response sequence that runs off automatically once initiated. The basic units of analysis chosen for this investigation are English words of known familiarity to the S as determined by Thorndike-Lorge (1944) frequency-of-usage

norms. Therefore, the review of the research on organization and memory will be restricted primarily to those studies using words as the test items. Those studies dealing with the organization of objects (e.g., Whitman and Garner, 1962) or with nonsense syllables (e.g., Johnson, 1969; Underwood and Keppel, 1963; Underwood and Erlebacher, 1965; Whitman, 1965) or digits (e.g., Bower and Winzenz, 1970) are considered to be beyond the scope of this investigation.

The second qualification deals with the nature of the organizational processes. In the basic free recall experiment, the S is presented a list of words to study (the input) and is then requested to reproduce as many words as possible in any order he wishes or in the order that they occur to him (the output). Of particular interest from an organizational viewpoint is the appearance of certain regularities in the S's recall protocol. For example, items that are related but not contigous during the study phase tend to be recalled together beyond a level due to chance. Similarly, "unrelated" items may be recalled together over consecutive tests. Bousfield (1953) refers to the former phenomenon as "clustering" and Tulving (1962) labels the latter example "subjective organization". Both researchers propose that the nonrandomness of emission order provides an index of organization as it operates in higher mental processes. Tulving (1968) suggests that organization inferred from the presence of clustering is a function of the Ss prior familiarity with the test items. He calls this "strong" or "secondary organization" and it occurs when "the output order of items is governed by the Ss prior, extra-experimental or intra-experimental acquaintance with the items constituting a list" (p.13). That is, the relationship among the items, as determined by the Ss' past

experience with the items, can effect substantial discrepancies between the input order and the output order of the items. It is these discrepancies which are presumed to reflect the operation of organizational processes.

It should be mentioned that Tulving (1968) distinguishes between primary organization and secondary organization, which is of major concern to this investigation. He defines primary or "weak" organization as the consistent discrepancies between the presentation order and recall order which are independent of the Ss prior familiarity with the test items. An example of primary organization is the serial position curve (e.g., Murdock, 1962) which reflects a strong tendency for items which are presented last in a list to have a higher probability of being recalled and also of being recalled earlier in the output phase than those items in either the initial or medial portion of the test list. Thus, regardless of a Ss prior familiarity with the items, a typical U-shaped function (which relates the serial position of the item to its probability of being recalled) will occur whenever immediate recall is required. However, because of the methodology employed in the present experiment - multiple study trials and delayed recall - the influence of primary organization is not considered to be of critical importance. Postman and Phillips (1965), have demonstrated that a delay of approximately 30 seconds from the end of the presentation phase to the start of the recall phase eliminates the serial-position effect. The decision to ignore the issue of primary organization further restricts the scope of the investigation to the free recall of English words as influenced by the Ss conceptual and associative habits.

There are many modifications to the basic free recall paradigm described above. For example, the presentation made may be aural or visual and correspondingly recall may be oral or written; the study and test trials may be paced or unpaced; there may be only one study and one test trial or varying numbers and combinations of each. Also, the study list may contain explicit experimenter-determined relationships or conversely, the experimenter (E) may actually try to eliminate any relationships among the items. In addition to modifying the basic design, the E is also free to manipulate a number of variables known to influence the degree of organization and recall. A thorough review of these variables is provided by Shuell (1969) and Tulving (1968). Several variables which are especially relevant to this investigation, such as the number of study trials, the number and nature of the conceptual categories, and the presentation order of the list of items, will be examined in detail.

In summary, the qualifications previously described should indicate that this investigation deals with the manner in which a S's past experience with certain words affects the manner in which he organizes those words. With this general overview of the investigation in mind attention can be directed to a more detailed consideration of the problem. The general literature review in the following section is focused towards this end. The review will progress through several stages. First, a definition of organization will be presented, followed by a discussion of two common techniques for measuring the amount of organization. Mandler's (1967) model of memory based on organizational principles will then be elaborated upon and integrated with Miller's (1956a; 1956b) concept of "chunking". Various studies which support a hierarchical model

of memory will be discussed within the context of certain predictions generated by the model. Finally, a consideration of the variables of (a) the number of study trials, (b) the number and nature of conceptual categories, and (c) the method of presentation should logically lead to the purpose of the research and the specific experimental hypotheses.

General Review

Definition of Organization

Mandler (1967) reviewed the Gestaltist origin and somewhat tattered reputation of the concept of "organization" as it applied to human memory. The Gestalt position, epitomized by Katona (1940, as cited in Mandler, 1967) that some type of organization or grouping is an important ingredient for efficient memory, was unproductive for many years. The failure was due to the lack of a specific testable definition of the concept of organization. Recently, more serviceable and testable formulations have been advanced. Mandler (1967) extends Garner's (1962) definition of "structure" to encompass organization as it applied to words in long term or permanent memory. He defined organization in the following manner:

> "A set of objects or events are said to be organized when a consistent relation among the members of the set can be specificed, and, specifically, when membership of the objects or events in subsets (groups, concepts, categories, chunks) is stable and identifiable" (Mandler, 1967, P. 330).

He further qualified this definition by adding that words are organized

> "when the functional aspects of a word, specifically its meaning, depend at least in part upon the set of words of which it is a member, and the relation of the members of the set to each other" (Mandler, 1970, P.102).

The critical words in the definition refer to the stability and

identifiability of the items. Typically, in Mandler's experiments the S was given a number of cards, each with a word printed on it, and asked to sort these cards into conceptual categories of his choice. Stability refers to the Ss ability to sort the cards in the same fashion on two consecutive sorts (i.e., he has organized the items). Identifiability is determined from the Ss recall protocol and is the measure of the tendency for words that were sorted together to be recalled together and is in effect the clustering phenomenon previously mentioned. From the measurement of the characteristics of the recall order it is possible to obtain an index of the amount of organization utilized by the S. Before proceeding with further elaboration of Mandler's position regarding organization and memory, a brief discussion of the quantification of organization is in order.

Measurement of Organization

The pioneering work in this area was done by Bousfield (1953) and was extended by Bousfield and his associates (e.g., Bousfield and Bousfield, 1966,; Bousfield & Cohen, 1956; Bousfield, Cohen & Whitmarsh, 1958; Cohen, 1963a, 1963b, 1966; Cohen and Bousfield, 1956; Puff and Bousfield, 1967). In his original study, Bousfield (1953) assumed that the clustering phenomenon was an index of the organization present in thinking processes and further that clustering could be quantified reliably. His study list included 60 nouns, 15 examples from four different conceptual categories (i.e. animals, names, professions and vegetables). The 60 items were randomized and presented orally at a rate of one item every three seconds. The 125 Ss were then given ten minutes for recall. Bousfield defined clustering as the number of category repetitions during recall. If, for

example, the S recalled "dog, cat, lion" in that order he would be credited with two repetitions from the animal category -- "dog-cat" being one repetition and "cat-lion" being the other repetition (from Adams, 1967). Summing over all categories, Bousfield obtained an index of the amount of organization. He compared the observed clustering to the clustering of a "statistical S" whose clustering was based on chance alone. The real S outperformed the "statistical S" for clusters of size 3, 4, 5, 6 and 7. The real Ss generated 164, 85, 38, 18 and 5 instances of the cluster size indicated whereas the statistical S generated only 87, 18, 4, 1 and 0 instances. In a later report (Cohen, Sakoda, and Bousfield, 1954 as cited by Bousfield et al, 1958) the measurement of clustering or the ratio of repetition (RR_w) was defined as

$$RR_w = \frac{R}{N-1} \qquad [1]$$

where R is the number of repetitions from an experimenter (E) determined category and N-1 is the total of number of words recalled minus one. The reason for using N-1 in the denominator derives from the fact that in no case can the first words in a recall sequence be counted as a repetition. It should also be noted that the maximum amount of organization possible may be less than one as the measure does not take into account the transition from one conceptual category to another which occurs even with perfect clustering at recall. The measure of RR_w is dependent upon the number of conceptual categories and the only situation in which RR_w can equal one is when only a single category is recalled (Shuell, 1969).

There is a basic weakness in Bousfield's approach to measuring organization. The S may discover, but choose not to use the experimenter-determined conceptual relationships. Since the E is forced to ignore any idiosyncratic organization on the part of the S, the amount of organization is usually underestimated. In an attempt to overcome this deficiency Tulving (1962) developed a measure of the subjective organization (SO) each S imposes on a list of items irrespective of the relationships among the items. The S's tendency to recall pairs of words adjacently over successive test trials is known as SO. Derived from information theory, SO is a ratio measure which relates the obtained organization to the maximum possible organization and can vary from zero to one. The limitation of this measure is that it only accounts for clusters of two words but not for clusters greater than two words in size. Again, the amount of organization is underestimated. Since the present investigation does not involve multiple test trials the measure of SO was introduced only to abstract the principle involved that is, the idiosyncratic formation and perseverance of higher order memory units.

Mandler (1967) combined the RR_w measure of Bousfield with Tulving's concept of SO in the following manner. Ss were given unrelated words to sort into their own conceptual categories. Repetitions were then defined as the number of repetitions from the subjectively organized categories. Since the categories were subjective and of unlimited size an accurate measure of the amount of organization was obtained. Although the quantification method is different from Tulving's (1962) the RR_w now measures the degree of subjective organization and not the amount of experimenter-

determined organization as was originally intended.

A Model of Organized Memory

To give a theoretical explanation of clustering, Bousfield originally combined the Hullian concept of 'habit strength' and the Hebbian notion of the superordinate. During recall the S recalls one word, for example, "horse", which in turn activates the superordinate category "animal". The activation of the conceptual name would elicit further instances of the class. The superodinate is arrived at through repeated exposure of the items comprising that unique class. Bousfield further postulates a "relatedness increment" which is an increment of habit strength added not only to the recalled word but also to other words of that conceptual category. He is suggesting, in fact, recall via a mediating concept.

Mandler (1967) rejects this Hullian view of the S as an essentially passive system who learns through the automatic strengthening effects of external reinforcement over a series of trials and views the S instead as an active, organizing agent. Until Mandler, most researchers seemed content to propose a correlational relationship between organization and recall. Mandler (1967) states his position much more emphatically with three general principles concerning the nature of memory:

> "first, memory and organization are not only correlated, but organization is a necessary condition for memory. Second, the organization of, and hence memory for verbal material is hierarchical with words organized in successively higher categories. Third, the storage capacity within any one category or within any level of categories is limited" (P. 328)

Mandler, (1967, 1968, 1970) views the organization of memory in terms of the structure of storage of the items, that is, subjectively determined categories which are hierarchically arranged. Recall, then becomes a search and retrieval process, the success of which depends upon the degree of organization of the items. The evidence Mandler presents to support this model of memory comes from a series of experiments in which the S determines his own categorization system. Either 52 or 100 unrelated words were printed individually on 3X5 index cards. The Ss were given the cards and told to sort them into categories using any rules or strategies they wished. In most of the experiments, Mandler restricted the maximum number of categories (NC) to seven in an attempt to minimize the forgetting of the entire category. However, in his Experiment B the Ss were randomly assigned to a specific NC while in Experiment C the Ss could use a maximum of 20 categories. Finally, in Experiment C the Ss used as many categories as they wished. The important point to note is that the Ss were not told that recall would follow. Instead they were told that the criterion was to sort the cards in the same fashion twice in succession. Over a wide range of experimental conditions, the average NC used by Ss to sort the cards was 4.6. This result has been replicated a number of times (Dong and Kintsch, 1968; Mandler and Pearlstone, 1966; Mandler, Pearlstone and Koopmans, 1969; Mandler and Stephens, 1967). The median correlation between NC and word recall in Mandler's experiments was +.70 ($P < .01$) and ranged from +.39 to +.95. That is, the free recall of the words was a linear function of the number of sorting categories used by the Ss. Conversely, the median correlation between the number of sorts to reach criterion and word recall was not significant (r=+.16).

Hence, it is the amount of subjective organization imposed by the S, and not the mere exposure or repetition of the words, which was the determining factor in free recall. In a follow up study utilizing a factorial design Mandler (1967) manipulated the presence or absence of instructions to categorize or recall the words. The instructions to categorize had essentially the same effect as instructions to recall the words. The group told only to categorize recalled 32.9 words; the group instructed only to recall yielded 32.8 words and the group given both recall and categorization instructions recalled 31.4 words. The final group comprising the design which was instructed to write the words in successive categories recalled 23.5 words. These results support Mandler's initial hypothesis that recall instructions are synonymous with organizing instructions and are treated as such by the Ss.

Another important finding of Mandler's (1967) research was the slope of the line of best fit relating NC to word recall. The average slope was 3.9, that is, for every additional category the S uses, an additional four words are recalled. Further analysis of the recall from categories into which a large number of words were sorted indicated that Ss were subcategorizing within a category. Quantitatively, the large categories contained, on the average two to five subcategories with a recall of about three to five words from each of the subcategories.

The amount of clustering in Mandler's (1967) series of experiments was near the maximum value possible. The average ratio of repition of words (RR_w) was .68 with the maximum possible being approximately .85. Allowing the S to impose his idiosyncratic structure on the material led to efficient organization and recall. These results led Mandler to describe

a model of organized memory in the following way:

> "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 organization. When all the 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 essential growth in the size of the system." (P.366)

The model proposed by Mandler had its roots in research on "chunking" conducted ten years earlier by George Miller.

"Chunking" or the Unitization Hypothesis

One of the more frequently involved explanations to account for the superior recall of subjectively organized verbal material is the chunking hypothesis of Miller (1956a; 1956b). Miller offered a new approach to the understanding of the storage and retrieval of verbal material. His analysis, based on an information processing theory, views the human memory as a communication channel with computer-like inputs and outputs and a restricted capacity for processing and retrieving information. In a review of discrimination and memory tasks he noticed that performance on a memory task was directly related to the number of units to be retained in memory. He analyzed this relationship and suggested that human memory displays two important characteristics: (a) a limitation to 7±2 units (e.g. words) that can be processed at any one time, and (b) the ability to overcome this limitation by a process called unitization (or chunking, or recoding). The unitization process functions in the following way. The S, faced with a list of items that exceeds his immediate memory span, cannot merely add

on seven new items. He must actively reorganize his material into seven new units which integrate the new items in a meaningful way. The original items are "enriched" in the sense that they now contain more information. These new information-rich units are called chunks. Miller does not place limits on the amount of information that each chunk can contain, but some recent evidence (Earhard, 1967) suggests that a cue can function efficiently for only six or seven items. Miller suggests that these first order chunks may themselves be unitized into higher order chunks --"superchunks"--as well. Thus, the S has learned a hierarchy of unitized items and his optimum recall strategy would be to recall the information-rich "superchunks" at the top and work his way down to hierarchy through the "chunks" to the individual items. Johnson (1968) states that Miller "appears to suggest that recall is a reconstructive process in which Ss use rules formed during learning to translate or decode mnemonic devices into specific responses" (p. 436.)

Although the models proposed by Miller and Mandler appear similar, they differ on a fundamental issue, the capacity within any level of categories. Mandler argues that the capacity of 7 ± 2 items which Miller proposes is in fact an inflated value due to a short term memory component inherent in the methodology used by Miller (cf. Waugh and Norman, 1965).

Another superficial difference resides in the terminology employed by both modelers. Mandler refers to "organizing" or "categorizing" while Miller uses "coding, unitizing, or chunking" to describe the differential covert responses that Ss are assumed to make to the verbal material. Both authors are presumably talking about "... the changes, transformations, additions, subtractions, adumbrations, and so on which occur to and between

verbal units as presented and which are, we assume, reflected in what is stored as memory" (Underwood and Erlebacher, 1965, P. 1). It should be emphasized that the present investigation is not interested in the coding or categorizing process per se, but rather the interest is in the end result of this process. As such the broad definition of coding quoted above should impart the sense in which the term is used in this research.

The other frequently involved explanatory concept for the apparent organization of recall protocols is the associative hypothesis. Basically, the associative position holds that items cluster, not because they are related conceptually, but rather, because they tend to elicit each other as free associates. Thus organization would depend upon the direct inter-item association among the items of the list. Research in this framework (Jenkins and Russell, 1952; Jenkins, Mink and Russell, 1958; Deese, 1961) generally utilize lists of items varying in interitem associative strength. Typically, the data indicate that as associative overlap increases, so do organization and recall.

Cofer (1965) showed the relative contributions of both categorization and associative strength to organization. When matched on associative strength, categorized words led to greater organization than non-categorized word pairs. Recently Foote and Pollio (1970) have shown that the extent to which categorization or association contribute to organization depends on a number of variables concerning the experimental task. Under discrete presentation, Ss capitalize on any relationship among the items and high associative strength leads to high recall. When the S is able to view all the items simultaneously and develop idiosyncratic relationships the advantage of high interitem associative strength is minimal.

In this regard Kendler (1966) advocated a dual process theory of organization in which both associative and coding processes are operative. Allen, Puff and Weist (1968) similarly suggest a two level process in which a S (a) attempts to determine and store information about the codability or conceptual relatedness of the items and (b) then forms associations among the items of a given conceptual category.

In review, the models of memory of Miller (1956a; 1956b) and Mandler (1967, 1968, 1970) view organization and hence recall as a process that depends on a S imposing or discovering the rules, relationships or codes which serve to categorize the words for him. At recall the S must first remember his code or category label and then retrieve the appropriate items from within the category. According to this view of memory, the coding and associative hypotheses are not necessarily incompatible. As Foote and Pollio (1970) have shown, their relative effect depends on the task the Ss are faced with. As mentioned earlier, this investigation is not concerned with the process but rather the concern is on the outcome of the organizational process and as such both the associative and coding hypothesis will be involved as the situation demands.

Given Miller's amd Mandler's model of organized memory previously presented the task now is to examine the evidence relevant to the model. The evidence in support is diverse but at times merely suggestive. Fortunately, the model lends itself to certain testable notions. For example, retarding the Ss ability to form subjective units of the words should impair his recall and organization scores. Similarily, if the S categorizes or codes, the words under the category label, then providing this cue at recall should facilitate his recall. Also, if the words are arranged in

a hierarchical fashion, then providing the S with the structure or plan of the hierarchy should facilitate both the learning and recall of the material. The remainder of the general review is directed towards evaluating these general hypotheses which are deduced from the model. Related to the first notion generated from the model, a number of experiments demonstrated that inappropriate SO can hinder learning. Basically the experiments involve transfer tasks in which "old" organized words added to a "new" list retarded the formation of new subjective units. Tulving (1966) had two groups of Ss learn a list of nine words. After 12 learning trials both groups learned a further 18 words. One group was given 18 new words while the other group had nine "new" words plus the nine previously learned words. The results indicated that the group given 18 new words learned the list more rapidly. Tulving argued that the impaired recall of the group with the "old" items was due to the Ss' unwillingness to modify their previously organized subjective units and to form new, relevant subjective units. Bower, Lesgold and Tieman (1969) showed essentially the same decrement in recall brought about by preventing the formation of stable, subjective units. Bower et al, used 6 groups of words with 4 unrelated words per group. The instructions to the Ss required them to relate the four words in an image. One group received the same four-word groupings from trial to trial while the other group had the structure of the word groups changed. The constant group recalled approximately 95% of the words after three trials whereas the group with the random grouping recalled approximately 75% of the words. Thus, in the random group in which Ss were continually forced to break up their subjective groups recall was inhibited.

The next area of support involves the facilitating effect of providing Ss with appropriate cues about the test list. Most of the evidence on the effects of cued recall involves lists of words with E - determined conceptual relationships, however it is assumed that similar processes are operating when the S imposes his own organization. Tulving and Pearlstone (1966) presented their Ss with lists of words belonging to explicitly designated conceptual categories. The lists contained both category titles (e.g., weapons) and category instance (e.g. bomb, knife, gun, etc.). The Ss were informed that they would be required to recall only the category instances. At recall, half the group was provided with cues at recall (CR). These cues were the category titles printed in their recall booklets. The other half of the group (NCR) did not receive any cues to augment their recall. As well as cued or non-cued recall, two other variables were manipulated. Lists varied in terms of length (12, 24 and 48 words) and the number of items per category (one, two and four words). Thus lists involving from three to 48 categories were obtained by combining a list length of 12 words with four items per category (IPC) and a list length of 48 items with one IPC. CR was greater than NCR for all experimental conditions. The facilitating effect varied directly with list length but inversely with the IPC. These results should become clearer if the results of the condition employing 24 words and NC equal to 24, 12 and 6 are considered. If the category title is provided for the group with NC equal to 24 and one item per category almost perfect recall of the category instances would be expected. If cues were not provided many of the categories, and therefore many of the items, would be forgotten. When the NC is reduced to six, there should be little forgetting of the category titles and consequently little forgetting of the category instances. The results of the condition employing 24 words and NC equal to 24,

12 and 6 are quite compatible with the model. The NCR group recalled 11, 12 and 13 words words respectively. The CR group recalled 22, 18 and 15 words respectively.

Dallett (1964) found that recall was facilitated if Ss had before them a list of the categories to aid in identification during presentation and as a memory aid during recall. As would be expected, recall was facilitated for NC equal to six but not for two and four categories, in agreement with Tulving and Pearlstone's (1966) findings. Two and four categories would have little chance of being forgotten and cueing would not facilitate recall. Cueing would facilitate recall for the NC equal to six and this NC is approaching the upper extreme of the model, i.e., 5 ± 2 categories.

Other experiments have shown the facilitating effects of cueing and have attempted to determine the locus of the facilitation. Wood (1967) showed that it was not necessary for the category names to be present during learning, as was the case in the Tulving and Pearlstone (1966) study. Wood provided category labels at recall which were not present during the learning phase. Compared to a group which did not receive cues, recall was facilitated (16 versus 11 words recalled out of 40). Wood suggests that organization functions as a retrieval variable as well as an input variable. He also suggests that the effectiveness of the cue depends upon the pre-established connection between the category name and the category instance. In other words, he interprets the data in terms of the associative strength between the items. Slamecka (1968) found no facilitation of recall by providing contextual cues to his Ss. It should

be noted that Slamecka used only one study trial which is inadequate for the formation of stable subjective units. This lack of facilitation also illustrates the discrepancy between an experimenter - determined cue and a subjectively-determined cue. Simply defining a word as a cue does not necessarily make it so. Wood (1969) suggests that the 30 items in Slamecka's list could easily have been organized and accessible at recall and therefore providing cues would not facilitate the recall. Wood further demonstrated that related words which were presented consecutively can be used as cues to facilitate recall, while unrelated words fail to facilitate recall.

In general, these studies support the notion that the S adopts a plan to organize, encode and store the general structure of the material and when cues are provided at recall the S has only to remember the rules he used to generate the structure. Perhaps the best study to illustrate the effects of cueing the Ss idiosyncratic structure is that by Dong and Kintsch (1968). They used the card-sorting technique of Mandler (1967) and when the Ss had reached the criterion of two identical, consecutive sorts, the authors formed four groups:

Relevant Cues This group of Ss was asked for a one or two-word descriptive title of each of their categories. This group was then given the same relevant cue on their recall sheets.

No Cues This group was asked to provide a title similar to group RC, but the cue was not provided to recall.

No Naming The Ss in this group were not asked to verbalize the category names after sorting the word cards.

Irrelevant Cues This group of Ss did not receive their own subjective category title and instead received irrelevant cues.

Each group sorted three different decks of cards each with 25 words for a total of 75 words. The group which received relevant cues at recall recalled significantly more words (47.3) than the other three groups (35.5, 35.0, 34.0 respectively) which did not differ from each other. This illustrates that subjective cues are an effective means of increasing recall and also suggests that the facilitation effect found by Wood (1967) may have been due to an overlap in experimenter cues and those cues utilized by the Ss.

Another implication of the hierarchical model of memory is that a tremendous facilitation in recall can be achieved by allowing the S to readily perceive the organization of the items. Bower, Clark, Lesgold and Winzenz (1969) constructed four hierarchies of words each with four levels of concepts within the hierarchy. For example, their "Mineral" hierarchy had "Mineral" as a level 1 word; "Metals" and "Stones" as level 2 words; "Rare", "Alloys", "Masonry", etc., as level 3 words; and level 4 words consisted of instances of the level 3 concepts. The four hierarchies contained a total of 112 words. These words were arranged in a hierarchical fashion so that Ss could perceive the relationship among the items, that is, the structure of the hierarchy. A control group received the same 112 words but the words were randomly placed on the nodes of the hierarchy or word tree. After two trials the experimental group recalled 106.1 words out of 112 whereas the control group recalled 38.9 words out of 112.

Further analysis of the data suggests that Ss use the structural

information as a retrieval plan for generating their recall. For example, the conditional probability of recalling a level-4 word was .66 when its level-3 word was recalled, but this dropped to .30 when the level-3 word was not recalled. This structure shows the S where to begin his recall, how to proceed from one category to another, and aids him in monitoring his output for errors. The obvious superiority of the experimental group cannot be denied or attributed to Ss spontaneously generating items.

There is evidence from a number of other sources which supports the hierarchical model of memory but which does not fit nearly under the rubrics of subjective organization, cueing and structure effects discussed thus far. These studies deal with recall strategies, the temporal aspects of recall and further notions about "chunking". For example the best evidence to support the interpretation that the Ss engage in a dual level search process in which the S first searches for the category title and then the category instances is provided by a study of Segal (1969). Each S was presented a random list of either 44 or 55 words at three-second intervals via tape recorder. The word lists contained a latent structure --11 categories with four items per category. The experimental groups had either category names added, category instances added, or nothing added. The paradigm was six alternating study and test trials. The results generally support the notion of a dual level search. Segal records that Ss usually recalled a category name followed by several of its members or several category labels were recalled in succession followed by instances of each of the categories, but in the order in which the categories labels were listed. The experimental group which had the category names included in the study list tended to recall the category

names first. For example, on trial 1, the category names were first on the protocol 56% of the time. By trial 2 this increased to 75% of the time and by the last trial, Ss were recalling the category name first 96% of the time. This evidence further supports the theory that memory is hierarchically organized.

Pollio, Kasschau and DeNise (1968) recorded and measured the oral free recall of their Ss. They then measured the interword-response times. The words comprising a fast output sequence were found to be highly associatively related more so than words of a slow output sequence. There was also a noticeable pause when Ss finished one cluster and before they gave their next fast burst. In a following study Pollio, Richards and Lucas (1969) presented Ss five categories with 5 items per category (from the norms of Cohen et al., 1957). The oral recall of the Ss was recorded and interword-response times were calculated. In spite of intersubject variability, the authors found that long interword-response times occurred between items from different categories and short interword-response times represented words from the same categories. This suggests a search for a category, quickly exhausting the pertinent instances of the category and then a search for the next category, etc. (Kintsch, 1970), which further substamtiates Mandler's and Miller's hierarchical model of memory.

Contemporary research on Miller's (1956a; 1956b) concept of "Chunking" or unitization is also in accord with the model, specifically as "chunking" relates to the capacity of the subjective categories. The major import of the research is that once "chunking" or unitization is performed the memory system then handles the chunk as one unit regardless of the number of items - or in Miller's language "the number of bits" - making up that

chunk. To illustrate this, Bower (1969) reasoned that a well-worn cliche, for example, "Tick-Tack-Toe" should function, not as three words, but as a single chunk. He further reasoned that a chunk should cause as much interference as a single word in the memory system. His test list consisted of 24 items, 12 items designated as "critical" items and the other 12 items were the filler items and of three different types:

One Word This group received 12 single, unrelated words.

Cliches This group received 12, three-word familiar cliches as their filler units.

Three word This group received 12 triplets of unrelated words. The items were then randomized and presented to the Ss.

As predicted critical item recall was the same for the "Single Word" (6.3) and "Cliches" (5.5) groups but significantly poorer for the "Three Word" group (3.7). The identical recall for single words and cliches shows these items are handled in the same fashion in the memory system. The amount of recall of the "Cliche" group is in agreement with Tulving's and Patterson's (1968) conclusion that highly related words are recalled as a functional unit.

Most of the other results on chunking are consistent with the model of memory presented previously. Tulving and Patkau (1962) defined a chunk as a number of words correctly recalled in an unbroken sequence. Using this as their measure, they found that the number of chunks recalled from 24-word lists was essentially constant (between five and six) for different approximations to the structure of English sentences. Tulving and Pearlstone (1966), over a wide range of experimental conditions, found that the number of items recalled per category was constant and approximately three words,

Dong and Kintsch (1968) found the same invariance within subjective categories. Their four groups (described previously) recalled 3.0, 3.3, 2.9 and 3.3 items per category respectively. McNulty (1966) defined a chunk as an unbroken sequence of letters and found that although the number of items recalled increased with their approximation to English, the number of chunks remained constant at approximately seven units.

Cohen (1963a; 1963b) makes a distinction between two types of word categories. A nonexhaustive (NE) category is one in which the instances only partially exhaust the entire pool of words belonging to the category. For example, "Dog, Lion, Horse, and Bear" are only four instances of the category labelled "Animals." These are but a small sample and many other animals could have been included in the list. Conversely, an exhaustive (E) category is one in which the instances completely exhaust the category, for example, the words "North, South, East and West" are the only words belonging to the category titled "Major points of the compass." In his first study on the recall of exhuastive and nonexhaustive categories, Cohen (1963a) presented lists of 70 words belonging to ten exhaustive and ten nonexhaustive categories. Various techniques, such as visually with a projector, orally with a tape recorder, and individually on teaching machines, were used to present the list. Cohen found that the mean number of categories recalled was the same for both types of categories with ten to 14 categories represented in recall. Also, significantly more words were recalled from the exhaustive category than nonexhaustive categories in each of the three studies. In his next study, Cohen (1963b) compared the recall of exhaustive categories to the recall of unrelated words and found that category recall equalled word recall. This

relationship held only if the two lists were matched for total presentation time. Cohen reported a consistent invariance in the recall of words within a category. Providing that at least one instance of the category was recalled, approximately 85% of the exhaustive category would be recalled. Cohen (1966) summarized his data with a "some-or-none"encoding hypothesis. Briefly, this proposes that either Ss (a) do not recall words of a given category (i.e., they forget the category label) or (b) when Ss do recall words of a given category then a constant proportion of the words are recalled. This research on the chunking hypothesis is compatible with the hierarchical model of memory. In the studies cited, both chunk recall and items per chunk (IPC) seem to be limited by the value 5±2. The Ss appear to organize the material within a chunk (i.e., IPC) to a limit of 5±2 and to organize the categories or chunks within the limits of 5±2. This in turn leads to the question posed earlier: "Are the chunks organized into "superchunks" within the limits of 5±2?

In summary, the evidence for a hierarchical model of memory is substantial. The evidence from research on subjective organization, cued recall, structure of material, temporal measurements and chunking is neatly integrated by a hierarchical model of memory. It is granted that most of the studies reviewed thus far have shown a correlation between organization and learning. However, this evidence has a cummulative effect, suggesting that Ss are active, structuring, planning agents and not the traditional, passive S of reinforcement theory. Bower et al. (1969) suggest that the "evidence does not yet prove that such a plan is necessary or required for S to produce high levels of recall, but this hypothesis is beginning to look increasingly attractive" (p. 342.)

Purpose of the Research

The purpose of the present research was to determine if chunks are organized into "superchunks" in a manner analogous to the way in which words are organized into chunks. Specifically, do the limits of 5 ± 2 hold for the organization of chunks into "superchunks"? Are the chunks recalled in an organized fashion, that is, is there a high ratio of repetition of the chunks? Finally, is there a positive correlation between the number of categories used by the S to sort the material and the number of chunks recalled? The model of memory espoused by Mandler demands a "yes" to each of these questions.

There are numerous variables which influence the amount of organization (cf. Bousfield and Bousfield, 1966; Shuell, 1969), however this investigation was mainly concerned with three of them. They are (a) the number of study trials (or reinforcements), (b) the number of categories used to organize the test items and (c) the method of presenting the items (blocked versus random). The findings of the research in these areas will be briefly reviewed. Based on the consideration of the effect of each variable, experimental and control groups of Ss will be described and the exact experimental hypotheses will be formulated.

The Number of Study Trials

The traditional learning position maintains that organization of the list items increases over trials as a direct function of the number of reinforcements or exposures. For example, Bousfield and Cohen (1953) varied the number of list exposures before requiring the Ss to free recall. Using five independent groups, they presented a list of 40 items from one to five times. For the group receiving one presentation, the mean

recall was 23.9 words while the group receiving five presentations recalled 37.9 words. Similarily, the index of clustering for the group receiving five reinforcements was double the amount of clustering for the one reinforcement group (Tulving, 1968).

Another approach to determine the influence of reinforcement is to recognize that different amounts of exposure to the stimulus words have occurred prior to the experimental session. An index of the prior exposure of an item is provided by the Thorndike-Lorge (1944) frequency of usage tables. It is expected, according to the reinforcement principle, that high frequency of usage words (HF) should have higher recall and clustering scores than low frequency words(LF). Bousfield and Cohen (1955) used the Thorndike-Lorge values to construct two lists of 60 items, one an HF list and the other an LF list. The words were presented orally at a word every three seconds and Ss were asked to recall the words in the order that they occurred. For the HF group, the mean recall was 26.6 words while the LF group recalled 22.2 words. The clustering hypothesis similarly received reasonable support. Related to this approach is the collection of taxonomic norms by Bousfield, Cohen, and Whitmarsh (1957, cited in their 1958 study). They had 400 Ss respond with the first four words they associated with each of 43 category names or descriptions. Thus, frequencies of occurrence of the various responses to a category name were tabulated. For their 1958 study Bousfield et al. abstracted from these norms four lists of 40 items each. Each list contained four categories of either high frequency words (HF) or low frequency words (LF). The presentation was visual and in randomized order at a rate of 2.5 seconds per item. The mean recall for the HF group was 24.8 words

and 18.7 for the LF group. Thus, within the framework of traditional learning theory, the concept of reinforcement or frequency of exposure was a vital parameter.

In comparison Mandler (1967) views the recall process as determined by the amount of subjective organization imposed on the material and not by the number of study trials. Recall that the median correlation between number of exposures or reinforcements and recall was only +.16. Thus in the present investigation, regardless of the experimental groups, Mandler's model of memory would lead to a prediction of a non-significant correlation between study trials and recall.

The Number of Categories

At present, the research relating the NC present in the stimulus list to the amount of recall and organization is somewhat equivocal. The majority of the studies have shown that recall and organization improve as NC increases. For example, Mathews (1954) had her Ss classify the names of 24 famous people into either two, three or six categories (e.g., artist, athlete, post, etc.). After ten minutes of interpolated activity, the Ss were required to recall. Mathews found that recall increased directly as a function of the NC the Ss were required to use. For the two, three and six categories, the recall was 11.8 (49%), 13.1 (54%) and 14.4 (60%) respectively, an overall significant difference. Unfortunately, Mathews was not concerned with organization so no data on clustering was available.

Similar results were found by Bousfield and Cohen (1956) in their Experiment 1 with naive Ss. They used 40 words per list with either two, four, or eight categories. Recall was directly related to NC with the Ss recalling 36%, 39% and 44% respectively. The measurement of

organization, the mean ratio of repetition, indicated that clustering similarly increased as NC increased. It is worthwhile to note that in their Experiment II, using sophisticated Ss (i.e. Ss who had previously categorized) recall decreased with the NC. The amount of clustering, however, paralleled the results of Experiment I: i.e., there was significantly more clustering for eight categories than for four categories, than for two categories.

Dallett (1964), in a series of five experiments, provided some conflicting evidence on the relationship between the NC and recall and organization. Of relevance here is Experiment II which utilized a 12 item list and NC equal to one, two, four and six. In Experiment IV, the list length was doubled to 24 items with NC equal to two, four, six, eight and 12. In both these experiments, recall decreased as NC increased. Dallett did not find that clustering scores paralleled recall scores as did Bousfield and Cohen (1956); however, Hudson and Dunn (1969) claim that a statistical bias was operating in Dallett's method of measurement of organization which negated the effect of NC.

In the studies just mentioned, both list length and NC varied across experiments. Items per category (IPC) would also change yet this variable has received little attention. Researchers, other than Tulving and Pearlstone (1966) typically indicate that NC vary but fail to mention that IPC also varies. For a given list length and two different NC any results attributed to differences in NC are confounded with the IPC variable. The hierarchical model of memory predicts that both NC and IPC are limited by the value 5 ± 2 and both variables must be considered when comparing the recall - NC relationship. Earhard (1967) believes that

category labels, (ie., retrieval cues) can only function efficiently for a limited number of category items. She suggested that "an efficient strategy for remembering must be some compromise between the number of cues used and the number of items assigned to each one" (P. 257). As mentioned previously a cue functioned efficiently for five or six category instances.

Further support for the conceptualization of the memory system as being limited by the number of subjective categories and the number of items within such categories comes from Weist (1970). He reasoned that the retrieval of items from memory involved both a category search and then retrieval of items from within the category, with both processes limited to a value of 5±2. He compared the organization and recall of Ss exposed to three lists of 24 words. The lists varied in NC and IPC. There were 4 categories with 6 IPC; 2 categories with 12 IPC; and 12 categories with 2IPC. Weist used a multi-trial free recall paradigm to test performance. The group which received NC equal to 4 showed significantly more clustering and more efficient recall performance than the other two groups which did not differ from each other.

In traditional terms the relation between NC and recall was interpreted in a relatively simplistic manner. Helson and Cover (1956) concluded that as the number of categories increased, the more specific each category label was and hence recall was facilitated. Since the present investigation views the S as an active, planning and organizing agent (after Shepard, 1966) the limiting value of 5±2 categories with 5±2 items within the category appears to be the optimum retrieval strategy and is expected to exert some influence on the Ss' sorting strategies.

The Presentation Mode of Stimuli

"Blocked presentation refers to the experimental situation in which all members of a category are presented contiguously in the stimulus list, for example, all of the examples of one category are presented before those of another category are presented" (Shuell, 1969, p. 363). Most of the early studies on clustering (e.g., Bousfield, 1953; Bousfield and Cohen, 1953; 1956) used a random presentation in which no two members of the same category could appear contiguously in the study list beyond a chance level. At an intuitive level a blocked presentation should allow the S to more easily detect the organization present in the list, or, as Bousfield described it -- to detect the superordinate. The empirical findings generally are in agreement with this hypothesis, although the results are not that large and are on the order of 13% (Bower, Clark, Lesgold, and Winzenz, 1969).

Dallett (1964) in four of his five experiments compared blocked versus random presentation for various list lengths and NCs. An example of the superior performance in recall and clustering (measured as deviations from the expected clustering) was the result of his Experiment V. List length was 24 items with eight categories. Mean recall and clustering for for blocked group were 14.5 words and 6.60 while the random group's performance was 10.5 words and 1.24. Cofer, Bruce and Reicher (1966) also support the previous findings. Comparing HF words (from Cohen, et al., 1957 norms) and LF words they found that the blocked mode facilitated recall and clustering, but to a greater extent in the lists of high taxonomic frequency. Puff (1966) varied the number of contigous category instances in a list before recall. There were 0, nine, 18 or 27 repetitions.

List length was 30 items with three categories. In effect, Puff created a continuum ranging from random (0 repetitions) to blocked (27 repetitions). Recall was 15.2, 15.7, 16.9 and 18.4 words respectively. With one exception, clustering also increased with the number of repetitions.

An unpublished study by Cohen (1968, as cited in Shuell, 1969), compared blocked and random presentation of a 70 word list of 20 categories. Although it is not indicated by Shuell, these categories are presumably the exhaustive categories mentioned earlier in the paper. There was a tendency for the blocked presentation to increase the number of words recalled per category while decreasing the number of categories recalled. Shuell suggests that "blocked presentation may facilitate the coding or organization of the predefined categories while decreasing the likelihood that stable inter-category associations will be developed" (p. 364). D'Agostino (1970) offered support to this position. The Ss in his study were exposed to five items from each of six conceptual categories in either a blocked or random fashion. The results indicated that IPC was higher for the blocked condition (3.31) as compared to the random condition (2.92) and in the direction specified by Cohen. The NC represented in recall were 5.20 for the blocked presentation and 5.57 categories for the random presentation again in the direction predicted by Cohen.

A consideration of these variables led to the selection of the experimental paradigm, test materials, groups of Ss, and ultimately, the experimental hypotheses.

The experimental paradigm used in this experiment was based on Mandler's (1967) card sorting task. Test items were sorted to a criterion

of two identical sorts in succession followed by an unpaced free recall of the items. The "chunks" in this investigation consisted of the exhaustive and nonexhaustive categories devised by Cohen (1963a). As presentation mode was considered a major variable at least two experimental groups were required. The Blocked-Exhaustive Group (BE) received all instances of the chunk on one card whereas the Random-Exhaustive group (RE) had one item per card with the cards presented in a randomized order. The intent was to give group BE one level of organization (i.e., words into chunks) and to determine if they could organize at the next higher level (i.e., chunks into "superchunks"). The BE group could then be compared with the RE group who were required to organize at both words and chunk levels. Organizing at two levels was considered to impose more stress on the memory system than organizing at one level and would be discernable as a difference in the number of sorts to criterion. In addition, the Random-Nonexhaustive (RNE) group was included in an explanatory sense. It was expected that the RNE group would impose still further strain on the memory system and require more sorts to criterion than either groups BE or RE. It was also expected that retrieval of the chunk instances would be more difficult for group RNE and accordingly it would have the lowest IPC recalled.

Two further groups were included mainly as a replication of Mandler's (1967) studies. Group Random Control I (RCI) received 86 unrelated words while group Random Control II (RCII) was required to sort only 25 unrelated words. Group RCII had the same number of words as the experimental groups had categories while group RCI had the same total number of words as the experimental groups. If exhaustive categories act as single units (i.e., chunks) then

the performance of groups BE and RCII should be identical when considering word recall for group RCII and chunk recall for group BE.

Experimental Hypotheses

The research in organization and memory suggested the following hypotheses:

1) The NC for groups RCI, BE and RCII will replicate the mean NC (4.6) found in Mandler's (1967) series of experiments. The random presentation mode and the nonexhaustive list make groups RE and RNE exploratory in nature. However, the NC used by these groups should also approximate 4.6.

2) The number of sorts to criterion will be (from highest to lowest) in this order: RCI, RNE, RE, BE and RCII.

3) Word recall will be (from highest to lowest) in this order: BE, RE, RNE, and RCI.

4) Chunk recall will be equal for groups RNE, RE, BE, and RCII.

5) The amount of organization of both words and chunks will be above the degree of organization attributed to chance alone.

6) The correlations between NC and items recalled (both words and chunks) will be high and positive. There will be a low and nonsignificant correlation between sorts to criterion and the number of items recalled.

7) Items per chunk will be highest for group BE, then group RE and the lowest IPC for group RNE.

METHOD

SUBJECTS

Originally, it was intended to use only paid Ss for the experiment. Posters, advertising the payment of $1.50 for approximately 1 hour in a psychology experiment, were distributed strategically throughout the University. An advertisement was also placed in the campus newspaper. The Ss were male and female students, between the ages of 18 and 25 years and had English as their native language. Thirty-two Ss participated, six criterion Ss in each of five groups. However, due to a lack of volunteers it was decided to use the subject pool from the introductory psychology courses. 72 Ss participated in this phase making a total of 20 criterion Ss in each group.

It should be noted that at the conclusion of the main experiment, a further 20 Ss were used to test for an unexpected strategy used by the Ss of group BE. Reasons for this will be elaborated below.

MATERIALS

Five sets of words were used. These words, along with their Thorndike-Lorge (1944) frequency of usage are contained in Appendix A to this report. All words were typed in upper case letters either individually or as categories on 3 inch by 5 inch blank, index cards. Each set is described in detail below.

The Blocked-Exhaustive List. This list, presented to the BE group consisted mainly of exhaustive categories of words described by Cohen (1963a; 1963b). Six additional categories, judged to be nearly exhaustive were constructed to bring the total to 25 categories. There were 14 three

word categories and 11 four word categories for a total of 86 words. Thus a card would contain either 3 words (e.g., blonde, brunette, redhead) or 4 words (e.g., north, south, east, west) in a columnar arrangement.

The Random-Control II List. This list consisted of 25 words chosen randomly, one from each of the 25 exhaustive categories described above. The words were centered, one per index card.

The Random-Exhuastive List. This list consisted of the same 86 words that were presented on the BE list. However, in this instance there was only one word per card. The cards were randomized so that no two category instances were consecutive.

The Random-Nonexhaustive List. This list consisted mainly of the nonexhaustive categories of words described in Cohen's (1962a; 1963b) articles. Cohen's categories were supplemented with six additional non-exhaustive categories chosen from the norms of Battig and Montague (1969). There was one word per index card and the cards were randomized so that no two category instances were consecutive.

The Random Control I List. This list, of 86 unrelated words, was chosen from a number of sources. Fifty words were chosen at random, one from each of the exhaustive and nonexhaustive categories described above. A further 20 words were obtained from Cohen's (1963b) list of unrelated control words. The remaining 16 words were chosen from the list in Mandler's (1967) article.

An attempt was made to equate all the lists in terms of Thorndike-Lorge frequencies. The mean frequencies of the exhaustive, nonexhaustive, random control I and random control II lists were, respectively 53.8,

51.3, 51.2 and 51.0. (AA words were assigned a value of 100 and A words were assigned a value 50.) A number of decks of each list were prepared to ensure that each S would have sufficient prearranged decks to sort to the criterion.

PROCEDURE

The Ss were randomly assigned to the treatment conditions and to one of two identical, experimental rooms. Each S was run individually by one of two E's. Upon entering, the S was seated at a table. In front of the S was one of the decks of words. Seven sorting categories were outlined on the table in masking tape. Each sorting category was slightly larger than a 3 X 5 inch index and they were numbered one to seven from the S's left to his right. A wooden divider partially screaned the E's table and the recording materials from the S.

The exact instructions to each S are contained in Appendix B. Basically each S was instructed to sort the cards into from 2 - 7 categories on any basis, by a system or set of rules with one exception. To minimize the incidence of non-content categorization the Ss were requested not to use the alphabet as the basis for categorization and that recall would follow. When they finished sorting one deck of cards they were given another deck of the same words but in a different order. The Ss were told of the criterion of two identical sorts in succession and when they reached this criterion, recall was requested. They were then handed a sheet of paper and asked to recall as many words as they could in any order they wished. They were asked to write the words in succession on blank lines numbered one to 90. Following recall, the E performed an informal quiz as to their system of categorization.

RESULTS AND DISCUSSION

The raw data from the experiment are contained in Appendix C where the appropriate groups are compared on each of the seven dependent variables.

The data from four Ss were discarded from the experiment. A paid Ss in group RCII counted the number of letters per word and sorted on this basis. Recall was obviously at a minimum. Another paid S in group RNE failed to reach criterion within one hour. The data from two unpaid Ss was also discarded. One S in group RE failed to reach criterion and a S in group BE failed to follow recall instructions. Instead of writing the words in succession at recall, the latter S placed words from one chunk adjacent to each other regardless of when she retrieved them. The instructions which informed Ss that (A) they would later be required to recall the words and (B) not to use the alphabet as their basis of sorting did much to reduce the 47% S attrition rate found in Mandler's (1967) study.

The paid and unpaid Ss were first compared on measures of number of sorts, NC and word recall. The analyses are shown in Appendix D. There were no significant differences between paid and unpaid Ss with respect to number of sorts, $F(1.90) < 1.0$; NC, $F(1.90) < 1.0$; and word recall, $F(1.72) = 2.09$. All interactions between pay/no pay and the three variables just mentioned similarly lacked significance. Since pay/no pay did not appear to be a major consideration it was decided to combine the data and perform the planned analyses.

Each S, depending on his experimental condition was scored on a number of variables. All the Ss were scored on the number of sorts to

criterion, the number of categories used, total words recalled and the ratio of repitition of the words. Originally, the number of relevant and irrelevant intrusions on the recall protocols was considered to be of interest, however, there were very few instances of these phenomena and no measures were taken. Groups BE, RE amd RNE were also scored on the ratio of repitition of chunks (RR_c) and items per chunk recalled (IPC). As these two measures were used in a somewhat novel sense further amplification, through the use of an example, is in order.

Ratio of Repetition of Chunks (RR_c). This measure is analogous to that developed by Bousfield and Bousfield (1966). It is defined as

$$RR_c = \frac{R_c}{Nc-1} \qquad [2]$$

where Rc is the number of times a chunk from a category follows another chunk from that category and Nc is the total number of chunks recalled. Following Cohen (1962a; 1963b) a chunk was considered to be recalled if at least one of the numbers of the chunk were recalled. Each chunk was counted only once, that is, if two members of the chunk were separated by a few other words or chunks, only the first member was counted as a chunk recalled. The second member was however, counted as a word recalled. This method of measuring organization may be best shown by the use of an example. Consider a S in group BE, who sorted and recalled five chunks in the following way:

Sorting		Recall	Category	Word Repitions	Chunk Repitions
Category 1	Category 2	Bass	1		
Bass	Army	Tenor	1	1	
Tenor	Navy	Alto	1	2	
Alto	Airforce	Masculine	1	3	1
Soprano		Feminine	1	4	
	Federal	Mother	1	5	2
Masculine	Provincial	Father	1	6	
Feminine	Municipal	Sister	1	7	
Neuter		Soprano	1	8	
		Army	2		
Mother		Federal	2	9	3
Father		Navy	2	10	
Sister		Airforce	2	11	
Brother		Provincial	2	12	
		Neuter	1		
		Brother	1	13	
		Municipal	2		

As indicated above, the S's recall protocol shows that 17 words and five chunks were recalled. The number of chunk repetitions (Rc) was three. Notice also that only "Army" followed by "Federal" was counted as a chunk repetition but that "Airforce" followed by "Provincial" was not. As Segal (1969) showed, some Ss prefer to recall all the category labels and then recall all the instances of each category. Thus to prevent an overestimate of the amount of organization it was decided to count only the first instances of the chunks in the chunk repetition score. "Airforce" followed by "Provincial"was, however, counted as a word repetition.

Substitutingin formula [2]

$$RRc = \frac{3}{5-1} = .75$$

For each S the obtained ratio of repetition of chunks is compared to the maximum and minimum values of RRc. In our example, the maximum number of chunk repetitions was also three and so the maximum amount of chunk organization was achieved by the S. It now remains to compare the obtained

organization to that organization which could have occurred by chance alone. The general formula for determining the number of repetitions assuming that the sequence is generated at random is that of Bousfield and Bousfield (1966):

$$E(Rc) = \frac{Mi^2}{N} - 1 \qquad [3]$$

$$= (M_1^2 + M_2^2 + \ldots + M_k^2)/N - 1$$

where Mi are the number of items (in our case chunks) recalled from category i as i takes on values from one to seven and N is the total number of chunks recalled. Recalling that three chunks were recalled from the first category and two chunks were recalled from the second category, E(Rc) = 1.6. Substituting this value back into formula [2] the degree of organization expected by chance alone is equal to .40. This process was repeated for each S of groups BE, RE and RNE. Finally, for each number of categories used by the Ss, the values of maximum, observed, and random organization were averaged across Ss.

Items per Chunk (IPC). This measure is simply the total number of words recalled divided by the number of chunks recalled. In the example above IPC is equal to 17/5 or 3.40 words per chunk recalled.

Before detailing the results of the experiment it is necessary to specify the reasons for the inclusion of an extra group (BE_2) which was tested after the main experiment. During the informal quiz of BE_1 Ss at the completion of their experimental session, it became obvious that the subjects were not attending to all the words comprising a chunk. They tended instead to pick out only one or two words and sort on this basis.

A similar phenomenon has been reviewed by Underwood (1963) as it applied to paired-associates learning. In general, the studies reviewed by Underwood show that Ss concentrate on some aspect of the stimulus rather than on the whole stimulus. In other words, the functional stimulus attended to by the S is different from the nominal stimulus defined by the E. A combination of the stimulus selection admitted to by Ss and the results of a pilot study in which group BE, performed at the same level as group RCII led to the decision to test a further 20 Ss. Instead of having all the words of a chunk on the same card, group BE_2 received the same chunks but each member of the chunk was on a separate card. The cards comprising a chunk were in a blocked or sequential order. It was hypothesized that the Ss of group BE_2 would attend more closely to each item and thus improve their performance relative to group BE_1. The data from group BE_2 were not included in any of the statistical analyses of variance. The possibility of violating the assumptions of random sampling led to this decision. However, the data are included in Tables 1 and 2 and Figure 5 mainly to illustrate the effects of controlling for what was assumed to be stimulus selection.

The results of the experiment are presented and discussed under seven major headings: (a) number of categories; (b) number of sorts to criterion; (c) word recall; (d) chunk recall; (e) organization of words and chunks; (f) interrelations of variables; and (g) the number of items per chunk. Where possible an attempt will be made to integrate the data with previous findings in the area as related to a hierarchical model of memory. This will obviously be speculative in nature; however, the data will be discussed on its own merits.

Number of Categories

The mean number of sorting categories (NC) used by each group is shown in Table 1. The analysis of variance yeilded significant differences among the groups (F (4,95) =6.71, p <.01). A Newman-Keuls test (see appendix D_4) indicated that group BE, used fewer categories than the remainder of the groups. Other than group BE, the NC used by the remaining groups appear to be higher than other studies in the area. For example, the mean NC for the seven experiments described in Mandler (1967) is 4.6. Other studies such as Mandler and Pearlsone (1966), Mandler and Stephens (1967) and Dong and Kintsch (1968) found mean NC to be 4.3, 5.0, and 4.9 respectively. It is interesting to speculate on some of the reasons for this discrepancy, if in fact a discrepancy does exist. Neither list length nor Thorndike-Lorge (1944) frequency of usage appear to be contributing factors. In the studies cited above list lengths ranged widely and included 15, 25, 52 and 100 words. Frequency of usage varied widely from a count of 11 words per million to AA, both within and between experiments. In all instances NC was relatively stable. A possible explanation for the discrepancy might be the degree of E imposed relationships among the words in the RNE, RE and RCI lists. In the studies by Mandler and his associates and Dong and Kintsch (1968) attempts were made to select words with no obvious relationships and to allow the S to impose his own categorization scheme. In the present study, groups RE and RNE were given lists with some obvious relationships. Their strategy may have been to use as many sorting categories as possible with the hope that they would exhaust the obvious relationships among the words in the list. In other words, the Ss' strategy may have been "I think there are only seven conceptual categories in the items". Once they

decided on a sorting strategy it was difficult to change as had been substantiated by Earhard (1969). Additional support for this interpretation comes from a pilot study conducted prior to this investigation. In the pilot study Ss were asked to sort 40 words comprising the four categories of animals, vegetables, names and professions used by Bousfield, Cohen and Whitmarsh (1958). There was a tendency though not significant, to sort into more than the four obvious categories. A similar interpretation may be applied to group RCI. The 85 words of the RCI list may have contained some unintentional but obvious relationships and the reader is invited to scan the random list contained in Appendix A and decide on the feasibility of this contention. A corollary of this hypothesis would predict that is there are obvious relationships among the words it should be easier to form a stable organization, that is, there should be fewer sorts to criterion. This is emphasized in the following section on the number of sorts to criterion.

Another way to approach this issue is to ask why the Ss did not use all seven sorting categories instead of asking why they did not use five categories. A conclusion from this line of reasoning is that a compromise was achieved between the tendency to sort into the optimal five categories and the strategy to use more sorting categories to determine the extent of the relationships among the items. As mentioned, once, a S embarks on a particular sorting strategy it is difficult to alter it.

It is conceivable that frequency of usage, list length and E defined organization interact is some complex fashion. The design used in this experiment is obviously inadequate to deal with the problem of interaction although further studies in this area could utilize factorial designs.

Table 1

Mean Scores on the Dependent Variables

as a Function of the Experimental Treatments

	Group					
	RCI	RNE	RE	BE_1	BE_2	RCII
Number of categories (NC)	6.3	6.1	6.5	4.5	5.6	6.1
Number of sorts to criterion	4.2	3.7	4.0	4.3	4.1	4.1
Word recall	55.7	49.9	60.8	52.2	61.4	----
Chunk recall	----	20.3	21.9	17.1	19.6	22.2
Ratio of repetition of words (RR_w)	.68	.75	.79	.82	.82	.67
Ratio of repetition of chunks (RR_c)	----	.56	.56	.45	.53	.67
Items per chunk (IPC)	----	2.45	2.79	3.06	3.13	----

Note: The data from group BE_2 was not included in any of the statistical analysis. RR_w and RR_c for group RCII was not included in the analysis of variance. Rationale for these decisions is given in the text.

Number of Sorts to Criterion

The mean number of sorts to criterion for each group is shown in Table 1. Analysis of variance showed there were no differences among the groups ($F(4,95)=0.78$). These results may be compared to those of Mandler (1967) who found a mean of 6.2 trials to criterion for list lengths of 52 and 100 words and to those of Dong and Kintsch (1968) who found a mean of approximately four trials for a 25 word list. As in the previous section, it is difficult to determine if in fact there exists a true difference. It suggests, however, that the 86 words of group RCI were organized easier than the 52 words used by Mandler (1967, Experiments A to E). The notion that the RCI list contained obvious relationships which facilitated the organization of the list appears to be supported. Group RCI required approximately the same number of trials (4.2) to organize 86 words as group RCII (4.1) required to sort 25 words.

The sorts to criterion for groups RNE, RE and Be (3.7, 4.0, and 4.3) were in the opposite direction, although not significantly so, than that predicted in the experimental hypothesis. A possible explanation of this would be to interpret the sorting task for each group as different levels of an interference task. Group BE which had one level of organization provided for it may have found the E - determined categorization incompatible with their subjective categorization system. Group RE would find less interference as they were able to organize at both the word and chunk levels. Finally, group RNE with the possibility of many more category instances than group RE may have been less constrained by the E - determined relationships. However, as the data stands (i.e., with equal sorts to criterion for groups RNE, RE and BE) a tentative conclusion might be that it is the number of conceptual categories (or chunks) and

not the presentation mode nor the exhaustiveness of the categories which determines the organizability of the material.

Word Recall

Table 1 shows the mean number of words recalled by each of the groups. Analysis of variance indicates a difference in recall ($F(3,76)=2.97$, $p<.05$) and the post hoc test (Appendix D_5) indicates that the overall significance is due to a difference between groups RE and RNE. The difference (60.8 versus 49.9 words) indicating a greater recall for group RE was in the direction specified in the experimental hypothesis. The word recall for group RCI was surprisingly high. The Ss recalled, on the average, 65% of the words. In Mandler's (1967) study, the Ss recalled an average of 50% of the 52 word list and only 40% of the 100 word list. In the experiment by Dong and Kintsch (1968) the Ss in the non-cued recall group recalled approximately 35 out of 75 words or 47%. These results may also be compared to those of Bousfield et al, (1958) who found that Ss recall approximately 24.8 words (62%) of a high frequency list of 40 words -- 10 words in each of 4 conceptual categories. The contention that the RCI list contained an inadvertent amount of prior organization which facilitated recall is again supported.

The lack of difference between the RE and BE_2 groups (60.8 and 61.4 words recalled) suggests that both E - determined and S - determined relationships facilitate recall once the S has reached the sorting criterion. The RE group was forced to organize at both the word and chunk level and this proved to be as advantageous as allowing the BE group to perceive the organization at the word level. The difference between groups RCI and RNE (55.7 versus 49.7) although not significant was in a direction opposite to that predicted in the experimental hypothesis. This suggests

that providing a S with weaker relationships, as in the RNE group, is offset by the subjective efforts carried out by group RCI Ss to organize the material.

Chunk Recall

Table 1 shows the mean number of chunks recalled for groups RNE, RE, BE and RCII. Following Cohen (1963a) a chunk is recalled if at least one member of the chunk is recalled, i.e., at least one instance of the exhaustive or nonexhaustive category is recalled. The analysis of variance yielded a significant difference among the groups ($F(3,76) = 11.71$, $p < .01$) and the post hoc test (Appendix D_6) showed that group BE_1 recalled and least number of chunks and that groups RNE, RE, and RCII did not differ from each other. The difference between groups RCII and BE was not anticipated, for in the pilot investigation previously mentioned, these groups performed at the same level. In the pilot study, the blocked presentation consisted of one word per card with the chunk instances contiguous. This method was employed again with group BE_2 and chunk recall increased to 19.6. As the data from group BE_2 were not included in the analysis of variance it is not worthwhile to pursue these findings in any specificity. It should be noted, however, that the significant differences are in effect artifactual and the low chunk recall of group BE_2 is due to the stimulus sampling effect previously discussed. It is quite possible that the same results would have been found if the original group BE, were required to say aloud each of the members of the chunk.

The experimental hypothesis predicted an equal recall from groups

BE, and RCII and the discussion will be restricted to reasons why this did not obtain. The first possibility was mentioned previously and involved the notion of stimulus selection by the Ss. The Ss admitted that they were cueing on one or two of the chunk instances rather than the entire chunk. The inferior recall of group BE_1 appears to be an example of the "associative priming effect" or relatedness increment described by Bousfield (1953) and Puff (1966) to account for the finding that the more repetitions of category instances then the greater the probability of category recall and the greater the amount of organization. The priming effect might account for inferior chunk recall of group BE, but cannot handle the findings that the organization of BE_1 was near maximum at several values of NC and that the items recalled per chunk was the highest of the groups. These latter points will be presented in detail in the sections on organization and IPC.

Cohen (1963b) showed that the number of chunks recalled from lists of 20 exhaustive or 20 nonexhaustive categories was the same as the recall of 20 unrelated words but only when the lists were equated for total presentation time. If the lists were not equated (i.e., a constant rate of 3 sec. per word for each list) then chunk recall surpassed word recall. Unfortunately in the present study, time to sort was neither controlled nor measured as a dependent variable. In defence of this omission, Mandler (1967) found no correlation between time to sort and word recall and perhaps the total time hypothesis is not relevant for this task. However, future investigations in this area should measure sorting time, if not to defend against this criticism, then at least to replicate Mandler's finding.

The Organization of Chunks and Words

It should be stated at the outset that the values of the mean ratio of repetition of words (RRw) shown in Table 1 are, at the best, gross indices of the organization present. Recent articles Dalrymple-Alford, 1970; Hudson & Dunn, 1969; Shuell, 1969) have described the inadequacy of comparing clustering measures across groups which used different NC and had different recall totals. For this reason the organization measures of group RCII, which recalled a mean of 22.2 words, were not compared with the remaining groups which had approximately double the recall. Although the criticism may not be as serious for the range of NC and items recalled of groups RCI, RNE and RE and BE, a conservative approach will be presented. The observed organization for each group was compared with the maximum and minimum amounts of organization possible. Figures 1 through 7 show the organization as a function of NC.

Figure 1 illustrates the obtained RR_w of group RCI at the three levels of NC. The observed organization is contrasted with both the organization present if the Ss had recalled the words from each category in one cluster and with a random amount of organization. Figure 1 essentially replicates the finding of Mandler (1967) which showed that as NC increases, Ss diverge more from a random model and approach the maximum organization possible. The clustering data showed that the subjective organization of the words produced a tendency for items of the same category to be recalled in a cluster. With smaller NC the Ss had a tendency during recall to switch from one category to another as compared to large NC where categories are exhausted in a more consistent fashion (Mandler, 1967).

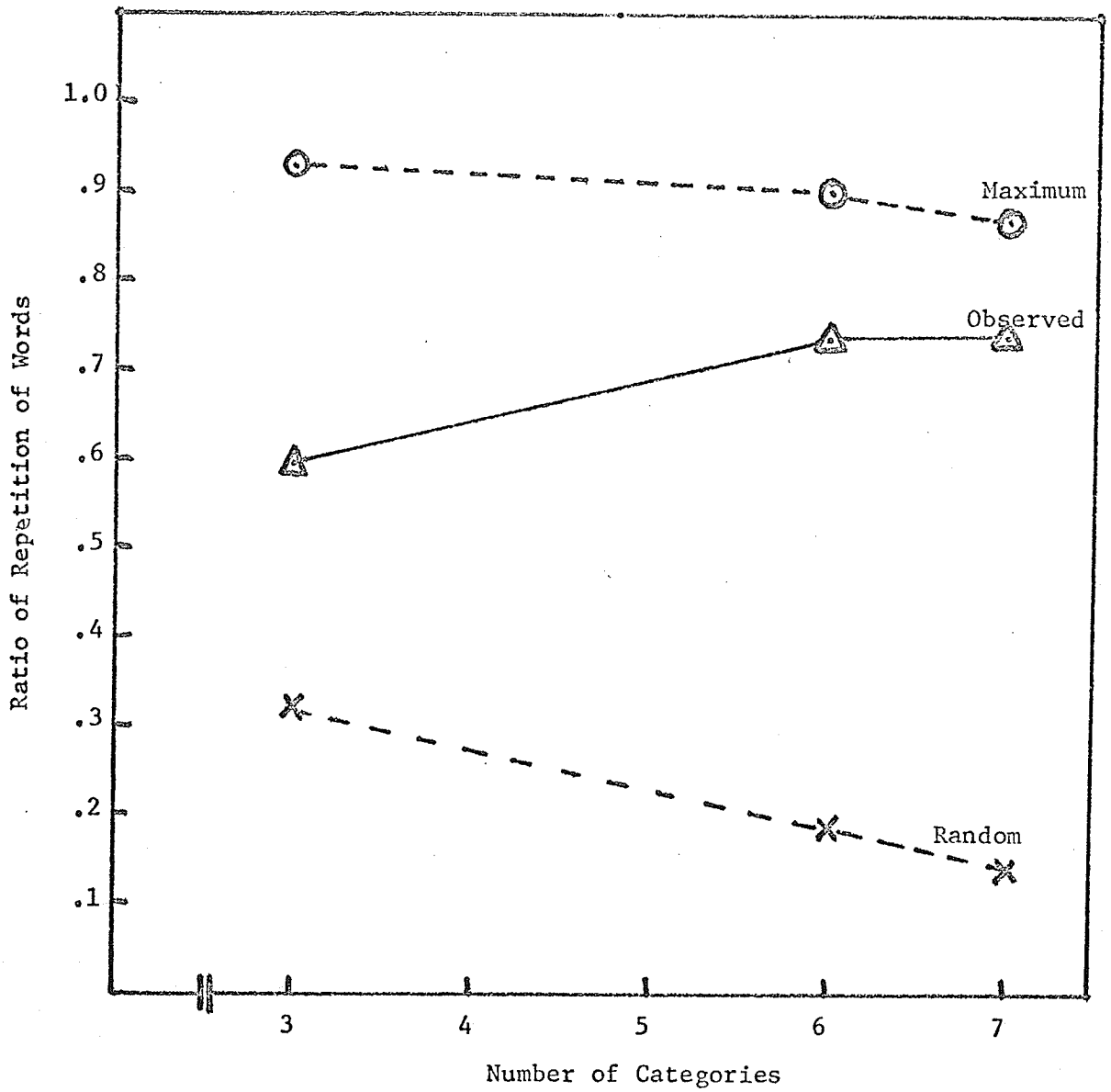

Fig. 1. Mean ratio of repetition of words as a function of number of categories used by group RCI. Solid line shows observed data; dashed line with circles shows maximum possible if perfect clustering had occurred; dashed line with crosses shows mean ratio of repetition for a randomly generated output.

The organization of chunks (RRc) of groups RNE and RE is shown in Figures 2 and 3 respectively. Again it can be seen that there was a strong tendency for chunks of the same category (i.e., the same superchunk) to be recalled together in a cluster. The organization was near maximum for both groups.

Figure 4 illustrates the amount of organization at each level of NC for group BE_1. As shown, Ss who used a low NC (two or three) tended to generate clusters at or below the random rate. As NC increased the organization also increased to near maximum. The organization of group BE_2 (Figure 5) depicts an opposite trend. Organization was maximum for a low NC and as NC increased the degree of organization departed from the maximum value but was still well above the random amount. This discrepancy is difficult to account for. The near random organization at low NC values of group BE_1 appears to be a result of inadequate "priming" as described earlier. If this group did select only one or two words of the chunk instead of three or four words then the "priming" theory would predict that the formation of a stable category did not occur and consequently recall and organization would be reduced. Yet as the NC increases organization also increases. Is it therefore conceivable that the stimulus selectors were the same Ss who sorted into only two or three categories? This problem awaits resolution.

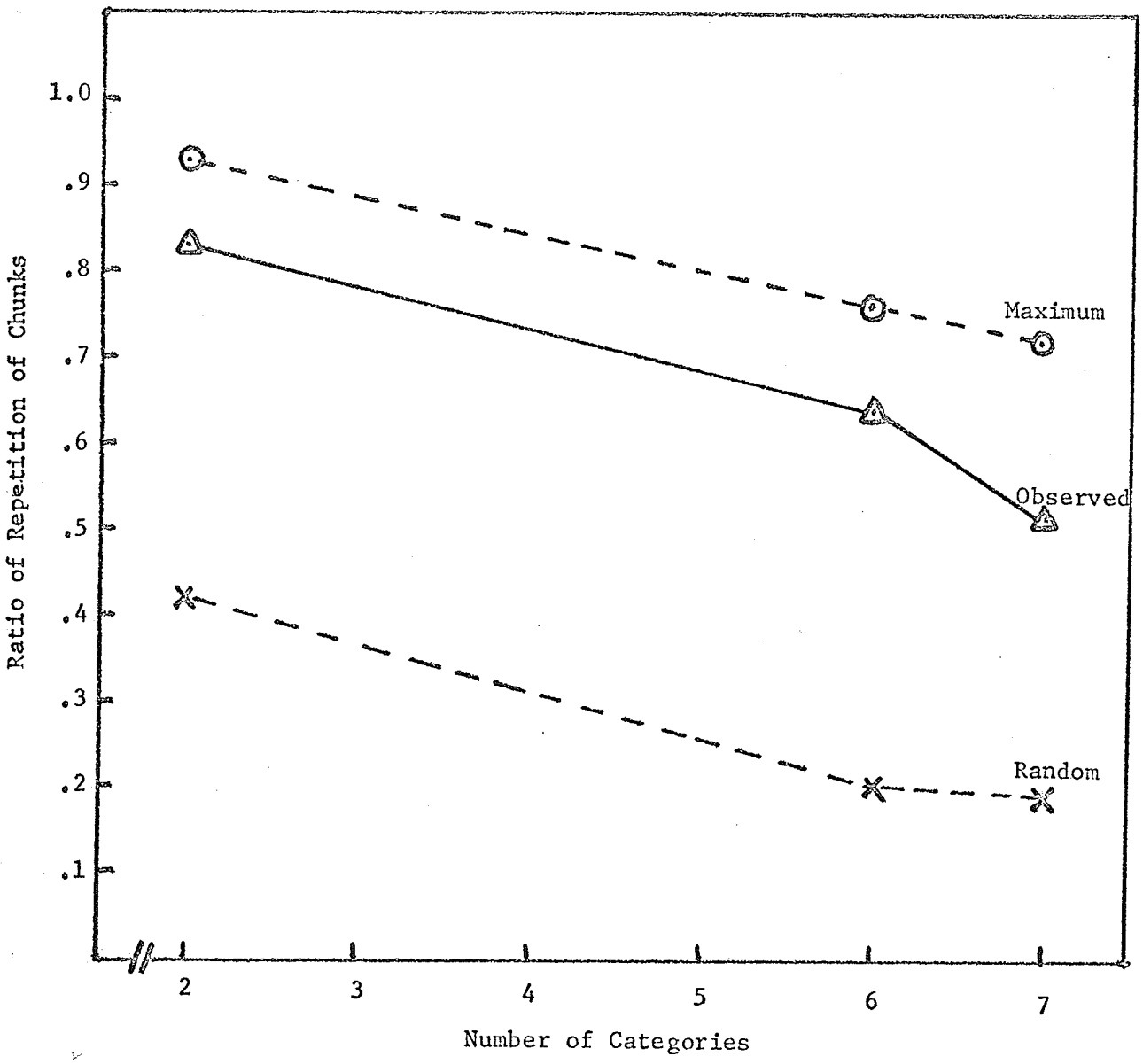

Fig. 2. Mean ratio of repetition of chunks as a function of number of categories used by group RNE. Solid line shows observed data; dashed line with circles shows maximum possible if perfect clustering had occurred; dashed line with crosses shows mean ratio of repetition for a randomly generated output.

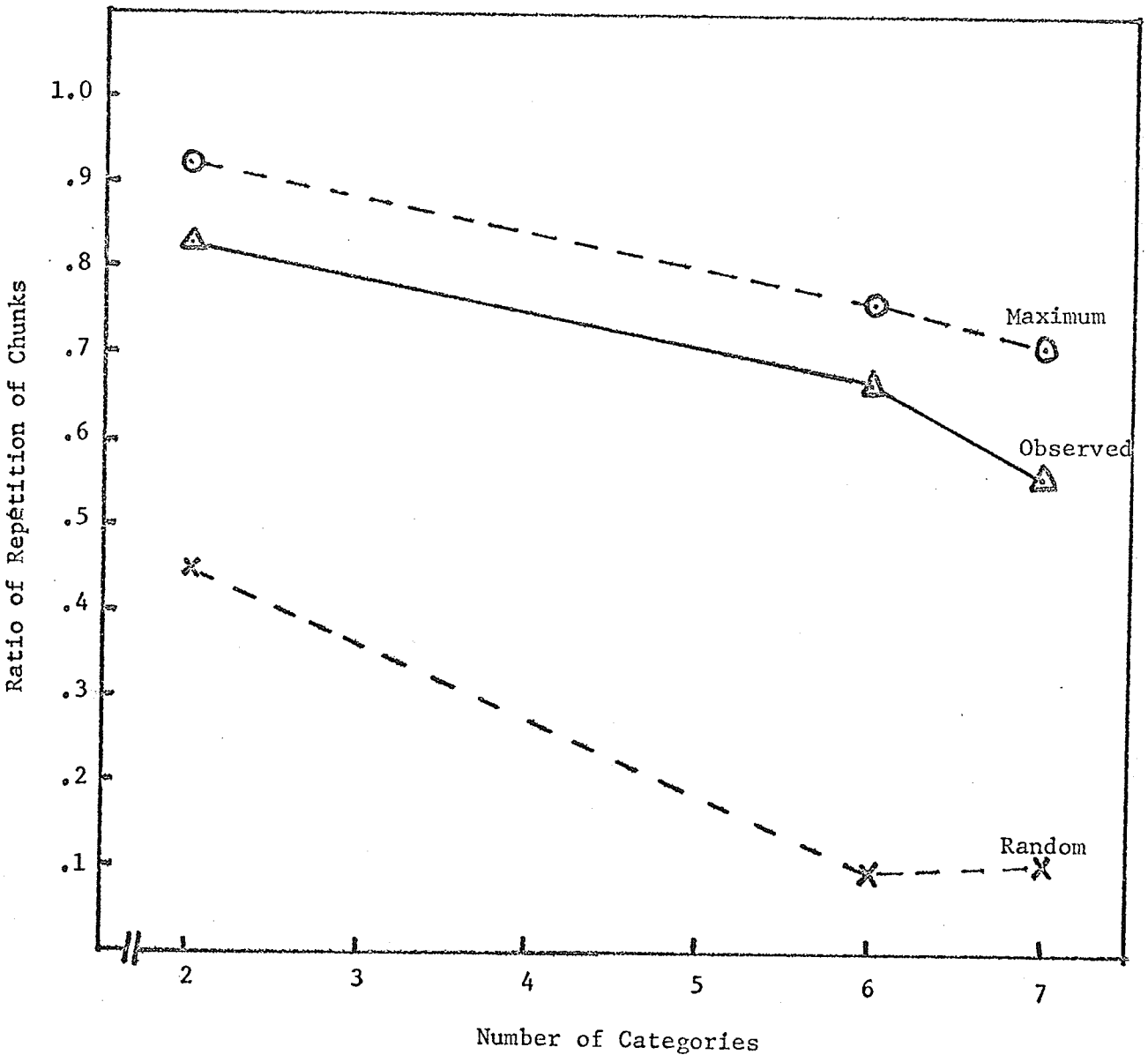

Fig. 3. Mean ratio of repetition of chunks as a function of number of categories used by group RE. Solid line shows observed data; dashed line with circles shows maximum possible if perfect clustering had occurred; dashed line with crosses shows mean ratio of repetition for a randomly generated output.

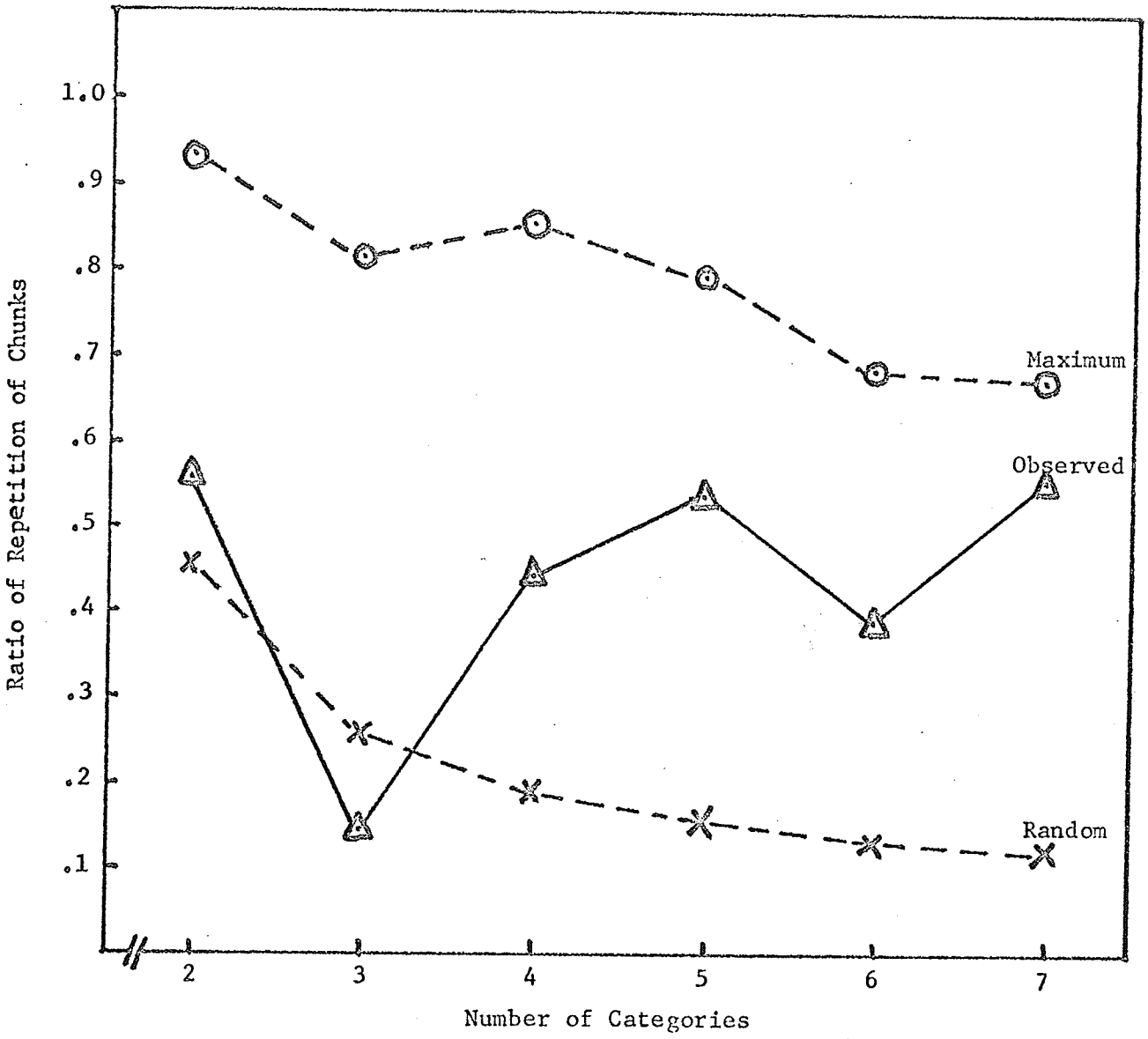

Fig. 4. Mean ratio of repetition of chunks as a function of number of categories used by group BE_1. Solid line shows observed data; dashed line with circles shows maximum possible if perfect clustering had occurred; dashed line with crosses shows mean ratio of repetition for a randomly generated output.

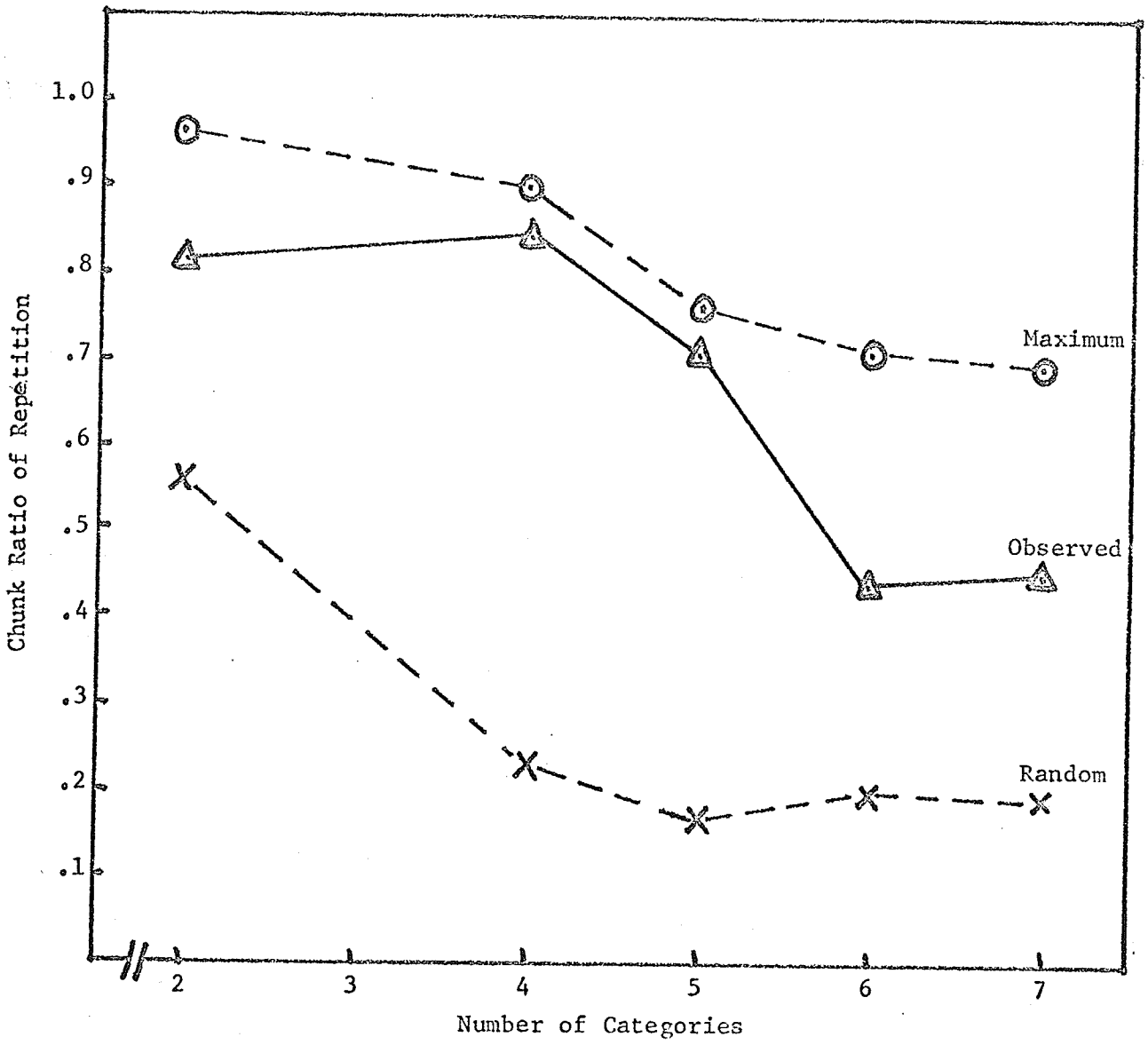

Fig. 5. Mean ratio of repetition of chunks as a function of number of categories used by group BE_2. Solid line shows observed data; dashed line with circles shows maximum possible if perfect clustering had occurred; dashed line with crosses shows mean ratio of repetition for a randomly generated output.

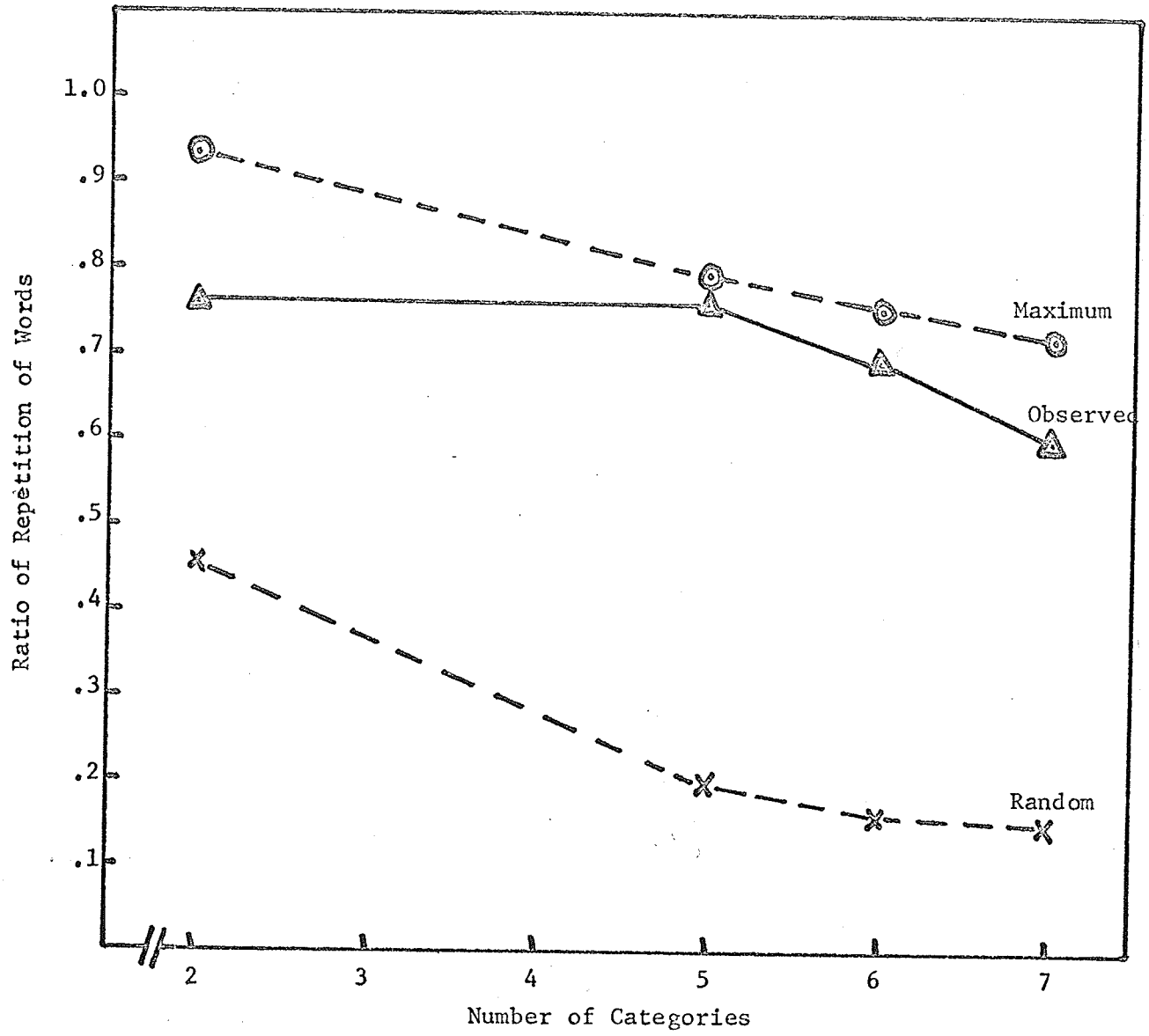

Fig. 6. Mean ratio of repetition of words as a function of number of categories used by group RCII. Solid line shows observed data; dashed line with circles shows maximum possible if perfect clustering had occurred; dashed line with crosses shows mean ratio of repetition for a randomly generated output.

Figure 6 shows that the amount of organization for the 25 word list of group RCII is near maximum for each of the NC used by the Ss. Other than those Ss in group BE_1 who used a low NC, the measures of organization were well above those expected by chance. The trend was generally for organization to increase as NC increased although there were exceptions. As mentioned earlier the different number of categories and words recalled would make comparisons between RR_w and RRc somewhat equivocal. Future studies might employ the modification suggested by Hudson and Dunn (1969) which allow comparisons to be made between groups differing on NC and items recalled. A recommendation made earlier, that S be assigned the NC they were to use, would eliminate data of the sort depicted in Figures 2 and 3, where Ss failed to use 3, 4 or 5 categories.

Interrelationships of NC, sorting trials and items recalled

Table 2 shows the interrelations of several of variables discussed thus far. Groups RNE and RE show a significant correlation ($p<.01$ and $p<.05$ respectively, one tailed tests) between NC and the number of chunks recalled. The correlation coefficient of +.78 shows a highly significant relation between the NC group RE used to sort the cards and the number of chunks recalled. Figure 7 shows that the more NC the Ss used, the greater the number of chunks recalled. The slope for the line of best fit (1.61) indicates that Ss add on the average 1.61 chunks to their recall for every extra category they use. Mandler (1967) found the average slope to be 3.9 and this value was expected in the present investigation. The slope of 1.61 does indicate that Ss do add 4.5 words (i.e. slope X IPC for group RE) for every additional category and it is interesting to speculate if the NC-word recall and NC-chunk recall functions involve different processes.

TABLE 2

CORRELATIONS BETWEEN WORD OR CHUNK RECALL AND

NUMBER OF SORTS OR NUMBER OF CATEGORIES

	Group					
Measures	RCI	RNE	RE	BE_1	BE_2	RCII
Number of categories - chunk recall	----	+.44*	+.78**	+.24	+.16	----
Number of sorts - chunk recall	----	+.03	+.05	+.01	+.10	----
Number of categories - word recall	+.39*	----	----	----	----	+.45*
Number of sorts - word recall	+.09	----	----	----	----	+.26

**$\underline{P} < .01$

*$\underline{P} < .05$

(Note: One tailed tests of significance were used)

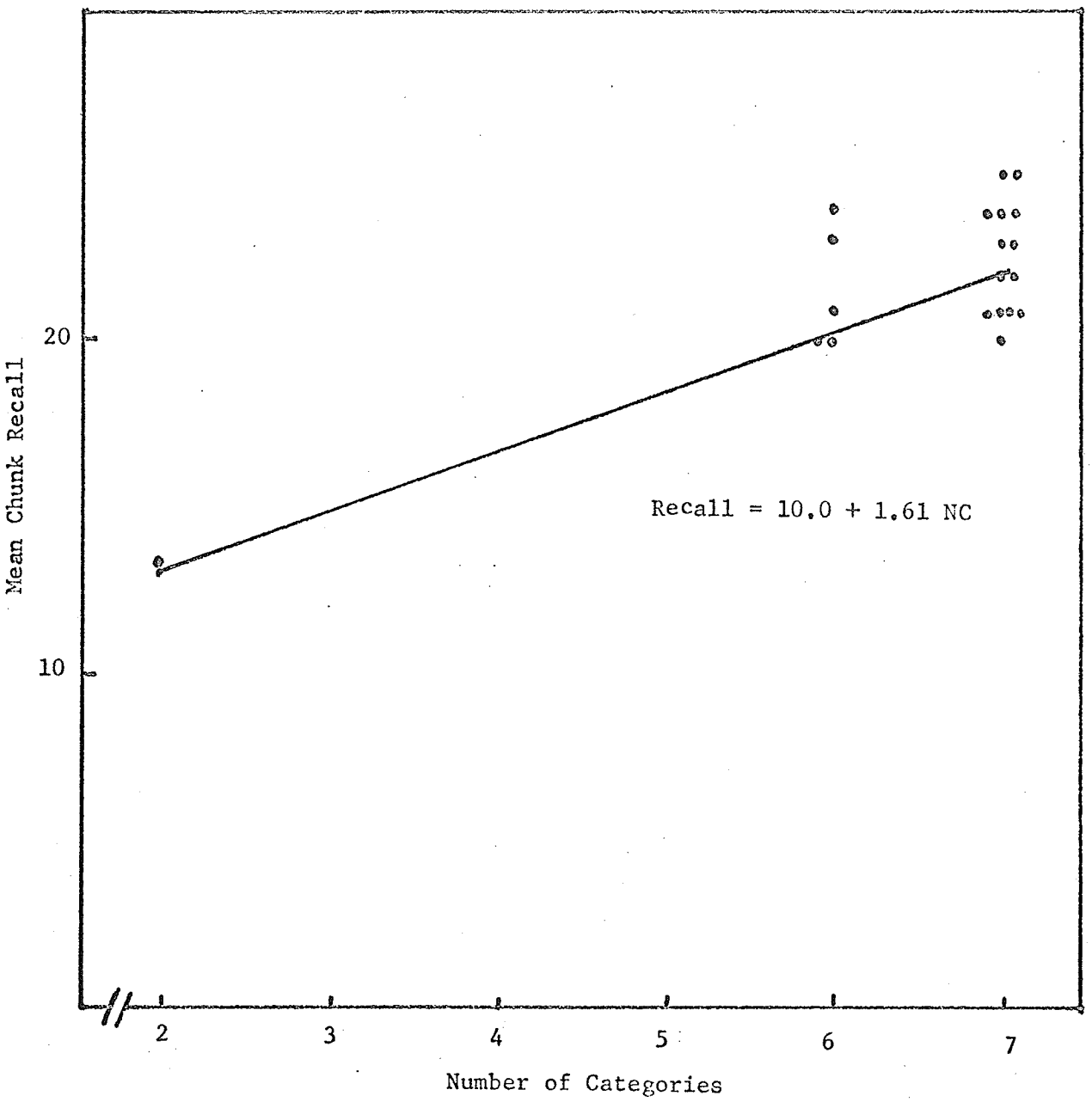

Fig. 7. Mean chunk recall as a function of number of categories used by group RE. The equation shown is for the line of best fit. The dots show the performance of the 20 Ss.

Perhaps the organization of chunks into superchunks cannot proceed at the same rate as the organization of words into chunks. Also, it is possible that it is really the number of words the S must integrate and not the number of chunks, and thus the slope of 5±2 would hold for words but not for chunks. Replication would indicate the validity of these speculations.

The meaning of the Y-intercept equal to 10 is somewhat ambiguous. It suggests, that if a S used zero categories he would recall ten chunks. Mandler (1967) suggests that the function between NC and items recalled is discontinuous between zero and one categories and the intercept provides an index of the amount of organization other than that determined by the NC variable. The Ss could use other organizational strategies to sort the cards. For example, the attributes of the chunks may have suggested rhymes (Fagan, 1969), associations, relational imagery or seriation (Mandler, 1970).

The failure to find the NC-chunk recall relationship for group BE_1 ($r=+.24$) and BE_2 ($r=+.16$) is perhaps the most severe departure from the model of memory espoused by Mandler (1967). Again, replication is in order as the methodology used in the present investigation may have hidden the NC-recall relation. Two alternatives are available for future investigations. The first would be to use the same stimulus card as group BE_1 but to require the Ss to verbalize each item on the card. In addition, the words within a chunk could be randomized as well as the order of the chunks. A more fruitful attempt (after Mandler, 1967) to determine if the NC-chunk recall relationship holds, would be to randomly assign an equal number of Ss to each NC between two and seven. This ensures the full range

of the NC of interest would be sampled.

The NC-word recall relationship for groups RCI and RCII (group RCII is depicted in Figure 8) are in the direction predicted in the experimental hypothesis. The correlations ($r=+.39$ and $r=+.45$ respectively) are less than the median correlation ($r=+.70$) found in Mandler (1967) studies. However, Dong and Kintsch (1968) report a similar correlation of $r=+.44$ ($p<.01$) by summing across Ss in four experimental groups. Mandler and Stephens (1967) also found a significant correlation ($r=+.52$) with words of a frequency not less than 14 per million by Thorndike-Lorge (1944) norms. Mandler and Pearlstone (1966) determined that the NC-word recall held for high frequency words ($r=+.96$) but not for low frequency words.

Finally, as shown in Table 2, there was no correlation between the number of sorts to criterion and the number of items recalled for any of the groups. The finding that Ss need a certain number of exposures to reach a criterion of organization but that it is the nature of the organization which determines recall has again been substantiated.

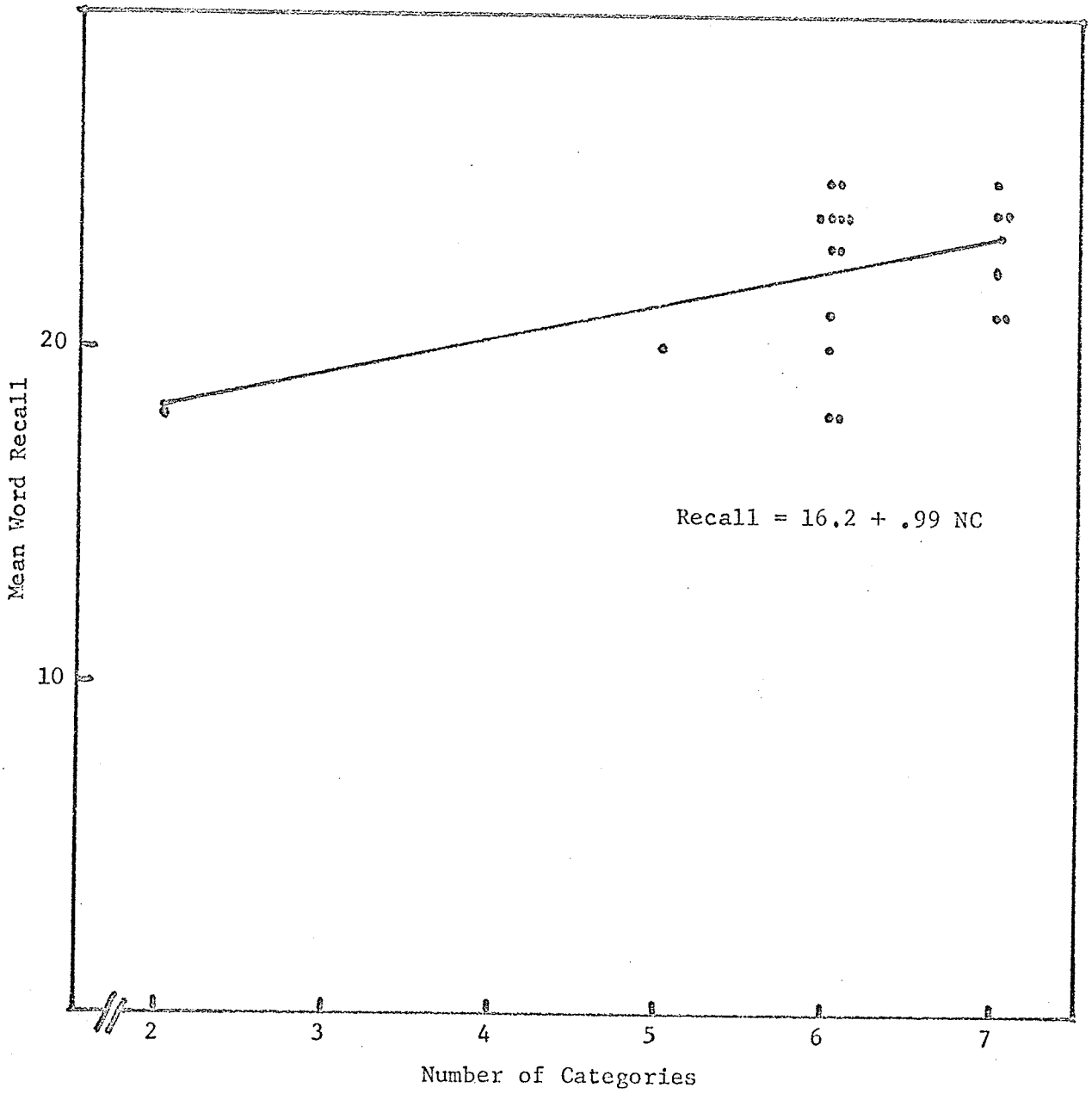

Fig. 8. Mean word recall as a function of number of categories by group RCII. The equation shown is for the line of best fit. The dots show the performance of the 20 Ss.

Number of Items per Chunk

IPC was previously defined as total number of words recalled divided by total number of chunks recalled (and a chunk is defined as the recall of at least one of the chunk numbers). Given a list length of 86 words made up of 25 chunks, the maximum value of IPC for groups RNE, RE and BE_1 is 3.44. Table 1 shows the IPC for the groups. Analysis of variance indicated a significant difference among the groups ($F(2,57)=18.08$, $p<.01$) and the Neuman-Keuls test (Appendix D_8) showed that group BE_1 had the highest IPC (3.06); group RE had the next highest (2.79) and group RNE had the lowest IPC (2.45) as predicted in the experimental hypothesis. Forcing the Ss of Group BE_2 to attend more closely to each chunk member yeilded the highest measure of IPC (3.13).

Groups RNE, RE and BE_1 recalled 71%, 81% and 88% of the words comprising a chunk. Cohen (1963b) found that the RNE and RE categories were represented by 63% and 85% of their words respectively. Although it may be possible to question whether the NE categories are in reality analogous to Miller's chunks, the E cateogries appear to have been handled in this fashion. The six extra exhaustive categories devised for this experiment appear to be relatively exhaustive of their category.

Cohen (in an unpublished study cited in Shuell, 1969) suggested that the blocked, as contrasted with the random presentation, facilitated intracategory (or intrachunk) coding and organization at the expense of intercategory (or interchunk) coding or organization. The present investigation offers limited support to his views. The blocked presentation had a higher IPC than the random presentation (3.06 versus 2.79 respectively); the blocked presentation had fewer chunks recalled compared to the random (21.85 versus 17.05). However, the organization appears to be approximately the same in

both groups.

SUMMARY AND CONCLUSIONS

The results offer partial support for Mandler's (1967, 1968, 1970) hierarchical model of organized memory and are in agreement with related research. The number of subjectively organized "superchunks" (i.e., sorting categories) was limited to a value of 5±2 however, the mean number of sorting categories used by four of the six treatment groups was higher (approximately 6.3) than previous studies (approximately 4.7) in the area. It was suggested that the experimenter-determined relationships in the lists led to the Ss a strategy in which they used extra categories in an attempt to exhaust the obvious relationships. As there were in fact 25 categories in the list this strategy proved to be unfeasible but presumably the Ss persisted in their strategy.

The relation between number of categories and chunk recall offered further support to the model. Significant positive correlations for groups RNE and RE were offset by the failure to find this relationship for Groups BE_1 and BE_2. Certain methodological inadequacies in the mode of presentation of the chunks may have overriden the NC-recall relationship. Tentatively, the results suggest that the S must actively organize from words to chunks to "superchunks" in order to yield this relation. The random groups, both E and NE fulfilled this requirement whereas the BE groups were given the word to chunk organization and had only to organize the chunks into "superchunks." Consistent with the hierarchical model of memory was the failure to find any relation between the sorts to criterion and chunk or word recall.

The measures of organization of the chunks offered the most convincing support to the model. With one exception (i.e., the low NC Ss of group BE_1) there was a distinct tendency for chunks within a superchunk (or category) to be recalled in clusters. The amount of clustering, with this one exception, was well above a randomly generated amount of clustering.

Several findings, incidental to the main findings just presented, are of interest. The combination of S - determined organization superimposed over E-determined relations led to high recall scores. The lowest group (RNE) recalled over 50% and the highest group (BE_2) recalled nearly 80% of the 86 word lists. Blocked presentation of chunks facilitated intrachunk coding and recall while random presentation facilitated interchunk coding as determined from measures of chunk recall and IPC consistent with the findings of Cohen (1963a; 1963b).

Future experiments in the study of organization of chunks through use of the card sorting technique should incorporate two modifications. First, Ss should be randomly assigned to a NC between two and seven to represent the full range of interest. Second, the BE_1 procedure should require Ss to verbalize each item of the chunk to reduce stimulus selection. A further control might be to randomize the order of items within a chunk as well as randomize the order of chunks. Also, investigations which require a comparison of organization among groups differing in the amount of recall and NC might profitably employ the standardized score technique of Hudson and Dunn (1969).

In conclusion, there was a degree of support to Mandler's model of organized memory. Several inadequacies in the present design led to

the inability to discriminate between limitations in the method of presenting the materials or limitations in the theory. Replications in the area could be directed towards solving this problem.

REFERENCES

Adams, J. Human Memory, New York: McGraw-Hill, 1967.

Allen, M., Puff, C.R., and Weist, R. Effects of Associative and Coding Processes on Organization in Free Recall. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 531-538.

Battig, W., & Montague, W. Category normals for verbal items in 56 categories. Journal of Experimental Psychology Monograph, 1969, 80, No. 3, Part 2.

Bousfield, A., & Bousfield, W. Measurement of clustering and of sequential constancies in repeated free recall. Psychological Reports, 1966, 19, 935-942.

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

Bousfield, W., & Cohen, B. The effects of reinforcement in the recall of randomly arranged associates. Journal of Psychology, 1953, 36, 67-81.

Bousfield, W., & Cohen, B. The occurence of clustering in the recall of randomly arranged words of different frequencies-of-usage. Journal of General Psychology, 1955, 52, 83-95.

Bousfield, W., Cohen, B., & Whitemarsh, G. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychological Reports, 1953, 4, 39-44.

Bower, G. Chunks as interference units in free recall. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 610-613.

Bower, G.H., Clark, M., Lesgold, A., &Winzenz, D. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 323-343.

Bower, G., Lesgold, A.M., & Tieman, D. Grouping operations in free recall Journal of Verbal Learning and Verbal Behavior, 1969, 8, 481-493.

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

Cofer, C., Bruce, D., & Reicher, G. Clustering in free recall as a function of certain methodological considerations. Journal of Experimental Psychology, 1966, 71, 858-866.

Cohen, B. An investigation of recoding in free recall. Journal of Experimental Psychology, 1963, 65, 368-376. (a)

Cohen, B. Recall of categorized word lists. Journal of Experimental Psychology, 1963, 66, 227-234. (b)

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

D'Agostino, P. The blocked-random effect in recall and recognition. Journal of Verbal Learning and Verbal Behavior. 1969, 8, 815-820.

Dallet, K. Number of categories and category information in free recall. Journal of Experimental Psychology, 1964, 68, 1-12.

Deese, J. Influence of inter-item associative strength upon immediate free recall. Psychological Reports, 1959, 5, 305-312.

Deese, J. Association and memory. In T.R. Dixon and D.L. Horton (Eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Dong, T., & Kintsch, W. Subjective retrieval cues in free recall. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 813-816.

Earhard, M. Cued recall and free recall as a function of the number of items per cue. Journal of Verbal Learning and Verbal Behavior. 1967, 6, 257-263.

Earhard, M. The facilitation of memorization by alphabetic instructions. Canadian Journal of Psychology. 1967, 21, 15-24.

Earhard, M. Storage and retrieval of words encoded in Mmeory. Journal of Experimental Psychology. 1969, 80, 412-418.

Fagan, J.F. Clustering of related but nonassociated items in free recall. Psychonomic Science, 1969, 16, 92-93.

Foote, R.A., & Pollio, H.R. Role of associative organization in free recall under 5 different presentation conditions. Psychonomic Science. 1970, 20, 205-207.

Garner, W., & Whitman, J. Form and amount of interval structure as factors in free recall learning of nonsense words. Journal of Verbal Learning and Verbal Behavior. 1969, 4, 257-266.

Helson, R., & Cover, A. Specificity-generality of classificatory categories as a variable in recall. Perceptual and Motor Skills, 1956, 6 233-236.

Hudson, R., & Dunn, R. A major modification of the Bousfield (1966) measure of category clustering. Behavior Research Methods and Instrumentation, 1969, 1, 110-111.

Jenkins, J.J., Mink, W.D., & Russell, W.A. Associative clustering as a function of verbal association strength. Psychological Report, 1958, 4, 127-136.

Jenkins, J.J., & Russell, W. A. Associative clustering during recall. Journal of Abnormal and Social Psychology, 1952, 47, 818-821.

Johnson, N. Sequential verbal behavior. In T. Dixon and D. Horton (Eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Johnson, N.P. Chunking, Associative chaining versus coding, Journal of Verbal Learning and Verbal Behavior., 1969, 8, 725-731.

Kendler, H.H., Coding: associationistic or organizational. Journal of Verbal Learning and Verbal Behavior., 1966, 5, 198-200

Kintsch, W., Learning, Memory and Conceptual Processes, New York: Wiley, 1970.

Mandler, G. Organization and memory. In K.W. Spence and J.T. Spence (Eds.), The Psychology of Learning and Motivation. New York: Academic Press, 1967.(A).

Mandler, G. Association and organization. In T.R. Dixon and D.L. Horton (Eds.) Verbal Behavior and General Behavior Theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968. (a)

Mandler, G. Words, lists and categories: An experimental view of organized memory. In J.L. Cowan (Ed.) Studies in Thought and Language. 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.

Mandler, G., Pearlstone, Z., & Koopmans, H. Effects of organization and semantic similarity on recall and recognition. Journal of Verbal Learning and Verbal Behavior. 1969, 8, 410-423.

Mandler, G., & Stephens, D. The development of free and constrained conceptualization and subsequent verbal memory. Journal of Experimental Child Psychology, 1967, 5, 86-93.

Mathews, R. Recall as a function of number of classifactory categories. Journal of Experimental Psychology, 1954, 47, 241-247.

McNulty, J. The measurement of adopted chunks in free recall learning. Psychonomic Science, 1966, 4, 71-72.

Miller, G.A., Human memory and the storage of information. IRE Transactions of information theory. 1956, IT-2, 129-137.

Miller, A. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-96.

Murdock, B. The serial position curve of free recall. Journal of Experimental Psychology, 1962, 64, 482-488.

Pollio, H., Kasschau, R., & DeNise, H. Associative structure and the temporal characteristics of free recall. Journal of Experimental Psychology, 1968, 76, 190-197.

Pollio, H.R., Richards, E., & Lucas, R. Temporal properties of category recall. Journal of Verbal Learning and Verbal Recall. 1969, 8 529-536.

Puff, C. Clustering as a function of the sequential organization of stimulus word lists. Journal of Verbal Learning and Verbal Behavior. 1966, 5, 503-506.

Puff, C.R., & Hysen, SPP. An Empirical comparison of two measures of intertrial organization in free recall. Psychonomic Science. 1967, 9, 329-330.

Rundas, D., & Atkinson, R.C. Rehearsal processes in free recall: a procedure for direct observation. Journal of Verbal Learning and Verbal Behavior. 1970, 9, 99-105.

Segal, E.M. Hierarchical structure in free recall. Journal of Experimental Psychology, 1969, 80, 59-63.

Shepard, R. Learning and recall as organization and search. Journal of Verbal Learning and Verbal Behavior. 1966, 5, 201-204.

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

Slamecka, N.J. An examination of trace storage in free recall. Journal of Experimental Psychology, 1968, 76, 504-513.

Tulving, E. Subjective organization in free recall of "unrelated" words. Psychological Review, 1962, 69, 344-354.

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

Tulving, E., & Osler, S. Effectiveness of retrieval cues in memory for words. Journal of Experimental Psychology, 1968, 77, 593-601.

Tulving, E., & Patkau, J.E. Concurrent effects of contextual constraint and word frequency of immediate recall and learning of verbal material. Canadian Journal of Psychology, 1962, 77, 239-248.

Tulving, E., & Patterson, E. Functional units and retrieval processes in free recall. Journal of Experimental Psychology, 1968, 77, 239-248.

Underwood, B.J., & Erlebacher, A.H., Studies of coding in verbal learning. Psychological Monographs, 1965, 79, (13).

Underwood, B.J., & Reppel, G. Coding processes in verbal learning. Journal of Verbal Learning and Verbal Behavior, 1963, 1, 250-257.

Wallace, W. Consistency order in free recall. Journal of Verbal Learning and Verbal Behavior. 1970, 9, 58-68.

Waught, N., & Norman, D. Primary memory. Psychological Review, 1965, 72. 89-102.

Weist, R.M. Optimal versus nonoptimal conditions for retrieval. Journal of Verbal Learning and Verbal Behavior. 1970, 9, 311-316.

Whitman, J.R., & Garner, W.R. Free recall learning of visual figures as a function of form of internal structure. Journal of Experimental Psychology. 1962, 64, 558-564.

Wood. G. Category names as cues for the recall of category instances. Psychonomic Science, 1967, 9, 323-324.

Wood. G. Retrieval cues and the accessibility of higher-order memory units in multi-trial free recall. Journal of Verbal Learning and Verbal Behavior. 1969, 8, 782-789.

APPENDIX A

Words used in the designated group along with their Thorndike-Lorge (1944) frequency of usage. The underlined words were randomly selected to comprise the list for group RCII.

Group					
Exhaustive		Nonexhaustive		Random	
Word	Freq.	Word	Freq.	Word	Freq.
army	AA	cotton	AA	army	AA
navy	49	linen	47	linen	47
airforce	-	satin	15	arrow	37
animal	AA	chair	AA	animal	AA
vegetable	A	desk	A	desk	A
mineral	38	couch	28	juice	37
fork	31	rose	AA	grass	AA
knife	A	violet	34	violet	34
spoon	33	daisy	38	spoon	38
volt	2	waltz	3	volt	2
ampere	-	tango	-	waltz	3
ohm	-	polka	-	haystack	1
breakfast	A	hat	AA	breakfast	A
lunch	39	coat	AA	paint	AA
dinner	AA	shirt	47	shirt	47

Group					
Exhaustive		Nonexhaustive		Random	
Word	Freq.	Word	Freq.	Word	Freq.
cigarette	22	truck	23	cigarette	22
cigar	16	taxi	17	truck	23
pipe	A	wagon	A	flag	A
caucasian	1	onyx	1	ruby	10
oriental	14	ruby	10	oriental	14
negro	47	pearl	48	library	A
blond	17	treason	15	blond	17
brunette	2	burglary	1	treason	15
redhead	-	larceny	1	veal	5
ice	AA	oil	AA	ice	AA
water	AA	coal	AA	coal	AA
steam	A	gas	A	winter	AA
Protestant	13	wrench	11	wrench	11
Catholic	41	file	43	noon	A
Jew	19	drill	21	Jew	19
masculine	8	violin	11	masculine	8
feminine	8	flute	9	violin	11
neuter	1	saxophone	1	cocoon	3

Group					
Exhaustive		Nonexhaustive		Random	
Word	Freq.	Word	Freq.	Word	Freq.
wall	AA	book	AA	end	AA
floor	AA	newspaper	AA	floor	AA
ceiling	23	novel	39	novel	39
principal	A	Jim	A	principal	A
student	A	Bob	A	Jim	A
teacher	AA	John	AA	class	AA
federal	49	pot	47	federal	49
provincial	8	saucer	7	saucer	7
municipal	14	refrigerator	11	garage	14
inch	AA	roof	AA	member	AA
foot	AA	window	AA	door	AA
yard	AA	door	AA	yard	AA
mile	AA	floor	AA	fear	AA
hearts	AA	cousin	A	niece	10
spades	13	niece	10	spades	13
diamonds	A	aunt	A	eye	AA
clubs	AA	uncle	AA	fact	AA
cocaine	1	aspen	1	cocaine	1
heroin	-	sequoia	1	referee	3
morphine	1	hemlock	7	hemlock	7
opium	7	spruce	11	creek	27

Group					
Exhaustive		Nonexhaustive		Random	
Word	Freq.	Word	Freq.	Word	Freq.
mother	AA	dog	AA	mother	AA
father	AA	horse	AA	queen	AA
sister	AA	bear	AA	support	AA
brother	AA	lion	A	lion	A
day	AA	blue	AA	wind	AA
week	AA	red	AA	green	AA
month	AA	green	AA	month	AA
year	AA	yellow	AA	century	AA
addition	A	Russia	A	farmer	AA
subtraction	1	Burma	1	mongrel	2
division	5	Peru	7	Peru	7
multiplication	A	Spain	A	multiplication	A
bass	7	bayonet	9	dagger	8
tenor	6	dagger	8	tenor	6
alto	1	cutlass	2	ballet	2
soprano	1	stiletto	-	glue	15

Group					
Exhaustive		Nonexhaustive		Random	
Word	Freq.	Word	Freq.	Word	Freq.
north	AA	mountain	AA	north	AA
south	AA	hill	AA	hill	AA
east	AA	valley	AA	taste	AA
west	AA	river	AA	paper	AA
winter	AA	arm	AA	leg	AA
summer	AA	leg	AA	summer	AA
spring	AA	head	AA	cross	AA
fall	AA	foot	AA	boat	AA
penny	38	bishop	40	penny	38
nickel	11	reverend	11	ink	20
dime	11	pastor	11	pastor	11
quarter	AA	minister	A	office	AA
hurricane	7	adjective	7	tutor	10
tornado	3	pronoun	4	tornado	3
typhoon	-	adverb	-	sofa	14
monsoon	1	noun	1	noun	1

APPENDIX B

INSTRUCTIONS

This is a study of how people organize words. In front of you are cards with words printed on them and outlined on the table are seven sorting categories. Your task will be to sort these cards, one at a time, into anywhere from two to seven categories on any basis, by any system or set of rules that you wish, with one exception. Do not use the alphabet as the basis for sorting the cards. Later, you will be required to recall as many words as possible. When you have finished sorting this deck of cards you will be given another deck of cards containing the same words but in a different order. Again, use any set of rules you wish, to sort these cards. However, to end the experiment, you must sort the cards in identical fashion twice in a row.

Once you place a card down--leave it down. You may change it on the next sort. Also look only at the top card on each of the sorting categories.

To review:

1. Turn over the cards, one at a time and use any set of rules except the alphabet to sort the cards into from two to seven categories.

2. After each sort you will be given another deck of the same words but in a different order. You are required to sort the cards in the same fashion twice in a row.

3. Look only at the top card on each pile.

4. Remember that you will be required to recall as many words as you can.

ARE THERE ANY QUESTIONS?

Recall Instructions

NOW RECALL AS MANY WORDS AS YOU CAN, IN ANY ORDER THAT YOU WISH. START AT THE TOP OF THE PAGE AND WORK YOUR WAY DOWN. FILL IN EACH BLANK IN SUCCESSION (WITH ONE WORD).

APPENDIX C

RAW DATA
NUMBER OF SORTING CATEGORIES

	Group					
Subjects	RCI	RNE	RE	BE 1	BE 2	RCII
1	7	7	6	2	7	6
2	7	7	6	3	7	6
3	7	5	6	6	7	6
4	7	5	2	6	7	6
5	7	7	7	7	6	6
6	7	7	7	5	4	7
7	6	7	7	6	6	7
8	6	7	7	2	6	7
9	7	7	7	6	6	6
10	6	6	7	4	7	6
11	7	7	7	5	2	6
12	7	7	7	2	7	6
13	3	4	7	3	6	6
14	7	7	7	6	4	6
15	7	6	6	7	2	5
16	7	7	7	4	5	7
17	3	5	7	6	6	2
18	3	6	7	2	7	7
19	7	3	7	4	6	6
20	7	4	6	4	5	7

NUMBER OF SORTS TO CRITERION

Subjects	Group RCI	RNE	RE	BE_1	BE_2	RCII
1	3	3	3	2	5	6
2	5	4	5	8	4	8
3	4	3	4	6	3	7
4	4	3	4	4	4	2
5	3	5	5	3	3	4
6	5	5	5	3	4	2
7	3	5	3	4	5	2
8	4	4	3	3	4	2
9	6	3	2	7	3	4
10	5	2	3	5	5	3
11	4	3	4	3	3	6
12	3	7	3	5	4	6
13	6	4	5	3	4	7
14	5	2	5	7	5	4
15	4	3	4	4	6	3
16	5	3	6	2	4	3
17	3	2	4	5	5	4
18	4	3	4	3	4	3
19	5	4	3	3	3	3
20	4	5	4	8	4	3

NUMBER OF WORDS RECALLED

Subjects	Group RCI	RNE	RE	BE_1	BE_2
1	57	32	48	74	73
2	65	66	74	48	57
3	62	48	59	44	70
4	55	55	27	61	52
5	64	31	79	70	45
6	69	57	66	51	79
7	64	56	56	43	73
8	60	67	66	41	82
9	53	59	62	33	71
10	65	63	50	58	49
11	27	64	55	63	56
12	73	52	64	44	69
13	60	45	60	25	56
14	45	57	60	48	54
15	62	38	63	61	66
16	70	36	52	42	46
17	29	53	60	70	52
18	38	45	72	45	73
19	49	37	73	50	53
20	47	38	70	69	52

NUMBER OF CHUNKS RECALLED

Subject	Group RNE	RE	BE_1	BE_2	RCII
1	20	20	22	23	25
2	22	23	15	18	25
3	17	20	14	21	24
4	22	13	21	17	21
5	14	25	21	16	20
6	22	20	19	24	21
7	24	21	15	23	21
8	24	24	13	24	25
9	25	21	11	21	24
10	21	23	22	16	18
11	21	22	19	17	18
12	22	21	15	22	24
13	19	24	8	20	23
14	21	22	16	18	23
15	17	21	19	20	20
16	20	21	13	15	24
17	20	23	22	18	18
18	19	25	15	23	24
19	17	24	18	19	24
20	19	24	23	16	22

WORD RATIO OF REPETITION

Subject	Group RCI	RNE	RE	BE_1	BE_2	RCII
1	.84	.32	.78	.67	.78	.71
2	.38	.68	.86	.75	.84	.79
3	.87	.85	.90	.81	.88	.78
4	.52	.83	.92	.92	.75	.75
5	.86	.53	.80	.90	.70	.74
6	.81	.55	.91	.84	.96	.70
7	.75	.80	.64	.76	.75	.45
8	.85	.81	.74	.85	.85	.67
9	.56	.81	.79	.72	.76	.65
10	.69	.90	.57	.84	.79	.41
11	.54	.83	.80	.86	.96	.65
12	.85	.83	.72	.88	.69	.78
13	.61	.88	.66	.75	.74	.77
14	.46	.89	.85	.77	.94	.68
15	.80	.76	.89	.85	.92	.79
16	.80	.53	.75	.78	.91	.74
17	.54	.89	.88	.80	.78	.77
18	.68	.71	.59	.91	.81	.65
19	.60	.75	.88	.92	.83	.57
20	.67	.84	.88	.75	.83	.36

CHUNK RATIO OF REPETITION

Subject	Group RNE	RE	BE_1	BE_2
1	.21	.53	.26	.45
2	.33	.59	.14	.59
3	.75	.75	.46	.60
4	.67	.83	.75	.19
5	.31	.67	.75	.27
6	.29	.68	.56	.87
7	.61	.40	.22	.27
8	.65	.65	.50	.59
9	.67	.35	.10	.32
10	.75	.18	.62	.40
11	.70	.57	.50	.88
12	.62	.45	.64	.24
13	.83	.44	.14	.32
14	.75	.62	.27	.82
15	.56	.70	.50	.74
16	.42	.45	.25	.71
17	.63	.68	.38	.47
18	.61	.29	.79	.55
19	.13	.65	.77	.67
20	.67	.70	.36	.69

NUMBER OF ITEMS RECALLED PER CHUNK RECALLED

Subject	Group RNE	RE	BE_1	BE_2
1	1.60	2.40	3.36	3.17
2	3.00	3.22	3.20	3.17
3	2.82	2.95	3.14	3.33
4	2.50	2.08	2.91	3.06
5	2.21	3.16	3.33	2.81
6	2.59	3.30	2.69	3.29
7	2.33	2.67	2.86	3.17
8	2.79	2.85	3.15	3.41
9	2.36	2.95	3.30	3.38
10	3.00	2.18	2.64	3.06
11	3.05	2.50	3.21	3.29
12	2.36	3.05	2.93	3.14
13	2.37	2.50	3.13	2.80
14	2.71	3.00	3.00	3.00
15	2.23	3.00	3.21	3.30
16	1.80	2.41	3.23	3.06
17	2.65	.261	3.18	2.89
18	2.37	2.88	3.00	3.17
19	2.17	3.04	2.78	2.79
20	2.00	2.91	3.00	3.25

APPENDIX D_1

ANALYSIS OF VARIANCE OF TRIALS TO CRITERION FOR PAID AND UNPAID Ss

Source of Variation	SS	DF	MS	F
Between Ss	13.38	9		
Sorts to Criterion	6.66	4	1.88	/1.0
Pay/No Pay	1.65	1	1.65	/1.0
Pay X Sorts	5.07	4	1.27	/1.0
Within Ss	197.12	90	2.19	

APPENDIX D_2

ANALYSIS OF VARIANCE OF NUMBER OF SORTING CATEGORIES FOR PAID AND UNPAID Ss

Source of Variation	SS	DF	MS	F
Between Ss	62.19	9		
Number of Categories	49.65	4	12.41	6.71*
Pay/No Pay	.72	1	.72	/1.0
Category X Pay	11.82	4	2.95	1.59
Within Ss	175.65	90	1.85	

* p/.01

APPENDIX D_3

ANALYSIS OF VARIANCE OF WORD RECALL FOR PAID AND UNPAID Ss

Source of Variation	SS	DF	MS	F
Between Ss	2018.98	7		
Word Recall	1142.80	3	380.93	2.51
Pay/No Pay	318.10	1	318.10	2.09
Word X Pay	558.08	3	186.03	1.22
Within Ss	10,943.91	72	151.98	

APPENDIX D_4

RESULTS OF THE NEWMAN-KEULS TEST OF THE NUMBER OF SORTING CATEGORIES (NC)

The Newman-Keuls procedure described in Kirk (1968, pp. 91-93) was used. The difference that a comparison must exceed, W_r, for the test is given by the formula

$$W_r = q_{r\,\alpha\,;\,r,\nu} \sqrt{\frac{\text{MS error}}{n}}$$

This formula is applicable to the following five post hoc tests where
MS error = 1.8495
n = 20
α = .05
ν = 95
r = number of steps means are apart = 2,3,4 or 5

W_2 = .84
W_3 = 1.01
W_4 = 1.11
W_5 = 1.19

DIFFERENCES AMONG MEANS

	RE	RCI	RCII	RNE	BE_1
$\bar{X}_{RE}$ = 6.50		.25	.44	.45	2.00*
$\bar{X}_{RCI}$ = 6.25			.19	.20	1.75*
$\bar{X}_{RCII}$ = 6.06				.01	1.56*
$\bar{X}_{RNE}$ = 6.05					1.55*
$\bar{X}_{BE}$ = 4.50					

* $p < .05$

APPENDIX D_5

RESULTSOOF NEWMAN-KEULS TEST OF WORD RECALL

In this case

MS error = 152.64
n = 20
α = .05
$\sqrt{}$ = 76
r = number of steps means are apart = 2,3 or 4

W_2 = 7.79
W_3 = 9.36
W_4 = 10.03

DIFFERENCES AMONG MEANS

	RE	RCI	BE	RNE
$\bar{X}$ RE = 60.8		5.1	8.6	10.9*
$\bar{X}$ RCI = 55.7			3.5	5.8
$\bar{X}$ BE = 52.2				2.3
$\bar{X}$ RNE = 49.9				

* $p < .05$

APPENDIX D_6

RESULTS OF THE NEWMAN-KEULS TEST OF CHUNK RECALL

In this case:

MS error = 9.433
n = 20
α = .05
ν = 76
r = number of steps means are apart = 2,3 or 4

W_2 = 1.87
W_3 = 2.33
W_4 = 2.56

DIFFERENCES AMONG MEANS

	RCII	RE	RNE	BE
$\bar{X}$ RCII = 22.20		.35	1.90	5.15*
$\bar{X}$ RE = 21.85			1.55	4.80*
$\bar{X}$ RNE = 20.30				3.25*
$\bar{X}$ BE = 17.05				

* $p < .05$

APPENDIX D_7

RESULTS OF NEWMAN-KEULS TEST OF WORD RATIO OF REPITITION (RRW)

In this case:

MS error = 343.69
n = 20
α = .05
ν = 76
r = number of steps means are apart = 2,3, or 4

W_2 = 11.69
W_3 = 14.05
W_4 = 15.45

DIFFERENCES AMONG MEANS

	RCI	RE	RNE	BE
RCI $\bar{X}$ = 67.0		10.8	11.3	22.2*
RE $\bar{X}$ = 56.2			.50	11.4
RNE $\bar{X}$ = 55.7				10.9
BE $\bar{X}$ = 44.8				

* $p < .05$

APPENDIX D_8

RESULTS OF NEWMAN-KEULS TEST OF ITEMS PER CHUNK (IPC)

In this case:

MS error = .1056
n = 20
α = .05
ν = 57
r = number of steps between means = 2 or 3

W_2 = .206
W_3 = .247

DIFFERENCES AMONG MEANS

	BE	RE	RNE
$\overline{X}_{BE}$ = 3.06		.27*	.61*
$\overline{X}_{RE}$ = 2.79			.34*
$\overline{X}_{RNE}$ = 2.45			

* $p < .05$