Theoretical Issues in Free Recall[1]

1

ENDEL TULVING

The task of the psychologist interested in memory, like that of any other scientist, is the description and explanation of phenomena subsumable under the general heading of his chosen subject matter. Although his ultimate concern lies in the understanding of these phenomena as they occur in the natural environment, the descriptive and explanatory statements he makes are based on the evidence gained from laboratory experiments. Experimental paradigms used in the study of memory represent miniature analogues of various kinds of "real life" situations.

When we ask a person to tell us the names of all the people he met the day before, or the names of all the books he read in 1965, or the names of the cities he has visited at least once in his lifetime, we are asking the person to utilize stored mnemonic information in the same manner as he does in free-recall experiments. In these experiments the subject is presented with a collection or series of stimulus items, and at some later time he is asked to reproduce from memory the names of as many of these items as he can, in any order he wishes. The analysis of the respondent's performance in both tasks begins with the assumptions that he has had a number of separate and temporally successive experiences and that his ability to recall these experiences depends on a multitude of factors. In laboratory experiments these factors can be identified, isolated, and controlled, and their effect on performance evaluated, but there are no compelling reasons to believe that the relevant variables determining the subject's mnemonic performance and the underlying processes involved in remembering past experiences in the laboratory are different from what they are outside the laboratory.

Human memory, of course, encompasses a great deal more than storage and retrieval of information about occurrences of certain events in a smaller or larger segment of a person's life. We may ask the person only to list the names of people he has met, or of books he has read, or of cities he has visited, but he undoubtedly remembers a great deal more about each event than just its general identifying label. In a free-recall experiment the subject likewise remembers more about the items he saw in the window of the memory

[1] The preparation of this chapter has been supported by the National Research Council of Canada (Grant No. APT-39) and the National Science Foundation (Grant No. GB-3710).

drum than just their names. But since the ability to reproduce aspects of individual experiences *qua* individual experiences is one of the fundamental properties of the human mind, and since more complex mnemonic skills of an articulate person both inside and outside the laboratory include this ability, the study of free recall constitutes a logical starting point in our quest for the understanding of human memory.

In this chapter we will review the present status of conceptual developments in free recall by focusing on those theoretical issues that seem to have generated more than passing interest. Because of the relative novelty of the subject matter of free recall, the chapter begins with a short excursion into the history of this phenomenon. The subsequent five main sections of the chapter are concerned with (a) some orienting attitudes and pretheoretical assumptions, (b) single-trial free-recall phenomena, (c) secondary organization in free recall, (d) repetition effects in multitrial free recall, and (e) storage versus retrieval processes. A summary will present a résumé of theoretical issues and an evaluation of the present state of our understanding of free recall.

Some Historical Antecedents

The method of retained members

Free recall as an experimental method of studying memory is the descendant of the method of retained members. The latter method was first used by several investigators before the turn of the century (e. g., Bolton, 1892; Kirkpatrick, 1894; T. L. Smith, 1896; W. G. Smith, 1896). It received its name from Ebbinghaus (1902), although the great master himself had no use for it. It involved either simultaneous or successive presentation to the subject of a collection of items—objects, pictures of objects, words, syllables, geometric figures, letters, digits, and the like—either once or more, and the recall by the subject of as many of the responses corresponding to the presented items as possible, usually immediately or a few seconds after the completion of the presentation. The order of recall in which the subject recalled the items was not specified by the experimenter.

The method of retained members proved to be quite popular with early students of memory and was adopted for the study of a variety of problems. In 1906, for instance, Pohlmann listed almost thirty publications reporting on the work done with this approach. These early experiments were concerned with recall of individual items as a function of variables such as characteristics of items, amount of material presented, sensory modality in which material was presented, number of repetitions or amount of study time, grouping of items, constant versus variable order of presentation, serial position, length of retention interval, nature of activities interpolated between presentation and recall, age and sex of subjects, and the like. This list of independent variables must look quite familiar to experimenters of the present generation, although many of the early experiments would not quite measure up to the more exacting standards of today's research.

The initial popularity of the method of retained members, however, was relatively short-lived. The orienting attitudes of investigators using it did not fit into the contemporary *Zeitgeist* that had emerged under the powerful impact of Ebbinghaus' and his followers' work. The method of retained members was geared to the study of retention and forgetting of individual items, whereas for Ebbinghaus the study of memory was synonymous with the study of acquisition and retention of associations among items. Theoretical concepts that had been derived from experiments in the Ebbinghaus tradition—such as Ebbinghaus' own general law of associations, Jost's law, Müller and Pilzecker's law of retroactive inhibition, and others—seemed to have no immediate relevance to the processes assumed to be involved in retention of individual items *qua* individual items. Since no alternate theoretical generalizations emerged from work done with the method of retained members, the method gradually lost its power to excite the imagination of psychologists interested in memory and was overshadowed by other experimental paradigms. Its sinking into oblivion was undoubtedly aided by Ebbinghaus' dogmatic verdict that it had only a limited usefulness and that it did not provide as sensitive a measure of retention as did his own method of savings, sentiments that were faithfully echoed by the next generation of writers (e.g., Busemann, 1911; Messer, 1922; Woodworth, 1938).

The method of free recall

Although scattered reports of experiments using the method of retained members appeared from time to time, the method was not an important part of the armamentarium of verbal learning laboratories between the First and Second World Wars. McGeoch's (1942) influential book and its successor (McGeoch and Irion, 1952) did not even mention the method by name. But McGeoch seems to have been the first writer to use the term "free recall" in discussing the experiments by Welch and Burnett (1924) and Raffel (1936). Thereafter, free recall came into more general use (e.g., Postman, Egan, and Davis, 1948; Postman and Jenkins, 1948), being regarded as "a test sensitive to intermediate degrees of associative strength" by Postman, Adams, and Phillips (1955, p. 10).

Under its new name, the method of retained members has gradually become respectable again. This newly found respectability, however, has little historical connection with the more remote past. Rather, it seems to be a consequence of the general broadening of outlook and interests of students of memory as well as of certain methodological and theoretical developments both within and outside the field of verbal learning.

In the first place, several mathematical psychologists found in multitrial free recall what they believed to be a relatively simple learning process whose various behavioral manifestations could be well described by mathematical models (Bush and Mosteller, 1955; Miller and McGill, 1952; Waugh and Smith, 1962). Secondly, other psychologists interested in mental organization (Bousfield, 1953; Bousfield and Cohen, 1953) or associative habits (Jenkins and Russell, 1952) started to use free recall as a vehicle for demonstrating the orderliness of the effects of past verbal experience. This line of work quickly underwent a reversal of its direction of orientation and came to be focused on free recall as influenced by mental organization and pre-existing associations (Bousfield and Cohen, 1956; Bousfield, Cohen, and Whitmarsh, 1958; Cohen and Bousfield, 1956; Gonzalez and Cofer, 1959; Jenkins, Mink, and Russell, 1958; Rothkopf and Coke, 1961). The third impetus to the development of the interest in free recall was provided by the conceptual separation of paired-associate learning into two stages, the response learning stage and the associative stage (Underwood, Runquist, and Schulz, 1959; Underwood and Schulz, 1960). Since response learning was thought—and by some still is thought—to occur independently of specific inter-item associations, free recall naturally suggested itself as a means for probing the response learning stage. Finally, with the general growth of the field of verbal learning and memory, many other experimenters began to employ the free-recall paradigm to study a variety of problems. Their findings have posed challenging problems for the theorists and have pointed to the potential usefulness of the free-recall paradigm in studying processes of memory, Ebbinghaus' peremptory protestations notwithstanding (e.g., Bruner, Miller and Zimmerman, 1955; Dallett, 1963, 1964; Deese, 1957, 1959a, 1960; Ekstrand and Underwood, 1963; Meyer and Miles, 1953; Miller and Selfridge, 1950; Murdock, 1960; Tulving and Thornton, 1959; Waugh, 1961, 1962).

Some Orienting Attitudes and Pretheoretical Assumptions

Free recall and other paradigms

Thus it is that we now find ourselves in possession of a number of reasonably reliable empirical facts gleaned from free-recall experiments and can afford to examine their theoretical implications. Many theoretical problems implicit in free-recall phenomena have their counterparts in the work done with other paradigms. The question therefore arises as to the advisability of breaking the domain of memory into compartments defined in terms of different experimental paradigms. Would it not be more profitable or strategically more advantageous to gloss over the mere methodological differences in data-collection procedures, and instead attempt to delineate the significant empirical facts and search for the comprehensive laws of memory as they apply to all paradigms?

Several closely related reasons argue against such an approach. The conceptual framework of serial and paired-associate learning has become quite formalized. Such a state of affairs readily provides analogies that the theorist can draw upon in his own efforts to speculate about the processes involved in free recall, but quite frequently the fitting of free-recall findings into the existing conceptual framework built around other situations could be achieved only by using Procrustean methods. Given even the remote possibility—and there are some who would argue that the possibility is indeed remote—that the traditional conceptual framework built around other paradigms has only limited generality, the shotgun wedding of free recall to the paradigms that have hitherto dominated the field may hinder rather than advance the progress of our thinking about free-recall phenomena. There are several rather clear-cut cases on record where the findings of free-recall experiments do not readily fit into existing theoretical molds. For instance, Waugh's (1961) finding, antedated to a limited extent by Abbott (1909), that multitrial free-recall performance is identical for constant and variable order input lists, could not be accommodated easily by existing theories that emphasize the importance of repeated contiguity of items for the development of associations between items. Similarly, Battig, Allen, and Jensen's (1965) demonstration that subjects in multitrial free-recall experiments tend to recall newly learned items prior to those items that have been recalled correctly on previous trials seems not to be easily reconcilable with principles according to which the order of recall is a direct function of "strength" of items (e.g., Underwood and Schulz, 1960). Finally, evidence provided by Garner and Whitman (1965) that learning of a part of the list following the learning of the whole list, and by Tulving (1966a) that learning of the whole list following the learning of a part of the list shows no positive transfer, also seems to lie outside the scope of the extant theoretical formulations based on traditional rote-learning paradigms.

For these reasons then, consideration of theoretical implications of free-recall phenomena is better initiated in relative isolation from the traditional conceptual formulations. The risks involved in divorcing ourselves from the dominant *Zeitgeist* are not as great today as they were fifty years ago. Some rapprochement between free-recall theories and, say, serial and paired-associate learning theories is undoubtedly possible even now and hopefully will be extended in the future. But such an alliance will probably require a certain modification (perhaps "liberalization" is the word) of the orienting attitudes that lie at the root of traditional formulations. Battig's assessment of paired-associate learning elsewhere in the present volume may become an important milestone on the road to such liberalization.

Learning in free recall

The minimal free-recall experiment consists of presentation of a series of items to the subject for study in the "input phase" and the measurement of the subject's recall in terms of the number of items he can reproduce in the "output phase" of a single trial. The minimal experiment of this kind will be referred to as a single-trial free-recall experiment. When two or more trials are given, each consisting of a single input phase and a single output phase, and the material presented in input phases is the same over trials, the experiment will be referred to as a multitrial free recall, or a free-recall learning experiment.

Other arrangements of sequences of input and output phases are also possible. A single input phase, for instance, may be followed by two or more output phases, in each of which the subject attempts to recall all presented items (Estes, 1960), or several input phases may be followed by a single output phase (Bousfield and Cohen, 1953; Raffel, 1934). Repeated "cycles," consisting of single input phases followed by two or more output phases, or of several input phases followed by a single output phase, have also been used (Tulving, 1966b).

Typical multitrial free-recall experiments have sometimes been referred to as free learning experiments (Underwood, 1964; Ekstrand and Underwood, 1963). Although this label has the merit of brevity, it implies that in multitrial free-recall experiments "learning" is free. Regardless of the meaning of the term "learning," it is difficult to argue that "learning" is any more "free" in this experimental paradigm than in any other paradigm. What is "free" in both single-trial and multitrial free-recall tasks is the order in which the subject may recall the remembered items. For this reason, the name "multitrial free recall," or perhaps "free-recall learning," seems to be preferable to the name "free learning."

The basic unit of analysis of the subject's performance in free-recall experiments is a single item. Exactly what constitutes a single item is not always easy to determine (Deese, 1961); some of the problems arising in this connection will be considered in a subsequent section. For the time being, we define a single item as an item that can be identified by the subject in terms of a highly integrated and invariant response, a response that could be readily elicited by stimuli other than the specific experimental stimulus (Mandler, 1966b), or a response that can be readily completed if a substantial part of it is given. Thus, common words with which the subject is familiar before entering the experiment are single items. A sequence of letters such as *PRAJ*, however, does not constitute a single item, even if the experimenter defines it so (Garner and Whitman, 1965). Memorization of sets of such verbal units involves both response integration and response selection. As a consequence, theoretical analyses of the processes underlying successful performance in free recall of such units are more complicated than those concerned with free recall of single items. In this paper we will limit our discussion primarily to free recall of lists of single items.

A single item is always "learned" when it is first presented, regardless of whether it occurs alone or in a series of other items, in the sense that the probability of its recall increases from a value near zero immediately prior to the presentation to unity immediately after the presentation (Tulving, 1964). "Learning" of individual items is equivalent to their perception, since the operations involved in testing learning and perception in this case are identical.

Thus, to talk about learning or degrees of learning of single items is meaningless, since at least at the present time learning cannot be measured independently of recall. Measures of recall, however, depend on many factors other than the hypothesized degree of learning, regardless of the temporal point at which recall is tested. For this reason, the question sometimes raised whether *learning* of a single item is all-or-none or incremental can never be answered. Single-trial free-recall experiments are experiments concerned with *remembering of perceived items* and not with learning of items.

The term "learning," however, does serve as a useful and nonredundant descriptive term in multitrial free-recall experiments. In this case it refers to the more-or-less systematic increase in the number of items recalled over successive trials or, more specifically, to the slope of the function relating number of items recalled to the ordinal number of the trial (Murdock, 1960). Whenever this slope is positive, learning can be said to have occurred over trials. The rate of increase in trial-by-trial recall can be specified quantitatively, and rates obtained under different experimental conditions can be directly compared. Since this rate, or the slope of the "learning curve," is a measure *derived* from measures of recall, it need not be identical with these measures. Degree of learning can be the same in different experiments or under different experimental conditions, even if recall scores on corresponding trials are not the same.

Thus, when "learning" is used as a descriptive term in multitrial experiments, it serves as a shorthand expression for the fact that recall increases over trials. To ask why learning occurs, or to ask what processes are involved in learning (Tulving, 1962, 1964), means to ask why recall scores increase from trial to trial, or what processes are involved in such an increase. Learning, defined in terms of increments in the number of items recalled as a function of trials, is always "incremental" by definition.

Measurement of free recall: Units of analysis

While specification of independent variables in free-recall experiments poses no problems different from those that arise in connection with many other verbal learning and memory tasks, definition and measurement of response variables do. What does "amount of material recalled" mean and how is it to be measured?

The long-standing tradition has been to count the number of experimentally defined single items. For the purpose of initial descriptions of relations between independent and dependent variables, specification of both input and output material in terms of units such as trigrams, two-digit numbers, or words may do quite well. But when we begin to speculate about underlying processes, the procedure becomes questionable, its usefulness depending on the nature of the independent variables whose effects are being investigated. The difficulty arises because of a possible discrepancy between single items defined as such by the experimenter and those that are handled as single items by the memory system. Let us designate these two kinds of units as nominal and functional units, or as E-units and S-units.

Ebbinghaus pointed out that individual elements may be perceived either as a collection of unrelated elements or as parts of a whole. When the elements belong to a whole, associations between them can be formed very rapidly. Ebbinghaus also noted that it is not the number of letters or the number of syllables that determines the ease of learning, but rather the number or kind of ideas these elements represent: a series of monosyllabic words is learned as readily as a series of dissyllabic words. Similar observations have been made by Holmes (1934).

As we noted earlier, experimenters who used the method of retained members or the method of free recall usually took pains to select "unrelated" items for inclusion in their lists. But such attempts were seldom successful. Smith (1896), for instance, whose subjects had to recall nonsense syllables, reported that most subjects showed a tendency to associate "foreign" ideas with the syllables or "to make the syllables into intelligible phrases." This tendency was "a very troublesome feature from beginning to end" with one individual. Pohlmann (1906) found that subjects recalling randomly ordered series of consonants or sets of two-digit numbers tended to output the consonants in the order in which they occurred in the alphabet or as groups of items of similar pronunciation. One group might consist of consonants *S*, *N*, *M*, and *F*, another of *D*, *G*, and *B*, still another of *H* and *K*, for which the German pronunciation is similar. Numbers were also frequently recalled in pairs such as 26 and 27, 42 and 84, 23 and 46.

Other early investigators even objected to the name of the method of retained members, since it conveys the impression that the observer is called upon to remember isolated items, while "in fact he much more frequently endeavors to remember the total impression and to retain the items only as component parts of the total impression" (Meumann, 1913). Meumann also suggested that "learning is not a mere matter of the number of elements but of the number of independent memorial units. . . . The only things which we remember are wholes; and particular things are remembered only as parts of unitary wholes" (p. 291).

Bowers, whose subjects recalled a list of 135 words, could also "not help being impressed by the frequency with which certain words were found paired together in the reproduction list. It would appear that the phrase 'words isolated from context' requires modification, for there is evidently a striving after artificial context when a natural one is not provided" (Bowers, 1931, p. 276). To eliminate the effect of "common associations" and to study the retention of words isolated from the context, Bowers constructed a matrix similar to the one named by Tulving "intersubject recall matrix," tabulated the frequencies with which a given word was followed by another word in recall protocols, and excluded such word-pairs from the subjects' recall scores. He noted that this method of "purifying" the data, however, did not tap "private associations" peculiar to the individual, which, he said, cannot be detected objectively. "Even introspection will frequently be at fault in uncovering the most subtle private associations" (Bowers, 1931, p. 277).

Such observations, repeated by many later investigators as well, strongly suggest that S-units do not always correspond to E-units, that E-units are frequently perceived and remembered as parts of larger wholes or higher-order units, and that it is difficult to discover what the S-units are.

The possible discrepancy between nominal and functional units does not affect the interpretation of data from certain types of experiments as much as it does in others. For instance, if the independent variable of interest is the length of the retention interval, and other conditions are held constant, it is reasonable enough to assume that differences in recall measured in terms of E-units also reflect differences in recall of S-units. It would be interesting, of course, to know whether the loss in recall scores over a retention interval is attributable to the loss of some intact units or to the disintegration of larger S-units, but such information would not affect the conclusion that less material can be recalled following a long retention interval than after a short interval.

The difficulty of interpretation is much greater in experiments where the independent variable is a task variable such as length of the list or the nature of the material. If one compares the recall of n meaningful words with the recall of n meaningless trigrams, one does not know how many S-units were presented in each case and how many were recalled. If subjects recall more words than trigrams, the conclusion that words are remembered better than trigrams ignores the possibility that recall might be identical if it were measured in terms of the number of S-units.

Miller and Selfridge (1950), in a well known experiment, showed that subjects can recall more words from long lists of higher order statistical approximations than from lists of lower order approximations. Higher order lists presumably contain fewer but larger S-units than lower order lists. Even if subjects recalled the same number of S-units from both kinds of lists, the number of correct word-responses would vary directly with the order of approximation.

In an experiment patterned after the Miller and Selfridge study, Tulving and Patkau (1962) examined subjects' recall in terms of "adopted chunks." An adopted chunk was defined as an uninterrupted sequence of words in the recall protocols corresponding to a similar sequence in the input list. Thus, given an input list of words A, B, C, D . . . Z, in this order, and a recall protocol of words K, L, M, A, G, H, in this order, the subject was said to have recalled six words and three adopted chunks—*KLM*, *A*, and *GH*. The results of the analyses showed that neither the Thorndike-Lorge word frequency nor the interaction between word frequency and order of approximation were significant sources of variance in the recall of adopted chunks, although both were highly significant when the response measure was the number of words recalled. McNulty (1966) also found that the number of "words" recalled at the first, third, and text order of approximation, as defined by Miller, Bruner, and Postman (1954), varied enormously with the order of approximation but that the number of adopted chunks remained invariant over the three orders.

The "adopted chunk," incidentally, corresponds to what Rozov (1964) has called an "original cluster." Original clusters represent the retention of the contiguity relations among items under conditions where no such retention is experimentally required—that is, in free recall. Such "contiguity transfer," according to Rozov, is not accidental, but due to the fact that the items presented together form some larger unit in recall. Rozov presented 30 words in six topic categories in a scrambled order in a single input phase and tested the subjects' recall in two successive output phases. An interesting finding was the fact that in all cases where original clusters found in the first output phase were broken up in the second phase, their component items were transferred into their proper topic clusters.

The invariance in the number of recalled S-units observed by Tulving and Patkau (1962) and McNulty (1966) is also apparent in the data reported by Jenkins, Mink, and Russell (1958). They used four lists of 12 associatively related pairs of words, lists varying in the strength of intrapair associations. Number of recalled words varied directly with the strength of associations over a range of 14.5 to 19.1. However, counting a related pair of words recalled in immediately adjacent output positions as one rather than two units, we find that the number of units recalled for the four lists, ordered in terms of increasing associative strength, becomes 13.4, 14.5, 13.2, 14.0, indicating no systematic differences among lists.

All these experiments suggest that the kinds of results we get and have to interpret depend very much on the kinds of units of analysis we use. Variation in the number of E-units recalled as a function of order of approximation or of inter-item associative strength would require an explanation different from that required by the finding of invariance in the number of S-units. The S-unit invariance illustrates limited recall capacity of human subjects which is independent of the number of presented units, provided that the total time of presentation is held constant (cf. Murdock, 1960).

One of the important problems facing experimenters and theorists interested in free-recall—as well as those concerned with other memory tasks—lies in the specification of the functional units of material that are remembered and recalled. We may be ignorant as to the exact S-units in any given situation, and we may have to temporize by counting such easily identifiable units as trigrams or words, but sooner or later we have to come to grips with the S-units in an objective fashion. In the long run, nothing will be gained by pretending that gaps between sequences of printed letters or between sequences of spoken phonemes define the units of information processed by the human memory system.

Stevens (1951) has suggested that there is only one problem in psychophysics, or indeed in all psychology—

that of the definition of the stimulus. In free recall, the problem of the definition of the response is equally vital.

Single-Trial Free-Recall Phenomena

Apart from the evidence on the relation between various characteristics of the material and the amount recalled, there are two major facts that have emerged from single-trial free-recall experiments in which single items are used as units of analysis. The first has to do with the relation between the ordinal position of an item in the input list and its probability of recall, while the second is defined by the relation between the amount of presentation time and the amount of material recalled. We will consider these two phenomena in turn and examine their theoretical implications.

The serial position curve: Evidence for two retrieval systems

If the input material consists of randomly ordered familiar items, if the presentation rate is not greater than two seconds per item, and if subjects have had some prior practice in the task, the function relating probability of recall of an item to its position in the input list, the serial position curve, is usually characterized "by a steep . . . primacy effect extending over the first three or four words in the list, an S shaped recency effect extending over the last eight words in the list, and a horizontal asymptote spanning the primacy and recency effects" (Murdock, 1962, p. 488). Murdock's experiment represents the most extensive investigation of the serial position curve under the conditions stated, but data similar to his have been reported by others as well (e.g., Deese, 1957; Deese and Kaufman, 1957; Postman and Phillips, 1965).

The primacy effect has so far eluded explanation. Two main lines of argument have been advanced to account for it, but neither has much unequivocal support. The first postulates that items from early input positions are recalled more readily than items from the middle input positions because the former are subject to smaller amounts of intraserial proactive inhibition. The second states that subjects tend to rehearse earlier input items while being exposed to middle items, with the consequence that earlier items are strengthened at the expense of middle items. Both hypotheses share the assumption that earlier items are in some sense "stronger" at the beginning of the output phase than are the middle items; this is a simple restatement of the fact that there is a primacy effect. No incisive experimental data to elucidate the processes involved have been reported so far.

The recency effect is much better understood. Most explanations that have been advanced constitute elaborations of the suggestion made by Robinson and Brown (1926) who, in discussing the pronounced recency effect reported by Welch and Burnett (1924), proposed that it was attributable to the shorter time between learning and recall of last items. Several investigators have demonstrated sizable negative correlations between probability of recall and output positions of individual items (Bousfield, Cohen, and Silva, 1956; Deese, 1957; Deese and Kaufman, 1957; Waugh and Norman, 1965). There are exceptions to this general rule in special cases (e.g., Green and Harding, 1962), but under the typical conditions of free recall of series of homogeneous items, the correlation between output order and probability of recall does imply that recall of terminal input items is high because they are recalled earlier than other items. Murdock's (1962) finding that the recency effect is independent of the rate of presentation further clarifies the picture by suggesting that it is not the length of the intratrial retention interval, but rather the number of items presented and recalled between the presentation and attempted recall of a given item, that is of critical importance in determining the probability of recall of the item.

The results of important experiments by Murdock (1962), Postman and Phillips (1965), Glanzer and Cunitz (1966), and Glanzer and Meinzer (1967) provide strong support for the hypothesis that at least two different recall mechanisms are involved in free recall. Murdock, as we have noted, found that the recency effect was independent of rate of presentation of items. He also found that it was independent of list length. Both of these variables, however, did affect recall of items from input positions preceding the range of the recency effect. Postman and Phillips demonstrated that the typical pronounced recency effects for lists of 10, 20, and 30 words were washed out when recall was delayed for 30 seconds, the interpolated interval being filled with activity assumed to prevent rehearsal. The rest of each serial position curve, however, was only little affected. Very similar data were obtained by Glanzer and Cunitz, who used lists of 15 words and delays of 10 and 30 seconds. The loss of recall over the interval of 30 seconds in the Postman and Phillips' experiment was approximately constant at two words for all list lengths. This finding is in good agreement with the hypotheses that the loss of recall involves primarily items from the terminal input positions and that the recency effect is independent of list length. Finally, the experiment by Glanzer and Meinzer demonstrated that overt repetition of words during the intervals between successive words in the input phase reduced the recall of words from the prerecency range of the serial position curve but did not affect the recall of terminal items.

All these findings, in showing that certain experimental variables affect one part of the serial position curve while leaving another part invariant, strongly suggest that the curve reflects the operation of at least two types of recall mechanisms. Waugh and Norman (1965) refer to these two mechanisms as primary and secondary memory, following James' (1890) distinction, and identify them with different storage systems. Glanzer and Cunitz also discuss their data in terms of two kinds of storage mechanisms, the short-term and the long-term storage. We will briefly consider the Waugh and Norman model as representative of the two-factor theories of memory having something to say about single-trial free recall.

According to the Waugh and Norman model, as an item is perceived by the subject, it enters into primary memory, which has a very limited capacity. Items in primary memory are displaced by other incoming items unless they are rehearsed. As long as they are rehearsed, they remain in primary memory and may be transferred into secondary memory, which has a much larger capacity. The two storage systems are independent, in the sense that the contents of one at a given time have no bearing on the contents of the other at the same time, and yet not mutually exclusive, in the sense that any given item can be in either or both of the two systems at a given time.

An important assumption Waugh

and Norman make is that any item in either system can always be recalled by the subject. They thus equate retention with recall, or storage with retrieval. The recency effect is determined by the contents of both the primary and secondary memory, while the asymptotic level of the serial position curve reflects the amount of material stored in secondary memory. The contents of primary memory can be estimated from the observed values of probability of recall of late input items and of the asymptote of the serial position curve, while the contents of the secondary memory can be inferred from the prerecency range of the serial position curve.

The easy retrieval of recently perceived items has been referred to by Waugh and Norman as the "echo box" phenomenon, "echo box" representing the primary memory storage. The existence of the "echo box" was known to early students of memory. Pohlmann (1906), for instance, pointed out that under the method of retained members recall of last items in the series may occur by virtue of their "*flüchtiges Nachklingen*"—fleeting reverberation —and recommended that an interval of a few seconds always be interpolated between presentation and recall to reduce the contamination of true memory effects by such *Nachklingen*.

Although the distinction between primary and secondary memory is heuristically useful, the identification of the two types of memory with different storage systems may not constitute the most promising stepping stone to further theoretical analyses. The very term "storage" or "store" implies a location of some sort in the nervous system in which information is held for future use. Do the two types of storage then imply two different locations where the engrams Lashley (1950) never found are tucked away?

To be sure, all models of memory represent only "as if" pictures of the human mind or the human nervous system, and the terms we use are just figures of speech that help us to relate something we know to something we do not quite understand. But, as Wechsler (1963) has recently reminded us again, the metaphors and analogies of which students of memory are so fond are frequently accepted as literal explanations. When this happens, the analogies may well stifle further inquiry. McGeoch's remark, made more than 35 years ago, that experimentalists have generally preferred to content themselves with analogies, such as that of the phonograph record, rather than to confront the real problem is probably applicable today, if we substitute "conditioning paradigm" or "electronic computer" for "phonograph record." Where shall we go next, once we have accepted the notion of two storage systems whose capacity and contents determine how much the subject can recall?

I prefer the view that all input information is stored in the same unitary storage system, whatever its nature, and that differences in recall of early, middle, and late input items reflect primarily differences in the accessibility of these items (cf. Miller, Galanter, and Pribram, 1960). Late input items may be retrieved more easily because certain kinds of additional auxiliary information, stored with each item at the time of presentation, are available for these items and not available for items perceived earlier. The acoustic trace of an item the subject hears—or the acoustic trace of an item the subject sees and recodes acoustically (Sperling, 1963)—may be one kind of such auxiliary information that might serve as a retrieval cue for the item. Such an acoustic trace may rapidly decay or be blotted out by other traces, even if its loss does not affect the rest of the information stored at the time of input. The item for which the acoustic trace is lost may still be potentially retrievable through other, less powerful but more permanent retrieval cues. Serial position or temporal dating of items in the input phase may constitute another kind of auxiliary information which can serve as a retrieval cue. Such information may be greater for early and late input items than for the middle ones and may be initially more powerful and yet decay more rapidly than other retrieval cues.

Regardless of the nature of retrieval cues affecting the recall of individual items, such cues must be postulated, since the alternative formulation of the organism or his homunculus "searching" for lost objects in memory (James, 1890) will lead us down the path of infinite regress. Until we know more about the retrieval mechanisms and retrieval cues, it is at best unnecessary and at worst misleading to identify primary and secondary memory with different storage systems. All we need to say is that at least two different types of retrieval systems must at this time be postulated. Eventually we may have not only primary and secondary memory, but as many different "types" of memory as we have found different types of retrieval cues.

Total presentation time

One of the most easily demonstrated facts about single-trial free recall is the direct dependence of the amount of material recalled on the amount presented. "Immediate" memory is limited in free recall as it is in many other tasks, but the limit, at least when measured in E-units, depends on the amount of material presented.

When items are presented successively at a fixed rate, the amount of material presented covaries directly with the total presentation time. Murdock (1960), on the basis of a series of experiments, suggested that this total presentation time is the critical variable responsible for the covariation between the amount of material presented and recalled. This general hypothesis was supplemented by a more specific one according to which the number of items recalled is a linear function of the total presentation time, with the intercept constant in the linear equation $R = kt + m$ representing the value of the memory span and the slope constant representing the number of items per unit of time that can be recalled over and above the span.

The boundary conditions of Murdock's two hypotheses have not yet been determined. Although a number of experimental reports exist in the literature providing data on the relation between amount of material presented and amount recalled, few experiments, apart from that of Murdock, have explored the relation over a sufficiently wide range of list lengths or presentation times to permit a fair check on the tenability of the linearity hypothesis. One of the experiments that comes closest to the necessary standards was done by Deese (1960). Deese used lists of 12, 25, 50, and 100 words and combined list length orthogonally with six different frequency-of-occurrence levels of words. For the two highest levels of word frequency, the relation between number of words recalled and list length was approximately linear, but for the lower four levels of frequency linearity did not hold. Cohen (1963b) also reported nonlinearity for lists of 20 words presented at the rate of 3 seconds and 10.5 seconds per word.

In some other experiments both the presentation-time and the linearity hypotheses run into difficulty. In one of Murdock's own experiments (1962), for instance, recall was no greater for 20-word and 40-word lists presented for 40 seconds than for 15-word and 30-word lists presented for a total time of 30 seconds.

The presentation-time hypothesis also breaks down when subjects have to recall names of objects shown in a simultaneous display. Michel (1923) reported a series of experiments in which immediate recall was studied as a function of number of objects presented—the number varying from 5 to 30—and the total presentation time, varying over a range from 5 to 150 seconds. Number of objects recalled was a direct function of the duration of display, but when the presentation time was held constant, the number of items recalled still increased systematically with the number of objects presented. One of the interesting findings reported by Michel was that even with very short exposure times recall was substantial. Thus, for instance, mean recall of 10 objects presented for a total duration of 2 seconds was 5.6, and mean recall of 25 objects presented for 5 seconds was 8.6. Hence Michel's method of a simultaneous display of a set of small objects recommends itself for the study of memory under conditions where reading-in time is to be minimized.

Despite the uncertainty as to the exact boundary conditions of the two hypotheses, "immediate" recall of a list of words does seem to be generally a positive function of the total amount of presentation time. Nothing much can be said about this phenomenon by way of theoretical speculation. Deese's (1960) suggestion, however, that the proportionality between recall over and above recall from the primary memory on the one hand, and the total presentation time on the other, can be explained in terms of interword associations constitutes a promising starting point for the search of underlying processes, provided that we define "interword associations" very broadly. The probability that two or more E-units are stored by the subject as a single S-unit is likely to increase with list length and with the total time the subject has available for the discovery or fabrication of relevant interitem relations. If groups of related items can be handled by the memory system as single S-units, longer lists and longer presentation times would be expected to yield larger recall scores if recall is measured in terms of E-units.

Secondary Organization in Free Recall

We have already noted that subjects usually produce their responses in recall in an order different from the presentation order. Discrepancies between input and output orders cannot occur in serial learning tasks, and when they occur in paired-associate tasks, they are of no theoretical significance since the output order is determined by the experimenter. In free recall, however, the output order is free to vary, and discrepancies between input and output orders, under conditions where output orders show systematic consistencies, provide evidence for organizational processes in remembering.

Organization defined in the weak sense refers to consistent discrepancies between input and output orders that are independent of the subjects' prior familiarity with a set of input items. We will label this type of organization "primary organization." The recency effect, for instance, being related to the subjects' tendency to recall terminal input items from a homogeneous list before recall of other items, regardless of the characteristics of these items, is one of the manifestations of primary organization.

In this section we will discuss some of the issues that emerge from observations of organization defined in the strong sense. Such organization occurs when the output order of items is governed by semantic or phonetic relations among items or by the subjects' prior, extra-experimental or intra-experimental acquaintance with the items constituting a list. We will label this type of organization "secondary organization." It can be measured in a variety of ways and can be related to both independent variables and other dependent variables, thus illuminating some of the underlying processes in free recall.

Clustering

Two types of measures of secondary organization as a response variable have so far been developed. The first type of measures can be obtained in situations where the total set of L items is constructed and conceived by the experimenter to consist of two or more mutually exclusive subsets of items. Each item within a given subset is assumed to be more "similar" to other items within the same subset than it is to any other item in other subsets. Subsets have been defined in terms of belongingness of items to a conceptual category, associative relations among items, parts of speech, etc. The items in the total list are presented in a random or quasi-random order, and organization is said to have occurred when items from a subset are recalled in immediately adjacent output positions more frequently than one would expect by chance.

Secondary organization measured in this manner has been referred to as clustering. A long line of studies by Bousfield and his associates, beginning with the experiments by Bousfield (1953) and Bousfield and Cohen (1953), as well as those by many other investigators (e.g., Cofer, 1959; Cofer, 1965; Dallett, 1964; Gonzalez and Cofer, 1959; Jenkins, Mink, and Russell, 1958; Jenkins and Russell, 1952; Postman, Adams and Bohm, 1956; Wicklund, Palermo and Jenkins, 1965) have been concerned with the phenomenon of clustering. Clustering has been shown to be a function of many variables, such as number of repetitions of the material, strength of associative relations among words, number of subsets, prior practice, instructions, and developmental level of subjects. In many cases, although not invariably, it has been shown to be directly related to the number of words recalled.

An obvious advantage of clustering measures is that they can be calculated for recall obtained in a single output phase, although Rozov (1964) has suggested that the practice of giving subjects only single recall trials may have prevented investigators from obtaining greater insight into the processes underlying clustering. Furthermore, if the single output phase follows a single input phase in immediate succession, subjects tend to recall the late input items first, regardless of the specific meaning of these items or their semantic relations to other items in the list. This primary organization probably interferes with, or at least attenuates, secondary organization. Bousfield (1953) and Bousfield and Cohen (1953) have shown that the function relating the degree of clustering to successive deciles of words produced in the output phase has the shape of an inverted U, clustering being strongest in the middle of the output period. Bousfield and Cohen (1953) have proposed a rather elaborate explanation for the low degree of clustering obtained at the beginning of the output phase, but the phenomenon may be at least partly attributable to the priority of retrieval of terminal input items, i.e., to primary organization. The fact that the degree of clustering is maximum in the first few deciles when recall is tested following five immediately successive input phases (Bousfield and Cohen, 1953) is consonant with this interpretation, since the recency effect is probably less pronounced under these conditions.

The main disadvantage of the clustering measures is related to the fact that they only tap organization of the kind that the experimenter is looking for and hence tend to underestimate the extent to which organization has occurred in a given situation. To take an extreme case, the failure to find clustering of words according to their form classes (Cofer and Bruce, 1965) does not mean that recall is unorganized; it only means that form class normally does not provide a basis of organization. Even when input material consists of words of different conceptual categories and there are good reasons to expect that secondary organization should be governed by the belongingness of words to such categories, S-units consisting of words from disparate categories may be formed. These units are not detected in a typical clustering analysis.

Subjective organization

The second type of measure of secondary organization, referred to as subjective organization, requires data from more than a single output phase, but it does not require that the experimenter know in advance of the experiment what items are to be grouped together. It is, therefore, applicable to any set of items.

Measures of subjective organization are defined in terms of the consistency of output orders, either for a single subject recalling the same material in two or more output phases or for a group of subjects recalling the same material in at least one output phase. When two or more items occur in close temporal contiguity in different output phases, they can be thought to represent elements of a larger S-unit which is being processed as a unit.

I have described a method for the measurement of subjective organization based on certain notions derived from information theory (Tulving, 1962). The measure of subjective organization (SO) is positively correlated with the number of words recalled, both across subjects with trials held constant and across trials with subjects held constant (Tulving, 1962, 1964).

A measure closely related to that of SO, labelled ITR (intertrial repetitions) has been used by Bousfield, Puff, and Cowan (1964). The subjects in their experiment had to memorize a list of ten words having zero interitem associative strength according to the Russell and Jenkins (1954) norms. Bousfield, Puff, and Cowan's main findings were that ITR increased over trials and that there was a significant agreement among subjects in the final order in which they recalled the words, this order being unrelated to the frequency of usage of the words.

Other methods of quantifying subjective organization have been described by Seibel (1965) and Ehrlich (1965). Seibel has used what he calls "the study sheet paradigm" for objectively defining the subjective clusters in recall. He has found that the degree of subjective clustering increases with learning and that it is highly correlated with the number of words recalled. Ehrlich has measured the development of organization over trials in terms of interitem distances (number of items intervening between any two items) in successive output phases, with the interitem distances in the final output period constituting the reference level. He has obtained high positive correlations between his "coefficient of structure" and the number of items recalled by individual subjects, the correlation increasing with the length of the list, and has found that, over successive trials, the number of correct responses is a linear function of the amount of structure, which in turn is a logarithmic function of the learning time.

Nature of secondary organization

Ordering of recalled items in terms of their associative relations, membership in obvious superordinate categories, phonetic or letter-structure similarity, degree of familiarity, and other such characteristics (Ehrlich, 1965; Tulving, 1962, 1965), a phenomenon that we have referred to by the generic term of secondary organization, is a ubiquitous phenomenon in free recall. It can always be found in any experiment when one looks for it. From the earliest days of experimental study of memory, it has intruded into experiments designed to examine the recall of individual items *qua* individual items, and it has been noted in many more recent free-recall experiments that were not directly concerned with the problem of organization (e.g., Kintsch and Morris, 1965; Miller and McGill, 1952; Waugh, 1961). How is this secondary organization to be explained or interpreted?

Bousfield, Steward, and Cowan (1964) have posed the question with respect to clustering, but it applies equally well to subjective organization: "There is the question of whether organization of verbal responses can be explained in terms of relatively simple associative connections between words, or whether it is necessary to invoke an additional principle such as superordination" (Bousfield, Steward, and Cowan, 1964, p. 206). The principle of superordination as an explanation of clustering had been put forth earlier by Bousfield and his associates (Bousfield, 1953; Bousfield and Cohen, 1953, 1955). The results of the Bousfield, Steward, and Cowan (1964) experiment again supported the principle of superordination—as did the results of the study of Bousfield and Puff (1964)—permitting the authors to conclude that response-produced cues function mediationally to promote clustering of words. They were careful to point out, however, that their findings did not in any way negate the "potency of direct associative connections for determining clustering."

Postman, in a brief review of studies of clustering, also has distinguished between associative and categorical clustering and has pointed out:

> The relationship between the two types of clustering remains uncertain. Associative clustering appears to reflect directly the associative structure of the words in the list, and no recourse to a mediational process is compelled by the data. On the other hand, at least some forms of category clustering strongly suggest mediation by cue-producing responses. The question now is whether it is necessary to assume two distinct types of clustering, and if it is, what their relative priorities are in determining sequential order in recall (Postman, 1964, p. 180).

A rather detailed discussion of associative and categorical bases for clustering has been presented by Cofer in two papers (1965, 1966). After considering a great deal of evidence relevant to the problem of mechanisms underlying clustering, Cofer concluded, "It now seems to me, and this represents a change of heart, that such a contrast is neither useful nor heuristic. In free-recall . . . subjects will use either or both of these bases to accomplish their recalls and will find ways to organize recalls even though the experimenter has not provided means in the list he presents" (Cofer, 1965, p. 271). In a subsequent paper, however, Cofer apparently found the distinction between associative and mediational clustering useful again and concluded that in some situations, such as Bousfield's four-category lists, associations are the dominant factor in clustering, while in some others the operation of a coding factor (in addition to association) seems highly plausible (Cofer, 1966).

In my opinion, the attempts to distinguish between associative and mediational mechanisms of clustering, even if only to assess their relative effects in various learning processes (Bousfield and Puff, 1964), are futile in the present state of the art. The first difficulty lies in the concept of association itself. It is a descriptive concept, not an explanatory one. Association is a name for the fact that one event leads to another, and as such it carries no theoretical implications (McGeoch, 1942). If the association between two items, A and B, is strong, the probability of A given B, or vice versa, is high. If the probability of A given B, or vice versa, is high, the association between A and B can be said to be strong. Thus, the demonstration that word B is given frequently as a free-association response to word A, and that the two words are recalled together when they are both members of a list, is a demonstration that similar events occur in different situations. It suggests that similar processes may be involved in both cases, but associations observed in free-association tests do not *explain* associations in free-recall tasks any more than the latter associations explain the former. Association between words A and B, whether observed in a free-association test or in a free-recall task, is a fact that remains to be explained.

The second difficulty in seeking the explanation of clustering in the distinction between associative and mediational mechanisms is related to the simple fact that any two words that can be shown to be associatively related—or any other two words, for that matter—are always members of some superordinate category that subsumes both. Thus, two or more words may function as a more or less intact unit either because of the direct associative relations among them, even if we do not know what the underlying mechanism is, or because of some indirect association (cf. Marshall and Cofer, 1963) indicating membership in a common category. To measure the strength of the direct association between two words, to compare it with the strength of the indirect association between the same words, and then to conclude that recall of the words as a unit is determined by direct associations or indirect associations in degrees reflected in the two strength measures would not serve any useful purpose. The conclusion one would draw from such a comparison would be no better than are the measures of strength of associations, and even if the measures were "perfect," we would still have only information about the correlation between events in two situations—free or restricted association tests and free-recall tasks. Such a correlation requires explanation; it does not constitute one.

The third difficulty in current attempts to explain the mechanisms of clustering is related to a large variety of apparent bases of organization observed in free-recall experiments. At Toronto, we have been experimenting with the application of a modified method of McQuitty's (1963) rank order typal analysis to the data from various free-recall experiments, with a view to extracting higher order S-units (Tulving, 1964, 1966a) from these data. An S-unit, most rigidly defined (although we have sometimes relaxed the criterion somewhat) is a subset consisting of two or more list items each of which is recalled more frequently in an output position immediately adjacent to other items constituting the S-unit than it is recalled in output positions immediately adjacent to items from other S-units. Some of the S-units gleaned from the data make good sense, while others do not. S-units such as *WARSHIP*, *MASSACRE*, and *SHRAPNEL; RUBBER*, *IRON*, and *SAND; AGREEMENT* and *SETTLE; NAME* and *CLIFF; JEWEL* and *JELLY; BALAP*, *VOLVAP*, and *ZUMAP*, all obtained from experiments using "unrelated" words, can be "explained" on various obvious grounds, as can S-units from a Bousfield type categorized list used in an unpublished study by Dronsejko (1963), units such as *BUGLE*, *HORN*, and *TRUMPET; BASS* and *VIOLA; BEE*, *HORNET*, and *WASP; CAVE* and *CAVERN; MOUNTAIN* and *VOLCANO; GUN*, *SPEAR* and *STICK*. But what are we to do with S-units such as *PEACE* and *SURVEY; CONQUEST*, *ORNAMENT*, and *DEMAND; SPOON* and *CHIMNEY; LATUK* and *TAROP*, that have also emerged from the data? Are these "determined" by direct or indirect associations?

Attempts to classify bases of organization are reminiscent of attempts to classify free associations into different types. After reviewing some of the early efforts at such classification, Ebbinghaus concluded that the results had been disappointing, partly because the classifying schemes that had been proposed were based on grammatical and logical rather than psychological considerations, and partly because the classification of many associations into disparate categories in most cases would not be possible anyway. If "coffee" and "tea" are associated, Ebbinghaus asked, is it because of the similarity of the circumstances under which coffee and tea are consumed, because of similarity of their effects, because of their belongingness to a class of potable things, or because of the frequency with which the two words have been seen or heard together in shop windows, on menus in restaurants, and in price lists? He concluded that all of these factors probably play some role in determining the association and that the assignment of an association between the two words to a single exclusive category would be misleading.

More recently, Deese (1962) has also rejected attempts to classify associations. Instead he defines the associative meaning of a word in terms of the total distribution of associative responses to that word. It would seem to be reasonable to suggest that secondary organization in free recall is at least partly governed by associative meanings of words constituting a list, but this statement again would only explicate a correlation between free-association and free-recall phenomena. Besides, it is conceivable that some of the "unrelated" words that are organized into an S-unit may never show any associative overlap in free-association norms, unitization being determined by the total intralist verbal context and accidental contiguity relations among words in input and output phases. To the extent that verbal contexts differ in free-association tests and in free-recall tasks, the correlation between the strengths of associations in these two situations will probably always fall far short of unity and the extent of the involvement of the associative meaning of words in determining secondary organization will remain uncertain. Bousfield, Puff, and Cowan's (1964) finding of subjective organization of words having zero interitem associative strength is a case in point.

Instead of attempting to explain the nature of the mechanism of secondary organization in terms of extra-experimental correlations, it seems more profitable at this time to start with the fact that certain words are more likely to be formed into clusters or S-units than others and to ask what *intra-experimental* conditions determine the ease with which such S-units are formed. For instance, would it be necessary for two prospective members of an S-unit to be presented in close temporal contiguity in the input list for unitization to occur? Does the probability that two words such as "man" and "woman" become members of an S-unit vary directly with the intraserial distance between them in input and output phases? If so, what is the effective range of distances over which possible unitization of the two words remains effective? If intraserial distance is an important variable, is it to be specified in terms of E-units or S-units? Would the subject have to be able to remember that "man" occurred earlier in the list to bring it into the S-unit with "woman" shown later in the list, or does the presentation of "woman" later in the list somehow resuscitate the trace of "man" even if the latter would not be recallable by itself?

Very little work relevant to these problems has been done, although experiments comparing the effectiveness of blocked and random presentation of word lists categorized on the basis of restricted word-association norms (Bourne and Parker, 1964; Cofer, Bruce, and Reicher, 1966; Dallett, 1964) or on the basis of free-recall data (Tulving, 1965) probably constitute steps in the right direction.

Repetition Effects in Multitrial Free Recall

The limited capacity of immediate memory is one of the basic puzzles confronting the student of mental processes; the increase in recall as a function of practice is another. How should learning—increments in recall over successive trials—be conceptualized?

Serial and paired-associate learning usually has been explained to be a consequence of strengthening of associations between the members of the series or individual pairs. As we have already noted, however, a statement that the strength of association increases over trials does not explain learning. Strength of associations cannot be measured independently of the frequency with which a response follows another response, and thus learning and strengthening of associations are inferred from the same data. The statement that associations in serial and paired-associate lists are strengthened is synonymous with the statement that learning occurs—it does not explain learning.

The situation is different in multitrial free-recall experiments. Learning is inferred from the positive slope of the function relating the number of items recalled to the ordinal number of the trial. Strengthening of associations among specific items, on the other hand, can be inferred from the increased probability that one item follows another item in recall. These two events, learning and strengthening of associations, are logically independent in free-recall tasks. Strengthening of associations can occur even if there is no learning, and learning can occur in the absence of strengthening of interitem associations. "Statistical" subjects in free-recall experiments can recall more and more items over trials independently of strengthening of interitem associations, a feat they cannot accomplish in serial and paired-associate learning experiments. If it is observed, therefore, that free-recall learning in human subjects is accompanied by strengthening of specific interitem associations, the fact has psychological implications of considerable importance.

Relation between secondary organization and learning

To the extent that measures of secondary organization reflect the strength of interitem associations, the findings of increased secondary organization over trials in free-recall experiments (Bousfield, Berkowitz, and Whitmarsh, 1959; Bousfield, Puff, and Cowan, 1964; Ehrlich, 1965; Tulving, 1962, 1964) can be interpreted to mean that interitem associations are strengthened as a consequence of repeated presentations and recall of the material. What are the implications of this fact for the problem of why learning occurs?

On intuitive grounds, if nothing else, the obvious answer seems to be that learning is a consequence of the strengthening of interitem associations. Such an interpretation has been put forth explicitly elsewhere (Tulving, 1962, 1964, 1966a). It was argued, by way of an extension of a position proposed earlier with respect to immediate memory by Miller (1956a, 1956b), that the subject can recall a limited number of S-units on each of the successive trials and that the learning curve expressed in E-units reflects only the increasing size of S-units. This hypothesis is not unlike Deese's (1961) conjecture that the number of words remembered on a single trial is the same under a variety of experimental conditions and that observed variations in the number of words recalled can be accounted for by associations evoked by the remembered "core" words as well as by subject's constructions of additional words on the basis of his knowledge of the language. The S-unit invariance hypothesis, however, holds that all items recalled are those *actually remembered* from previous input and output phases. Extralist intrusions, sometimes highly predictable from the nature of the material (Deese, 1959b), probably represent remembered associative responses made to list-words in the input phase rather than associative constructions at the time of recall. They are intrusions only in the sense that they do not correspond to the experimenter's tally sheet, and not necessarily in the sense that the ideas they represent are not *remembered* from earlier parts of the same or similar experiments.

Although it seems reasonable to assume that the correlation between measures of subjective organization and number of items recalled over trials indicates the dependence of increments in recall on the strengthening of specific interitem associations, alternative interpretations are possible. The first such interpretation holds that increments in organization depend on increments in recall. This position was implicit in Carterette and Coleman's (1963) suggestion that increments in subjective organization *follow* increments in recall. The second alternative is that both increments in recall and increments in organization are parallel but otherwise unrelated manifestations of a common underlying process and that one has no more effect on the other than the second has on the first. So far no champion of this point of view seems to have stepped forth.

Availability of items and contextual associations

If free-recall learning occurs independently of increments in organization, what other mechanisms can be postulated to account for learning? Asch and Ebenholtz (1962) have suggested that free recall is not a function of interitem associations and that it depends solely on "availability" of items. They have even claimed, on the basis of the results of several experiments involving memorization of lists of eight nonsense syllables, that no interitem associations are established in the course of repeated free recall of items, but such a claim is not logically necessary for the integrity of their main assertions. Availability, according to Asch and Ebenholtz, "refers to the accessibility of an item to recall; the index of availability is the ease with which a datum of past experience can be produced. Central to our conception is the definition of availability as a phenomenon of recall, distinct from associations between items; availability refers to a property of memory traces" (p. 21). Availability as a fundamental condition of recall, and presumably of increments in recall, is said to be determined by factors such as frequency, primacy, recency, and isolation of individual items in a series. Thus, an item that has been perceived frequently and recently is more available for recall than another item perceived less frequently and less recently.

Asch and Ebenholtz's (1962) concept of availability seems to be another descriptive label for recall probability, and as such it does not take us beyond the immediate empirical data. Moreover, as we shall see presently, some experimental data clearly contradict the assumption that probability of recall unconditionally depends on frequency and recency.

The assumption that free-recall learning occurs in the absence of interitem associations also has been made by all authors of stochastic models of free-recall learning (Bush and Mosteller, 1955; Miller and McGill, 1952; Kintsch and Morris, 1965; Waugh and Smith, 1962), but this assumption is made only for the sake of mathematical convenience. All these writers have pointed out that clustering is one of the most persistent characteristics of free-recall data. Phenomena of secondary organization, or, for that matter, primary organization, as they occur in free recall have not yet been incorporated into mathematical models. Miller and McGill (1952) did deduce the prediction from their model that associative clustering should affect only the variability and not the rate of learning, but this deduction seems to be at variance with the data that have since been accumulated by many experimenters.

A different approach to the problem of learning in free recall is implicit in the writings of those psychologists who have postulated associations between list-items and the general experimental context to account for the response-learning stage in paired-associate learning (Keppel, 1964; McGovern, 1964; Underwood, 1964; Underwood, Runquist, and Schulz, 1959; Underwood and Schulz, 1960). According to these formulations, increments in recall of individual items *qua* individual items over trials reflect the strengthening of such contextual associations, the context being defined in terms of "all stimuli, other than the specific items, which contribute to the recall of the response terms under the conditions of free recall" (Keppel, 1964,

p. 92). Postman (1964) also has stated, "Rehearsal will strengthen the connection between each item and the situational context, but it will serve other functions as well, depending on the characteristics of the items in the list" (p. 176). Unlike Asch and Ebenholtz, and like the authors of stochastic models, these writers do not deny the possible role of interitem associations in determining free recall. Underwood and Schulz (1960) and Postman (1964) have stated explicitly that associations among items in the list may facilitate recall.

Evidence for organizational factors in learning

Is it necessary to postulate the strengthening of both contextual and specific interitem associations to account for learning in free recall, or can the whole burden be carried by organizational processes alone? This is not an easy question to answer, and it may be that the issue will not be satisfactorily settled for a long time. But I will argue that at least at the present time there is no compelling need to postulate the strengthening of contextual associations to account for trial-by-trial increments in free recall and that the available evidence points only to secondary organization or strengthening of interitem associations as the critical ingredient of free-recall learning. This is not to say that contextual associations—which I would like to define as associations between different aspects of the subject's total experience in an experimental situation and the to-be-remembered items—play no role at all in free recall. It is logically necessary that at least one item—and psychologically plausible that a small but fixed number of items—be initially associated with the context to make later recall of the list possible. What seems to be unnecessary is the assumption of the *strengthening* of direct contextual associations; it is at least unnecessary to assume that they play anything more than a minor role in memorization of a list.

The first reason for this view is that while it is exceedingly simple to demonstrate the development and strengthening of interitem associations in the course of trial-by-trial practice—or in the course of a prolonged study period—and only a little more difficult to show that these associations facilitate recall, it is quite difficult, if not impossible, to prove that associations between items and the general context are strengthened *independently* of specific interitem associations, as long as items are presented in the context of other items, as they are in free-recall experiments.

The second and a more germane reason for the suggestion that strengthening of contextual associations plays only a minor role in free-recall learning lies in some experimental evidence reported by several investigators. We will consider this evidence briefly.

Murdock and Babick (1961) measured probability of free recall of a single critical word presented on a number of successive trials, but always in the context of new sets of words. In this situation the association between the critical word and the general nonverbal context can become stronger over trials, since the context remains stable; but interitem associations cannot be strengthened, because the verbal context changes from trial to trial. The results were quite clear in showing that the recall of the critical word remained constant over trials. This finding is directly contrary to the hypothesis that repetition strengthens associations between the general nonverbal context and individual words within the verbal context. It provides indirect support to the notion that repetition effects are mediated through interitem associations.

The next source of relevant evidence is a series of experiments we have done at Toronto (Tulving, 1966a). In one type of experiment, subjects read aloud a list of 22 words, all words being presented six times in a quasi-random order. Immediately after the completion of this task the subjects were presented the same set of 22 words for learning on 12 trials under the typical free-recall conditions. The resulting learning curve was identical with that obtained from another group of subjects who learned the identical list after reading a different set of words. This finding demonstrates that frequency and recency alone are not sufficient to increase the availability of responses, although it does not negate the possibility that general background associations could be strengthened under conditions where subjects are instructed to memorize the material. However, evidence is available from at least two other experiments on single-trial free recall that does cast considerable doubt on this possibility.

Mechanic (1964, Exp. III) had two groups of subjects learn a list of 24 pronounceable trigrams under typical learning instructions. Trigrams were presented at the rate of 11 sec. each. One group had to pronounce each trigram over and over again until the next one was shown in the input series, while the second group was left free to do whatever they wished in the intervals between items. The mean number of trigrams recalled was significantly higher in the second group, despite the fact that the first group made many more overt responses corresponding to list items. Similar findings have been reported by Glanzer and Meinzer (in press) in an experiment already cited. One group of their subjects also was asked to repeat words overtly during interword intervals in the input phase of a single-trial free-recall experiment, while the other group was left to their own devices in trying to remember the material. Again, subjects who did not have to repeat the words showed higher recall than those who were instructed to do so. Thus, whatever else overt repetition does, it does not seem to contribute very much to recall. It probably only prevents subjects from organizing the material, thus producing lower recall than that of subjects who are free to organize the material.

Coming back to multitrial experiments, we note the second type of experiment reported by Tulving (1966a). Experimental groups learned a list of words under typical free-recall conditions and then learned another list, twice as long as the first one, which contained all the words learned previously. The control groups learned the same second lists as the experimental groups but different first lists. The learning curve of the experimental group was only initially higher than that of control groups. Following four to six trials the curves crossed, the control subjects recalling more words than the experimental subjects. This finding again is contrary to the hypothesis that strengthening of associations between individual list words and the stable nonverbal context takes place and is an important determinant of learning in free recall.

The next line of evidence relevant to the issue comes from a series of experiments by Mandler (1966a). Subjects were first asked to sort sets of words into a number of categories of their own choice. After they had arrived at a consistent set of categories, indicated by their ability to produce two identical sorts in succession, they were asked to recall as many words as they remembered. The number of words recalled was a linear function of the number of categories used by subjects but independent of the number of trials required for reaching the sorting criterion. Again it seems that it is not the frequency with which each word had been responded to by the subject in the presence of the general experimental stimuli that determined recall, but rather the number of higher-order units into which words were organized.

The final bit of relevant evidence, though somewhat less direct, can be gleaned from data published by Bousfield and Cohen (1953). In their experiment, different groups of subjects were presented with a categorized list of 40 words from one to five times. The number of words recalled increased directly as a function of the number of presentations, but the increase was attributable solely to the formation of larger clusters with repeated presentations. As shown in their Table 5, the number of nonclustered items, as well as the number of small clusters, showed no systematic trend over the range of one to five repetitions, although such a trend was very pronounced for "statistical" subjects. The effect of repetition lay in the production of larger S-units and not in the strengthening of background associations of individual items.

While all this evidence is still rather fragmentary, these findings are difficult, if not impossible, to reconcile with the hypothesis that repetition strengthens associative connections between individual items and the general nonverbal contextual stimuli. They make good sense, however, if we assume that repetition is effective to the extent to which it permits the subjects to organize E-units into larger S-units. I have been unable to think of any evidence that would unequivocally support the hypothesis that repetition strengthens contextual associations of individual items independently of interitem associations. Until such evidence becomes available, theoretical conceptions of the free-recall process and of the effect of repetition on free recall, which attribute increments in availability of individual items to frequency and recency stimulation, or to strengthening of associative bonds between individual items and their nonverbal context, must be rejected in favor of the view that repeated presentation of the material over trials, or increasing study time, produces increments in recall because individual items can be organized into larger cohesive units and are processed as parts of these units.

Storage versus Retrieval Processes

Melton (1963) has defined the domain of the theory of memory in terms of two broad classes of problems, one having to do with storage of traces and the other with utilization or retrieval of traces. Some of the more specific issues Melton lists under the general rubric of problems of retrieval seem to be as relevant to storage as to retrieval. But at least one of them, namely the dependence of retrieval on the reinstatement of original stimulating conditions, illustrates the feasibility of analytical separation of storage and retrieval processes (or retention and recall) and implies that retrieval depends not only on the contents of the storage at any given time, but on some other factors as well.

Theories of free recall have had little to say about retrieval processes. We have already seen that some theoretical formulations (e.g., Asch 1964; Asch and Ebenholtz, 1962; Waugh and Norman, 1965) sidestep the issue completely by assuming that recall is determined by the availability of items in the storage or by the availability of traces of items. Others suggest that recall of individual items is primarily dependent on the development of higher-order S-units (Tulving, 1962, 1964) or on the total number of categories into which material has been organized (Mandler, 1966a, 1966b), but they have not been explicitly concerned with retrieval processes, independently of storage processes.

The heuristic usefulness of the distinction between storage and retrieval or, more specifically, between potential availability of information in the storage and accessibility of this information is indicated by certain experimental data. First, it is a well-known fact that many items a subject cannot recall can nevertheless be readily recognized as members of a previously seen experimental set. Michel (1923), for instance, found in one of his experiments that subjects could correctly recognize all items they had failed to recall. Such demonstrations, however, do not prove that information sufficient for the *reproduction* of items is sometimes potentially available in the storage even though it is not accessible. Scores on a recognition test are determined by different retained information from that measured by recall tests.

The second and more relevant fact is provided by demonstrations that subjects can recall more items at a later time than at an earlier time without the benefit of interpolated presentation. In an early experiment (Nicolai, 1922) subjects were shown a simultaneous display of ten small objects. Free recall tests were given twice, once immediately after the presentation and the second time following various intervals up to 96 hours. The number of objects recalled was invariably greater in the second test than in the first. The largest difference occurred in the 96-hour group, in which the mean immediate recall score was 7.5 and the second test score was 9.4.

It might be argued that the first test responses provide additional practice or that subjects rehearse the material during the retention interval and that such practice is responsible for increased recall from the first to the second test, but unless certain additional assumptions are made about the retrieval process, these arguments are beside the mark. How would practice or rehearsal of items the subject could recall on the first test lead to subsequent recall of items the subject could not recall on the first test? In a recent experiment at Toronto (Tulving, 1966b) subjects were shown 36 words and asked to recall as many words as they could remember in each of three successive output phases. A total of six such "cycles" were given with the same list. One interesting finding was that only approximately 50 per cent of all the words a typical subject could recall at least once in any of the three intracycle output phases were recalled consistently in all three phases, even though the total number of words recalled in each output phase within a cycle was relatively constant. These findings were interpreted in support of the view of memory as a limited capacity retrieval system in which the limit is set by the number but not by the nature of the contents of accessible memory units. It is as if retrieval of one unit makes access to another unit impossible even though the latter is available in storage.

All these findings indicate that in free-recall situations the retrieval mechanism does not have access to all of the relevant information available in the storage. This indeed was one of the reasons why Ebbinghaus rejected the method of retained members as not particularly suitable for the study of memory. He failed to recognize that subjects' inability to recall all the material they have retained poses one of the truly challenging problems for the theory of memory. Why can the subject not recall everything he has "learned" and retained?

There is nothing in the free-recall literature that would provide a very meaningful answer to this question. About all we can say at the present time is that accessibility of individual items seems to depend on the degree of organization of items into higher-order units. Miller and McGill's (1952) casual observation that "many pairs of words are recalled together or omitted together on successive trials" (p. 389) suggests that one member of a higher-order S-unit serves as a retrieval cue for other members. One higher-order unit may similarly serve as a retrieval cue for another unit, in keeping with the hierarchical nature of organization described by Miller (1956a) and elaborated by Mandler (1966a). Exactly which of the available items within S-units are recalled would then depend on the order in which retrieval proceeds.

The role of experimentally manipulated retrieval cues in recall of words from a categorized list was examined in an experiment by Tulving and Pearlstone (1966). Lists consisted of words belonging to explicitly designated conceptual categories. Three levels of list length (12, 24, and 48 words) were combined with three levels of number of words per category (1, 2, and 4 words) to yield nine input conditions. Both category names and words belonging to categories were presented in the input phase, but subjects were told that only their recall for words would be tested. For half the subjects in each of the nine input conditions, recall of words was tested in the presence of category names as retrieval cues, for the other half recall was not cued. Since all subjects in a given input condition were treated identically until the end of the input phase, availability of information at the beginning of the output phase was equated for subjects in the cued and non-cued recall groups. Any differences in recall between these groups could thus be attributed to differences in accessibility.

Cued recall was greater than noncued recall for all nine input conditions, the difference varying directly with list length and inversely with the number of words per category. The word-recall scores in all groups were further analyzed in terms of two components of such scores—number of accessible categories, i.e., categories for which at least one word was recalled (Cohen, 1963a, 1963b) and number of words within accessible categories. This analysis yielded two important findings. First, the presentation of category names as retrieval cues affected only the number of accessible categories but not the number of words within accessible categories. Thus, given that the subject could recall at least one word from a category, the number of additional words recalled for that category was the same regardless of whether he remembered the category on his own or was reminded of it by the experimenter. Second, the number of accessible categories was considerably greater for 48-word lists than for 24-word lists, but again the number of words within accessible categories was identical for both list lengths. Similar results have been reported by Cohen (1966). Cohen also found recall of words within accessible categories to be independent of the length of the lists varying from 35 to 70 words, although the total number recalled varied directly with the length.

These findings demonstrate the independence of word-recall within higher-order units from the recall of higher-order units. The proportion of words accessible within higher-order units is not influenced by the number of accessible higher-order units or by the availability of names of higher-order units as retrieval cues. Although these data were gleaned from experiments in which organization of the material was suggested to subjects by the experimenter, there is no reason to believe that the basic processes are different in situations in which individual E-units are organized into higher-order units subjectively. In either case the functional significance of organization lies in its facilitating effect on retrieval—organization makes individual list items more accessible to recall. Whether or not it also affects the amount of mnemonic information available in the storage, and whether or not it helps to conserve storage capacity, must for the time being remain open questions.

Summary

The free-recall paradigm is the descendant of the method of retained members which was widely used in the experimental study of memory around the turn of the century. The early work on free recall yielded a fair amount of evidence on first-order relations between a variety of experimental variables and recall performance, but at the theoretical level it remained quite sterile. As a consequence, the method was overshadowed by other paradigms, notably serial-anticipation and paired-associate learning, and for a while almost disappeared from the scene. In the past decade and a half, prompted by several parallel developments in the field of verbal learning, it has been revived, and it is now gradually evolving as an area of investigation in its own right.

While the more recent work has produced a number of relatively stable facts, and while initial attempts at theoretical analysis of some of the phenomena have been made, research in free recall still seems to lack a firm sense of purpose. This state of affairs is perhaps to be expected in any new area of research, but in the case of free recall it probably also stems from the heterogeneity of orienting attitudes of investigators and from the ambivalence of the logical status of free recall. Is it just another "method" of measuring retention? Is it an alternate proving ground of empirical relations and phenomena established in the context of other paradigms? Or does it constitute a somewhat different and in some sense a unique memory task whose analysis opens up new theoretical vistas and dictates the introduction of new concepts? The orienting attitudes reflected in these questions can all be discerned in the recent literature. Their very existence tends to retard the crystallization of important theoretical issues and to militate against the emergence of unifying focal points of experimental inquiry.

Typical free-recall experiments have little to do with learning of individual items. Rather, they are concerned with retention and recall, or with storage and retrieval, of mnemonic information of the sort that permits later reproduction of desired aspects of earlier experiences. The subject may be said to retain and recall individual list items only insofar as the experimenter adopts as a measure of the subject's performance the number of correct responses; that is, responses corresponding to experimentally designated discrete units of material.

Evidence shows that individual list items are not always processed independently of one another, thus hinting at the discrepancy between nominal and functional units of material, or between E-units and S-units. If biological memory systems, such as that of a human being, have limited capacity for processing information, and if the limitation depends on the number of units of information that can be processed within a given period of time—two fundamental assumptions which for good reasons seem to be widely accepted—it would also seem to be reasonable to assume that the limit of free recall is set by the number of functional units rather than nominal units. Some experimental findings suggest that this is indeed the case. An important theoretical problem in free recall lies in the identification of functional S-units.

The discrepancy between E-units and S-units is most difficult to detect under conditions where the associative relations among individual list items are minimized, where the amount of time available to the subject for thinking about the items is very limited, and where recall of a set of items by a given subject is tested only once. If, under these conditions, frequency of recall is plotted against the position of items in the input list, characteristic serial position effects are found. Evidence shows that different parts of the serial position curve are affected by different independent variables. These findings support the hypothesis that at least two different kinds of underlying process, referred to as primary and secondary memory, are involved in recall. Some theorists have identified primary and secondary memory with separate storage systems. This formulation is contrasted with an alternative view that storage is unitary and that differential recall of items from different serial positions and the differential effects of various independent variables on such recall are to be attributed to different kinds of retrieval mechanism. Whether the two approaches will generate testable hypotheses, and how these will fare in the laboratory, remains a problem for future research.

The liveliest theoretical activity in the context of free-recall experiments has been focused on organizational processes. Organization as a response variable refers to sequential constraints in output sequences that can be shown to be independent of the order of items in input sequences. Two types of organization may be distinguished. Primary organization is relatively independent of both the associative meaning of input items and the subject's prior familiarity with a given list as a list. Primary organization lies at the root of the recency effect and may also be related to the primacy effect. Secondary organization is largely determined by the preexperimental and intraexperimental (intertrial) experiences the subject has had with individual input items.

Measures of secondary organization can be divided into two broad classes—clustering and subjective organization. The two classes of measures are predicated on somewhat different assumptions and are not equally applicable in all situations. Clustering can be assessed only for lists consisting of items belonging to specifiable subsets, while measures of subjective organization can be obtained for any set of items that are recalled by subjects more than once. Nevertheless, both types of measures presumably reflect the effects of common underlying processes. In keeping with the tradition of "two-factor" theories of psychology, explanations of clustering—and by implication, of subjective organization—have also postulated two different kinds of mechanism: organization based on direct associative relations among items, and organization based on indirect associations or on mediation through response-produced cues. This conceptual distinction reflects the empirical distinction between associative and category clustering, which in turn is determined by the relations of list items to normative data collected in free and restricted word-association tests.

While the importance of the subject's prior verbal habits in determining what specific list items are organized into which S-units is beyond dispute, attempts to distinguish between mechanisms of secondary clustering on the basis of extra-experimental word-association data seem to be rather pointless. At the stage of our present ignorance as to the mechanisms underlying associative relations between verbal units in any situation, it appears more fruitful to begin with the observation that some items are more likely to be organized into a given higher-order unit than are others, and then to seek the explanation of such organization in intra-experimental conditions, such as the contiguity relations, both in input and output sequences, between potential members of a functional S-unit.

The fact that recall increases systematically over trials in multitrial experiments constitutes a theoretical puzzle of long standing, not only in free recall, but in other paradigms as well. Some writers have suggested that individual list items are processed independently of one another and that repeated presentation or recall, or both, increases the availability of items for subsequent recall. Others have conceptualized trial-to-trial learning in terms of strengthening of traces. And still others have implied that such learning reflects the strengthening of associative bonds between individual items and the general experimental context in which rehearsal occurs. None of these formulations have been stated explicitly enough to make it possible to derive critical testable hypotheses from them. More often than not they simply represent restatement of the fact that trial-to-trial learning occurs and that it is influenced by those variables that influence them.

A more promising approach consists in the postulation of processes underlying secondary organization as the mechanism of learning. According to this view, the retrieval system can have access to only a limited number of functional S-units in a given output phase, and any increase in the recall of nominal E-units reflects the increase in the size of the accessible S-units as a consequence of secondary organization. Evidence supporting this type of formulation is provided by findings that secondary organization increases over trials and that it shows sizable correlations with the number of recalled E-units. Other evidence contradicts the hypothesis that repeated presentation of items in presence of stable nonverbal contextual stimuli alone is sufficient for the strengthening of traces or general background associations of items or for the enhancement of their availability for recall.

The conceptual distinction between retention and recall, or between storage and retrieval, has only recently been translated into operations that allow for experimental separation of these two stages of remembering as they occur in free recall. The results from a number of studies suggest that the amount of relevant information available in the memory storage and thus potentially usable for reproduction of list items is greater than the amount of information accessible to the retrieval system. Accessibility of relevant information, and hence the amount of recall, depends not only on the contents of the storage, but also on retrieval cues. For a long time, identification of retrieval cues was the Achilles' heel of theoretical formulations of free recall, but now some progress has been made in remedying this weakness. Initial experimental results suggest that the retrieval system operates independently at different levels of organization of the material to be remembered.

In general, at the present time we are nowhere near a cohesive theory of memory which would explain free-recall phenomena in all their complexity. Theoretical speculation about these phenomena is mostly restricted to restatements of known empirical facts. The major concepts of General Behavior Theory have been conspicuously absent in such restatements. It looks as if conceptual analyses of free recall have been developed not just in isolation, but almost in defiance of the traditional S-R models of behavior. Whether this isolation is only semantic, and hence temporary, or whether it represents a substantive and more permanent break with the past is a question which can be answered only by future research.

References

Abbott, E. A., "On the Analysis of the Factor of Recall in the Learning Process," *Psychological Monographs*, **11** (1909), 159–177.

Asch, S. E., "The Process of Free Recall," in C. Scheerer, ed., *Cognition: Theory, Research, Promise*. New York: Harper and Row, 1964, pp. 79–88.

———, and S. M. Ebenholtz, "The Principle of Associative Symmetry," *Proceedings of the American Philosophical Society*, **106** (1962) 135–163.

Battig, W. F., M. Allen, and A. R. Jensen, "Priority of Free Recall of Newly Learned Items," *Journal of Verbal Learning and Verbal Behavior*, **4** (1965), 175–179.

Bolton T. L., "The Growth of Memory in School Children," *American Journal of Psychology*, **4** (1892), 362–380.

Bourne, L. E., and B. K. Parker, "Inter-item Relationships, List Structure, and Verbal Learning," *Canadian Journal of Psychology*, **18** (1964), 52–61.

Bousfield, W. A., "The Occurrence of Clustering in the Recall of Randomly Arranged Associates," *Journal of General Psychology*, **49** (1953), 229–240.

———, H. Berkowitz, and G. A. Whitmarsh, "Associative Clustering in the Recall of Minimally Meaningful Geometric Designs," *Canadian Journal of Psychology*, **13** (1959), 281–287.

———, and B. H. Cohen, "The Effects of Reinforcement on the Occurrence of Clustering in the Recall of Randomly Arranged Associates," *Journal of Psychology*, **36** (1953), 67–81.

———, and ———, "The Occurrence of Clustering in the Recall of Randomly Arranged Words of Different Frequencies-of-Usage," *Journal of General Psychology*, **52** (1955), 83–95.

———, and ———, "Clustering in Recall as a Function of the Number of Word-Categories in Stimulus-Word Lists," *Journal of General Psychology*, **54** (1956), 95–106.

———, ———, and J. G. Silva, "The Extension of Marbe's Law to the Recall of Stimulus Words," *American Journal of Psychology*, **69** (1956), 429–433.

———, ———, and G. A. Whitmarsh, "Associative Clustering in the Recall of Words of Different Taxonomic Frequencies of Occurrence," *Psychology Reports*, **4** (1958), 39–44.

———, and C. R. Puff, "Clustering as a Function of Response Dominance," *Journal of Experimental Psychology*, **67** (1964), 76–79.

———, ———, and T. M. Cowan, "The Development of Constancies in Sequential Organization during Repeated Free Recall," *Journal of Verbal Learning and Verbal Behavior*, **3** (1964), 449–459.

———, J. R. Steward, and T. M. Cowan, "The Use of Free Associational Norms for the Prediction of Clustering," *Journal of General Psychology*, **70** (1964), 205–214.

Bowers, H., "Memory and Mental Imagery," *British Journal of Psychology*, **21** (1931), 271–282.

Brown, W., "To What Extent Is Memory Measured by a Single Recall?" *Journal* *of Experimental Psychology*, **6** (1923), 377–382.

Bruner, J. S., G. A. Miller, and C. Zimmerman, "Discriminative Skill and Discriminative Matching in Perceptual Recognition," *Journal of Experimental Psychology*, **49** (1955) 187–192.

Buseman, A., "Lernen und Behalten: Beiträge zur Psychologie des Gedächtnisses," *Z. angew. Psychol.*, **5** (1911), 211–271.

Bush, R. R., and F. Mosteller, *Stochastic Models for Learning*. New York: John Wiley and Sons, Inc., 1955.

Carterette, E. C., and E. A. Coleman, "Organization in Free Recall." Paper presented at the meeting of the Psychonomic Society, Bryn Mawr, Pa., August, 1963.

Cofer, C. N., "A Study of Clustering in Free Recall Based on Synonyms," *Journal of General Psychology*, **60** (1959), 3–10.

———, "On Some Factors in the Organizational Characteristics of Free Recall," *American Psychology*, **20** (1965), 261–272.

———, "Some Evidence for Coding Processes Derived from Clustering in Free Recall," *Journal of Verbal Learning and Verbal Behavior*, **5** (1966), 188–192.

———, and D. R. Bruce, "Form-Class as the Basis for Clustering in the Recall of Nonassociated Words," *Journal of Verbal Learning and Verbal Behavior*, **4** (1965), 386–389.

———, ———, and G. M. Reicher, "Clustering in Free Recall as a Function of Certain Methodological Variations," *Journal of Experimental Psychology*, **71** (1966), 858–866.

Cohen, B. H., "An Investigation of Recoding in Free Recall," *Journal of Experimental Psychology*, **65** (1963), 368–376.

———, "Recall of Categorized Word Lists," *Journal of Experimental Psychology*, **66** (1963), 227–234.

———, "Some-or-None Characteristics of Coding," *Journal of Verbal Learning and Verbal Behavior*, **5** (1966), 182–187.

———, and W. A. Bousfield, "The Effects of a Dual-Level Stimulus-Word List on the Occurrence of Clustering in Recall," *Journal of General Psychology*, **55** (1956), 51–58.

Dallett, K. M., "Practice Effects in Free and Ordered Recall," *Journal of Experimental Psychology*, **66** (1963), 65–71.

———, "Number of Categories and Category Information in Free Recall," *Journal of Experimental Psychology*, **68** (1964), 1–12.

Deese, J., "Serial Organization in the Recall of Disconnected Items," *Psychology Reports*, **3** (1957), 577–582.

———, "Influence of Interitem Associative Strength upon Immediate Free Recall," *Psychology Reports*, **5** (1959), 305–312. (a)

———, "On the Prediction of Occurrence of Particular Verbal Intrusions in Immediate recall," *Journal of Experimental Psychology*, **58** (1959), 17–22. (b)

———, "Frequency of Usage and Number of Words in Free Recall: The Role of Association," *Psychology Reports*, **7** (1960), 337–344.

———, "From the Isolated Verbal Unit to Connected Discourse," in C. N. Cofer, ed., *Verbal Learning and Verbal Behavior*. New York: McGraw-Hill Book Company, 1961, 11–31.

———, "On the Structure of Associative Meaning," *Psychology Review*, **69** (1962), 161–175.

———, and R. A. Kaufman, "Serial Effects in Recall of Unorganized and Sequentially Organized Verbal Material," *Journal of Experimental Psychology*, **54** (1957), 180–187.

Dronsejko, K., "The Role of Organization in Verbal Learning." Unpublished M. A. thesis, University of Toronto, 1963.

Ebbinghaus, H., *Grundzüge der Psychologie*. Leipzig: Von Veit, 1902.

Ehrlich, S., "Le rôle de la structuration dans l'apprentissage verbal," *Psychologie Française*, **10** (1965), 119–146.

Ekstrand, B., and B. J. Underwood, "Paced versus Unpaced Recall in Free Learning," *Journal of Verbal Learning and Verbal Behavior*, **2** (1963), 288–290.

Estes, W. K., "Learning Theory and the

New 'Mental Chemistry,'" *Psychology Review*, **67** (1960), 207–223.

Garner, W. T., and J. R. Whitman, "Form and Amount of Internal Structure as Factors in Free-Recall Learning of Nonsense Words," *Journal of Verbal Learning and Verbal Behavior*, **4** (1965), 257–266.

Gonzalez, R. C., and C. N. Cofer, "Exploratory Studies of Verbal Context by Means of Clustering in Free Recall," *Journal of Genetic Psychology*, **95** (1959), 293–320.

Glanzer, M., and A. R. Cunitz, "Two Storage Mechanisms in Free Recall," *Journal of Verbal Learning and Verbal Behavior*, **5** (1966), 351–360.

———, and A. Meinzer, "The Effects of Intralist Activity on Free Recall," *Journal of Verbal Learning and Verbal Behavior* (1967).

Green, R. T., and G. Harding, "Sequential Recall of a Mixed List," *British Journal of Psychology*, **53** (1962), 389–396.

Holmes, F.L.D., "Syllabic Length, Recognition and Immediate Recall," *Journal of Applied Psychology*, **18** (1934), 831–841.

James, W., *The Principles of Psychology*. Vol. 1. New York: Holt, Rinehart and Winston, Inc., 1890. Reprinted by Dover Publications, Inc., 1950.

Jenkins, J. J., W. D. Mink, and W. A. Russell, "Associative Clustering as a Function of Verbal Association Strength," *Psychology Reports*, **4** (1958), 127–136.

———, and W. A. Russell, "Associative Clustering during Recall," *Journal of Abnormal and Social Psychology*, **47** (1952), 818–821.

Keppel, G., "Facilitation in Short- and Long-term Retention of Paired Associates Following Distributed Practice in Learning," *Journal of Verbal Learning and Verbal Behavior*, **3** (1964), 91–111.

Kintsch, W., and C. J. Morris, "Application of a Markov Model to Free Recall and Recognition," *Journal of Experimental Psychology*, **69** (1965), 200–206.

Kirkpatrick, E. A., "An Experimental Study of Memory," *Psychology Review*, **1** (1894), 602–609.

Lashley, K. S., "In Search of the Engram," in *Symp. Soc. Exp. Biol.*, No. 4 (1950), pp. 454–482.

Mandler, G., "Organization and Memory," in K. W. Spence and Janet T. Spence, eds., *The Psychology of Learning and Motivation*. New York: Academic Press, 1966. (a)

———, "Verbal Learning," in *New Directions of Psychology*, III. New York: Holt, Rinehart & Winston, Inc., 1966. (b)

Marshall, G. R., and C. N. Cofer, "Associative Indices as Measures of Word Relatedness: A Summary and Comparison of Ten Methods," *Journal of Verbal Learning and Verbal Behavior*, **1** (1963), 408–421.

McGeoch, J. A., "Memory," *Psychology Bulletin*, **27** (1930), 514–563.

———, *The Psychology of Human Learning*. New York: Longmans, Green & Co, Inc., 1942.

———, and A. L. Irion, *The Psychology of Human Learning*. New York: Longmans, Green & Co., Inc., 1952.

McGovern, J. B., "Extinction of Associations in Four Transfer Paradigms," *Psychological Monographs*, Vol. **78**, No. 16 (1964).

McNulty, J. A., "The Measurement of 'Adopted Chunks' in Free-Recall Learning," *Psychonomic Science*, **4** (1966), 71–72.

McQuitty, L. L., "Rank Order Typal Analysis," *Educational Psychology Measurement*, **23** (1963), 55–61.

Mechanic, A., "The Responses Involved in the Rote Learning of Verbal Materials," *Journal of Verbal Learning and Verbal Behavior*, **3** (1964), 30–36.

Melton, A. W., "Implications of Short-term Memory for a General Theory of Memory," *Journal of Verbal Learning and Verbal Behavior*, **2** (1963), 1–21.

Messer, A., *Psychologie*. Stuttgart: Deutsche Verlagsanstalt, 1922.

Meuer, H., "Experimentelle Beiträge zur Lehre vom Wortgedächtnis nach der Methode der Polyeidoskopie," *Arch. ges. Psychol.*, **47** (1924), 14–44.

Meumann, E., *The Psychology of Learning*, trans. J. W. Baird. New York: Appleton-Century, 1913.

Meyer, D. R., and R. C. Miles, "Intralist-Interlist Relations in Verbal Learning," *Journal of Experimental Psychology*, **45** (1953), 109–115.

Michel, O., "Experimentelle Untersuchungen über das Gedächtnis. Reproduktion und Wiedererkennen von optischen Eindrucken," *Arch. ges. Psychol.*, **44** (1923), 244–271.

Miller, G. A., "Human Memory and the Storage of Information," *IRE Transactions on Information Theory*, **2** (1956), 129–137.

———, "The Magical Number Seven Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychology Review*, **63** (1956), 81–96.

———, J. S. Bruner, and L. Postman, "Familiarity of Letter Sequences and Tachistoscopic Indentification," *Journal of General Psychology*, **50** (1954), 129–139.

———, E. Galanter, and K. H. Pribram, *Plans and the Structure of Behavior*. New York: Holt, Rinehert & Winston, Inc., 1960.

———, and W. J. McGill, "A Statistical Description of Verbal Learning," *Psychometrika*, **17** (1952), 369–396.

———, and J. A. Selfridge, "Verbal Context and the Recall of Meaningful Material," *American Journal of Psychology*, **63** (1950), 176–185.

Murdock, B. B., Jr., "The Immediate Retention of Unrelated Words," *Journal of Experimental Psychology*, **60** (1960), 222–234.

———, "The Serial Position Effect in Free Recall," *Journal of Experimental Psychology*, **64** (1962), 482–488.

———, and A. J. Babick, "The Effect of Repetition on the Retention of Individual Words," *American Journal of Psychology*, **74** (1961), 596–601.

Nicolai, F., "Experimentelle Untersuchungen über das Haften von Gesichtseindrucken und dessen zeitlichen Verlauf," *Arch. ges. Psychol.*, **42** (1922), 132–149.

Pohlmann, A., "Experimentelle Beiträge zur Lehre von Gedächtnis." Berlin: Göttinger Diss., 1906.

Postman, L., "Short-term Memory and Incidental Learning," in A. W. Melton ed., *Categories of Human Learning*. New York: Academic Press, 1964, 145–201.

———, P. A. Adams, and A. M. Bohm, "Studies in Incidental Learning: V. Recall for Order and Associative Clustering," *Journal of Experimental Psychology*, (1956), 334–342.

———, ———, and L. W. Phillips, "Studies in Incidental Learning: II. The Effects of Association Value and of the Method of Testing," *Journal of Experimental Psychology*, **49** (1955), 1–10.

———, J. P. Egan, J. Davis, "Rate of Recall as a Measure of Learning: I. The Effects of Retroactive Inhibition," *Journal of Experimental Psychology*, **38** (1948), 535–546.

———, and W. O. Jenkins, "An Experimental Analysis of Set in Rote Learning: The Interaction of Learning Instruction and Retention Performance," *Journal of Experimental Psychology*, **38** (1948), 683–689.

———, and L. W. Phillips, "Short-term Temporal Changes in Free Recall," *Quarterly Journal of Experimental Psychology*, **17** (1965), 132–138.

Raffel, G., "The Effect of Recall on Forgetting," *Journal of Experimental Psychology*, **17** (1934), 828–838.

———, "Two Determinants of the Effect of Primacy," *American Journal of Psychology*, **48** (1936), 654–657.

Robinson, E. S., and M. A. Brown, "Effect of Serial Position upon Memorization," *American Journal of Psychology*, **37** (1926), 538–552.

Rothkopf, E. Z., and E. H. Coke, "The Prediction of Free Recall from Word Association Measures," *Journal of Experimental Psychology*, **62** (1961), 433–438.

Rozov, A. I., "Formation and Dynamics of Topic Clusters in Free Recall," *Soviet Psychology and Psychiatry*, **2** (1964), 19–26 (transl. from *Voprosy Psikhologii*, 1961).

Russell, W. A., and J. J. Jenkins, "The Complete Minnesota Norms for Responses to 100 Words from the Kent-

Rosanoff Word Association Test," *Technical Report No.* 11, *Contract N8-ONR-66216*, Office of Naval Research and University of Minnesota.

Seibel, R., "Organization in Human Learning: The Study Sheet Paradigm and Experiments One and Two," paper presented at the meeting of the Psychonomic Society, Chicago, Oct., 1965.

Smith, T. L., "On Muscular Memory," *American Journal of Psychology*, **7** (1896), 453–490.

Smith, W. G., "The Place of Repetition in Memory," *Psychology Review*, **3** (1896), 21–31.

Sperling, G., "A Model for Visual Memory Tasks," *Human Factors*, (1963), 519–531.

Stevens, S. S., "Mathematics, Measurement and Psychophysics," in S. S. Stevens ed., *Handbook of Experimental Psychology*. New York: John Wiley & Sons, Inc., 1951, pp. 1–49.

Tulving, E., "Subjective Organization in Free Recall of 'Unrelated' Words," *Psychology Review*, **69** (1962), 344–354.

———, "Intratrial and Intertrial Retention: Notes towards a Theory of Free Recall Verbal Learning," *Psychology Review*, **71** (1964), 219–237.

———, "The Effect of Order of Presentation on Learning of 'Unrelated' Words," *Psychonomic Science*, **3** (1965), 337–338.

———, "Subjective Organization and Effects of Repetition in Multi-Trial Free-Recall Learning," *Journal of Verbal Learning and Verbal Behavior*, **5** (1966), 193–197

———, "The Effects of Presentation and Recall of Material in Free-Recall Learning, "*Journal of Verbal Learning and Verbal Behavior*, **6** (1967), 175–184.

———, and J. E. Patkau, "Concurrent Effects of Contextual Constraint and Word Frequency on Immediate Recall and Learning of Verbal Material," *Canadian Journal of Psychology*, **16** (1962), 83–95.

———, and Z. Pearlstone, "Availability versus Accessibility of Information in Memory for Words," *Journal of Verbal Learning and Verbal Behavior*, **5** (1966), 381–391

———, and G. B. Thornton, "Interaction between Proaction and Retroaction in Short-term Retention," *Canadian Journal of Psychology*, **13** (1959), 255–265.

Underwood, B. J., "The Representativeness of Rote Verbal Learning," in A. W. Melton ed., *Categories of Human Learning*. New York: Academic Press, 1964, 47–87.

———, and R. W. Schulz, *Meaningfulness and Verbal Learning*. Chicago: J. B. Lippincott Co., 1960.

———, ———, W. R. Runquist, "Response Learning in Paired-Associate Lists as a Function of Intralist Similarity," *Journal of Experimental Psychology*, **58** (1959), 70–78.

Waugh, N. C., "Free versus Serial Recall," *Journal of Experimental Psychology*, **62** (1961), 496–502.

———, "The Effect of Intralist Repetition on Free Recall," *Journal of Verbal Learning and Verbal Behavior*, **1** (1962), 95–99.

———, and D. A. Norman, "Primary Memory," *Psychology Review*, **72** (1965), 89–104.

———, and J. E. K. Smith, "A Stochastic Model for Free Recall," *Psychometrika*, **27** (1962), 141–154.

Wechsler, D., "Engrams, Memory Storage, and Mnemonic Coding," *American Psychologist*, **18** (1963), 149–153.

Welch, G. B. and C. T. Burnett, "Is Primacy a Factor in Association Formation?" *American Journal of Psychology*, **35** (1924), 396–401.

Wicklund, D. A., D. S. Palermo, and J. J. Jenkins, "Associative Clustering in the Recall of Children as a Function of Verbal Association Strength," *Journal of Experimental Child Psychology*, **2** (1965), 58–66.

Woodworth, R. S., *Experimental Psychology*. New York: Holt, Rinehart & Winston, Inc., 1938.