AD 696 668

REHEARSAL, INTERFERENCE, AND SPACING OF PRACTICE IN SHORT-TERM MEMORY

Alexander Warren Pollatsek

Michigan University
Ann Arbor, Michigan

July 1969

AFOSR 69-2798TR

08773-41-T

HUMAN PERFORMANCE CENTER
DEPARTMENT OF PSYCHOLOGY
The University of Michigan, Ann Arbor

Rehearsal, Interference, and Spacing of Practice in Short-Term Memory

ALEXANDER WARREN POLLATSEK

Technical Report No. 16

JULY 1969

THE HUMAN PERFORMANCE CENTER

DEPARTMENT OF PSYCHOLOGY

The Human Performance Center is a federation of research programs whose emphasis is on man as a processor of information. Topics under study include perception, attention, verbal learning and behavior, short- and long-term memory, choice and decision processes, and learning and performance in simple and complex skills. The integrating concept is the quantitative description, and theory, of man's performance capabilities and limitations and the ways in which these may be modified by learning, by instruction, and by task design.

The Center issues two series of reports. A Technical Report series includes original reports of experimental or theoretical studies, and integrative reviews of the scientific literature. A Memorandum Report series includes printed versions of papers presented orally at scientific or professional meetings or symposia, methodological notes and documentary materials, apparatus notes, and exploratory studies.

THE UNIVERSITY OF MICHIGAN

COLLEGE OF LITERATURE, SCIENCE AND THE ARTS

DEPARTMENT OF PSYCHOLOGY

REHEARSAL, INTERFERENCE, AND SPACING OF PRACTICE IN SHORT-TERM MEMORY

Alexander Warren Pollatsek

HUMAN PERFORMANCE CENTER--TECHNICAL REPORT NO. 16

July, 1969

This work was partially supported by the Advanced Research Projects Agency, Department of Defense, monitored by the Air Force Office of Scientific Research, under Contract Nos. AF 49(638)-1235 and AF 49(638)-1736 with the Human Performance Center, Department of Psychology, University of Michigan, and by NIH Grant GM-01231-05.

PREFACE

This report is an independent contribution to the program of research of the Human Performance Center, Department of Psychology, on human information processing and retrieval, supported by the Advanced Research Projects Agency, Behavioral Sciences, Command and Control Research, under Order No. 461, Amendments 3 and 5, and monitored by the Behavioral Sciences Division, Air Force Office of Scientific Research, under Contract Nos. AF 49(638)-1235 and AF 49(638)-1736.

The doctoral dissertation committee was: Drs. R. A. Bjork (Chairman), C.H. Coombs, D. H. Krantz, and C. B. Moler. The author acknowledges the assistance of Ruven Brooks in conducting Experiment II and Stanley Bielby for designing and building the equipment.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

ABSTRACT

Short-term memory as exhibited in a variety of experimental paradigms is heavily influenced by variations in the time allotted for rehearsal, in the time allotted for rehearsal-preventing interfering activity, and in the interval separating successive presentations of an item. These findings and others have led to the emergence of both a general conceptual framework distinguishing between sensory, short-term, and long-term memory storage and some specific quantitative models that account quite well for subsets of the experimental findings.

Although there has been considerable research on the effects of rehearsal, interference, and the spacing of presentations in short-term memory, there has been little systematic empirical or theoretical work on the interactions among these variables. This dissertation was designed to investigate systematically such interactions in an effort to derive constraints on an adequate theory of short-term memory beyond those imposed by prior research.

Two experiments were conducted, each of which employed variations on the Brown-Peterson experimental paradigm. There were 72 S̲s̲ in each experiment and the items to be remembered were word trigrams consisting of three four-letter one-syllable nouns presented visually. The basic condition common to both experiments can be labeled PRIT where P denotes the presentation of a word trigram, R denotes a variable rehearsal period, I denotes a variable period of interpolated activity (counting backwards), and T denotes a test for recall of the trigram. This Basic condition was designed to assess the interaction of the effects of rehearsal and interference on performance.

In Experiment I there were two additional conditions as follows: $P_1R_1I_1P_2R_2I_2T_2$ (Double-Presentation) and $P_1R_1I_1T_1I_2T_2$ (Double-Test). The Double-Presentation condition was designed to assess the joint interaction of rehearsal, interference, and spacing of presentations and the Double-Test condition was designed to assess the permanence of the memory traces of items recalled on a first test in relation to amount of prior rehearsal and interfering activity.

In Experiment II there were also two conditions in addition to the basic condition: $PR_1I_1P'R_2I_2T$ (Forget) and I_1PRI_2T (Control). In the Forget condition, a different stimulus was presented at P' than at P and the S̲s̲ were required to remember only the second item. The Forget and Control conditions were designed to assess inter- and intra-trial proactive interference in relation to variations in rehearsal and interference.

The principal experimental findings of Experiments I and II can be summarized as follows. (a) Rehearsal of an item greatly improved its retention in the Basic condition, whereas rehearsal of the to-be-forgotten item in the Forget condition led to slightly better retention of the to-be-remembered item. (b) With increasing rehearsal the observed forgetting curves became more sigmoid in shape. (c) Increasing I_1 (the spacing interval) resulted in markedly better retention in the Double Presentation condition while increasing I_1 in the Forget and Control conditions resulted in only slightly better retention. (d) Retention was better in the Control condition than in the Forget condition. (e) There was almost no forgetting from the first test to the second test in the Double Test condition.

The results of Experiments I and II together with the results of other experiments suggest the following constraints on an adequate theory of memory: (a) there are at least two memory stores; (b) there is response competition in the long-term store (LTS); (c) some items in LTS are much more retrievable than others.

The data from Experiments I and II were analyzed in detail with respect to a specific theory designed to be consistent with the above constraints. Two major conclusions were drawn from the analysis: (a) rehearsal provided opportunity for more unique coding of LTS traces but did not necessarily strengthen them; (b) neither reduction of proactive interference nor consolidation could explain the effects of spacing in the Double-Presentation condition, but an explanation in terms of increased coding effectiveness with increased spacing seemed plausible.

CHAPTER I

INTRODUCTION

Research in short-term memory during the last ten years has produced data that have led to increasingly detailed theories of memory. Hopefully, each new experimental finding will contribute to the development of a general picture of the memory process by reducing the range of plausible theories. It is the goal of this dissertation to clarify several specific aspects of that general picture. A brief review of the basic theoretical issues in memory will provide a framework for the experimental work to be reported later.

Some Theoretical Issues in Short-Term Memory

Although there are many interesting issues in short-term memory, the following stand out as being the most central to this dissertation.

Number of Memory Stores

Although the debate on the number of different memory stores started out as a debate on whether there were one or two memory stores (a long-term store plus a separate short-term store), the issue has become more complex in the last ten years.

Work by Sperling (1960) and Averbach and Coriell (1961) on the visual system demonstrated the existence of a very-short-term store (with decay time less than a second) in which the memorial trace is a fairly literal visual copy of the stimulus. Neisser (1967) argues that work by Treisman (1964) demonstrates the existence of a similar sensory store in the auditory system in which a quite literal replica of the stimulus is stored. (In Treisman's work, the decay time is about 2 seconds.) Although the inference of such a store from Treisman's work is somewhat less compelling than from Sperling's work, the data from the two experiments make it reasonably likely that there is such a raw sensory store for each of the sensory modalities. The question still remains, however, whether there are two different functional stores in memory besides the various possible sensory stores.

The general conception of the two-store theories is that, in addition to the sensory store, there is a short-term store which more or less corresponds to one's "working memory" and there is a relatively permanent store. Another possibility is that owing to rehearsal mechanisms, the acoustic sensory store could be maintained for periods of 30 seconds or longer and, hence, there is no additional short-term store besides the sensory stores. The one-state theorists (e.g., Melton, 1963) claim that there has been little functional distinction made between short- and long-term memory and therefore memory is a continuum of states rather than several discrete states. In one-store systems, the phenomenological entity, working memory, is largely ignored.

Interference vs. Decay

This issue is usually coupled with the number of stores issue outlined above, with interference theorists believing in one store and decay theorists believing in two stores. The issues are, however, logically separable. The extreme form of the interference position is that all memory loss of the correct response is due to interference from competing responses associated with the stimulus situation, while the extreme form of the decay position is that all memory loss is caused by an autonomous weakening or disappearance of the memory trace. Although the extreme interference position is adhered to by some researchers, the extreme decay position is not. A typical "decay" position asserts that processes such as associative interference are important in the long-term store, but that loss of memory in the short-term store is largely accounted for by decay.

The associative interference position described above should not be confused with the central limited-capacity store idea (Broadbent, 1958). In the latter conception, memory loss in short-term memory occurs largely because the limited-capacity store becomes overloaded and information is lost. In contrast to interference theory in which similarity of intervening activity is considered a crucial variable, processing difficulty of the intervening task is considered to be a crucial variable.

Strengthening of Memory Traces

Another issue is whether memory traces can be strengthened and, if so, by what mechanism. Three positions seem to define the range of thought. The first says that memory traces are either there or they aren't there (all-or-none), the second says that the traces can assume a continuum of strength values (strength), and the third says that one can create one or more copies of a memory trace, but that each trace is all-or-none (multiple copy). There are, of course, other possible positions (e.g., a multiple-copy strength model).

As in the interference vs. decay issue, one could visualize different processes in the two stores. One could visualize an all-or-none short-term store (STS) with a multiple-copy long-term store (LTS) or a strength short-term store with an all-or-none long-term store, etc.

Consolidation

Perhaps the most general statement of the consolidation hypothesis is that the stimulus has some chance to increase in strength or permanence each moment that it is held in the memory system. The most generally explored mechanism of consolidation is that the stimulus has some probability of going to the LTS from the STS whenever the stimulus is in the STS. A secondary question that is both important and difficult to answer is whether the consolidation process can go on during the performance of a task that prohibits the subject from paying attention to the to-be-remembered item.

Differential Coding

Another issue is whether the S's ability to form long-term traces is influenced by what state the memorial trace is in at the moment. One specific form of this general notion that has been of considerable interest is that having a trace in the short-term store works to inhibit formation of a permanent trace in some way.

Some Relevant Experimental Findings in Short-Term Memory

With this brief framework provided, a consideration of some of the relevant phenomena of short-term memory will give shape to the issues outlined in this section. The following review will be in two parts. The first part will review work done with the Brown-Peterson (B-P) paradigm, which is the paradigm used in the experiments to be discussed later. The second part will relate these findings to other short-term memory findings to discover the generality of regularities found in the Brown-Peterson research. Readers familiar with the short-term memory literature can proceed without loss of continuity to the summary section of this chapter.

The Brown-Peterson Paradigm

One of the simplest paradigms created to study short-term retention was developed by Brown (1958) and elaborated and popularized by Peterson and Peterson (1959). A description of Peterson and Peterson's task will serve to illustrate the paradigm.

On any one trial, the S is presented a three-consonant stimulus. He is given 2 seconds to study the stimulus, after which he is presented a number and is required to read off the number and count backwards by 3's (e.g., 671, 668, 665, etc.) until given a signal to recall the consonant trigram. When the recall signal is given, the S attempts to recall the letters in the correct order. He is then given a brief rest period before the start of the next trial. The material to be remembered can, of course, be varied as can the nature of the interfering task. The outline shared by all variants of the basic paradigm is as follows: presentation of the stimulus, intervening activity, and test for recall of the entire stimulus. The following outline of the unique properties of this paradigm will help to define the experiments of interest to this paper.

The verbal items used in the B-P paradigm are well within the immediate memory span. That is, if recall of the item is tested immediately after presentation, Ss perform essentially without error. The immediate memory span for most verbal materials is about three or four independent units or "chunks" to use Miller's (1956) terminology. Although it is possible to employ stimuli that are longer than the immediate memory span, it would complicate the paradigm with no apparent immediate dividend and there are few such studies.

A key property of the paradigm is that the interfering activity does not necessarily have to be composed of materials that are similar to the to-be-remembered item. Although experiments have shown (to be discussed later) that the more similar the interfering activity is to the to-be-remembered item the greater detriment in performance, similar intervening material is not necessary in order to produce significant forgetting in a matter of a few seconds. The most important feature of the intervening tasks typically used in the B-P paradigm is that overt rehearsal of the item to be remembered is precluded. Whether overt attention to the item and/or covert rehearsal of it are also precluded is a point that will be discussed later.

Another feature of the paradigm is that the time of presentation is sufficiently long for the S to be able to read off the stimulus without error. Whether this allows us to assume that the stimulus has passed through the sensory stores into some more permanent form of storage is not clear, however.

A usual but not necessary part of the paradigm is that there are no experimentally provided cues to help the S recall the item. In all the experiments to be reported, the S's only cue available is recency.

Retention interval. -- The most striking fact about the B-P paradigm is that significant forgetting occurs over quite short intervals of time (3-6 seconds). Although the amount of forgetting is influenced by variables such as amount of time allowed for rehearsal before the onset of interfering activity, number of items to be remembered, amount of rest time before presentation of the stimulus, and position in the experimental session (all to be discussed later), there is some forgetting in almost all experiments if there is interfering activity between presentation and recall. Performance (percent of items correctly recalled) steadily declines over time, reacing non-zero asymptote at the end of about 20-30 sec. The curve is usually exponential in shape: $\%E = K(1-e^{-kt})$. Some typical forgetting curves are displayed in Figure 1.

Rehearsal. -- A second variable of interest is the amount of time given the S to view the stimulus and/or the amount of time allowed for rehearsal. There are two studies in which the rehearsal and/or stimulus presentation time were systematically varied. The first (Peterson and Peterson, 1959) included two conditions: (a) Ss rehearsed the stimuli

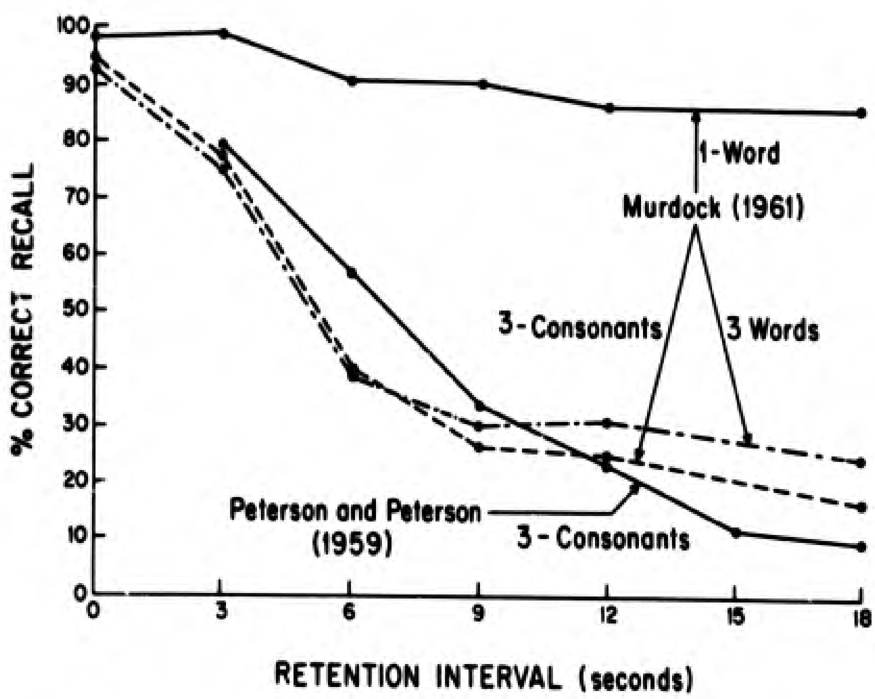

Fig. 1. Percent completely correct recall of 3-consonant trigrams (Peterson and Peterson, 1959) and 1-word and 3-word units (Murdock, 1961). Reprinted from an article by Melton (1963).

(3 consonant trigrams) overtly in time to a metronome (1 repetition per second) before performing an intervening activity (counting backwards by 3's) and (b) Ss were given time to rehearse the stimuli to themselves before the counting, with no specific instructions to rehearse. For the "aloud" group, increasing the rehearsal time improved performance at all levels of interference time, but there was no significant effect of rehearsal time for the "silent" group. Since there was no guarantee that all the Ss in the silent group were rehearsing, the latter result is inconclusive. Two features of the "aloud" group data (Figure 2) deserve additional comment: (a) the apparent asymptote of the curves increases with increasing rehearsal and (b) the curves all have roughly the same shape.

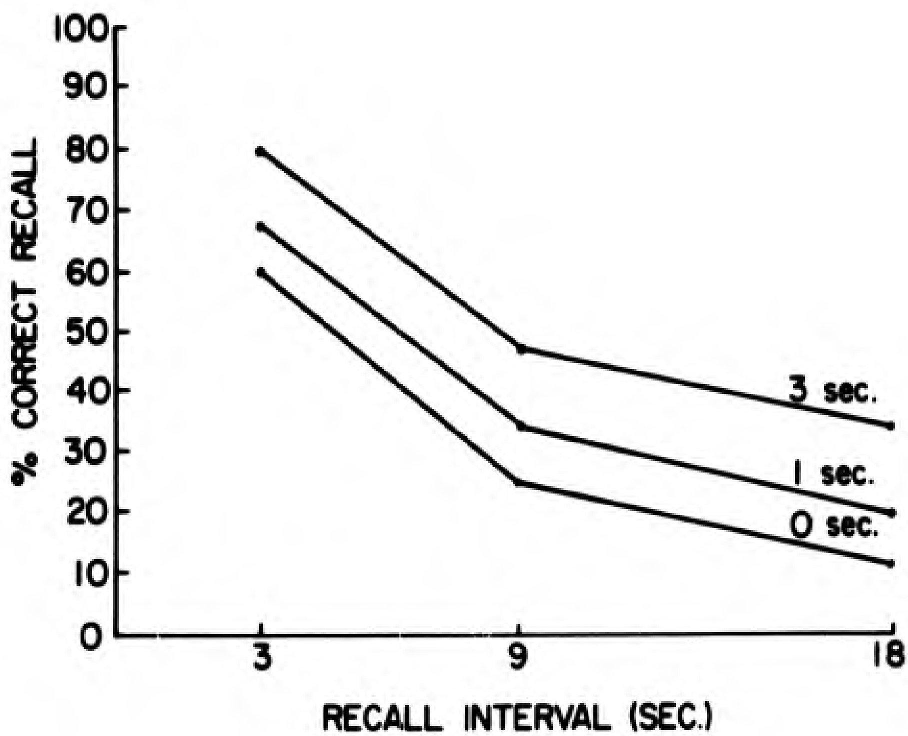

Fig. 2. Percent completely correct recall of 3-consonant trigrams as as function of overt rehearsal of the trigrams (Peterson and Peterson, 1959).

Two studies have been done with strings of digits well over the memory span (Pollack, 1963; Sanders, 1961). In both, significant improvement in performance after an interfering task was obtained if the Ss were given time for rehearsal. Because the stimulus items were much longer than in B-P experiments, further details of the results will not be presented.

A second study (Hellyer, 1962) varied the stimulus presentation time rather than the rehearsal time. In this study, the Ss were not given any instructions to rehearse the items during the extended presentation time, but unlike the results of the Peterson and Peterson "silent" condition, performance did improve with increasing stimulus presentation time. The curves also seem to asymptote at higher levels of performance for longer presentation times but the curves seem to be a bit more sigmoid (S-shaped) in Hellyer's experiment (see Figure 3). A possible cause for this difference is that the rate of the interfering activity is slower in Hellyer's experiment. In sum, the data from the Peterson and Peterson overt rehearsal condition and from Hellyer's experiment are strikingly similar.

Number of units. -- Another variable that effects performance in the B-P paradigm is the number of units or "chunks" to be remembered. In an experiment by Murdock (1961), an effect similar to the effect of increasing rehearsal occurs (see Figure 1). As the number of items increases, performance gets worse. Furthermore, the performance asymptotes at different levels. This is borne out in a more complete parametric study done by Melton (1963) the results of which are shown in Figure 4.

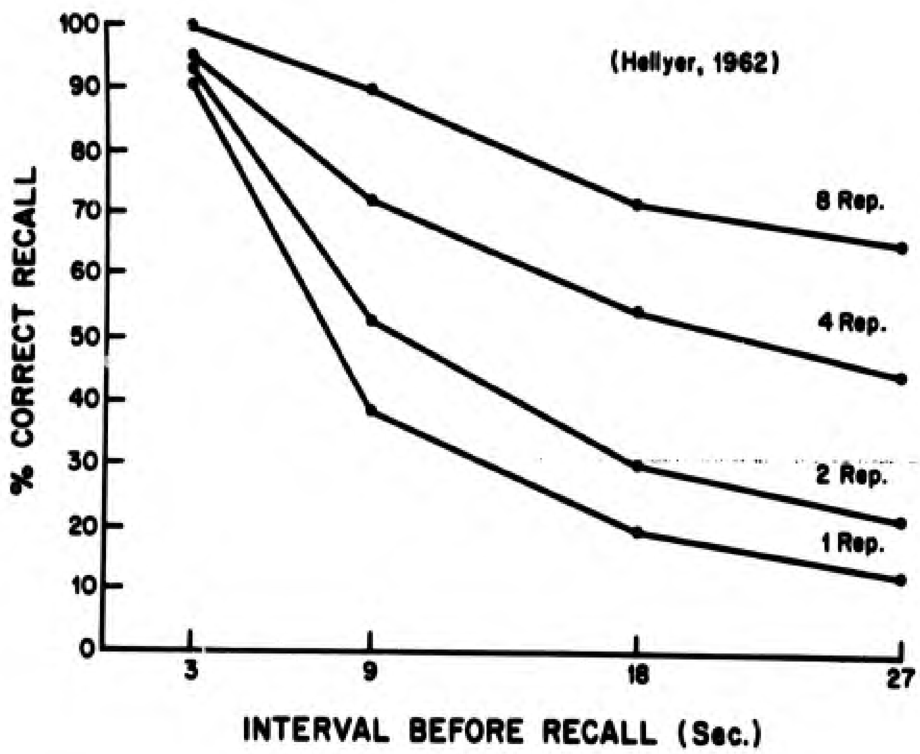

Fig. 3. Percent completely correct recall of 3-consonant trigrams as a function of the frequency of 1-sec. presentations of the trigram before beginning the retention interval (Hellyer, 1962).

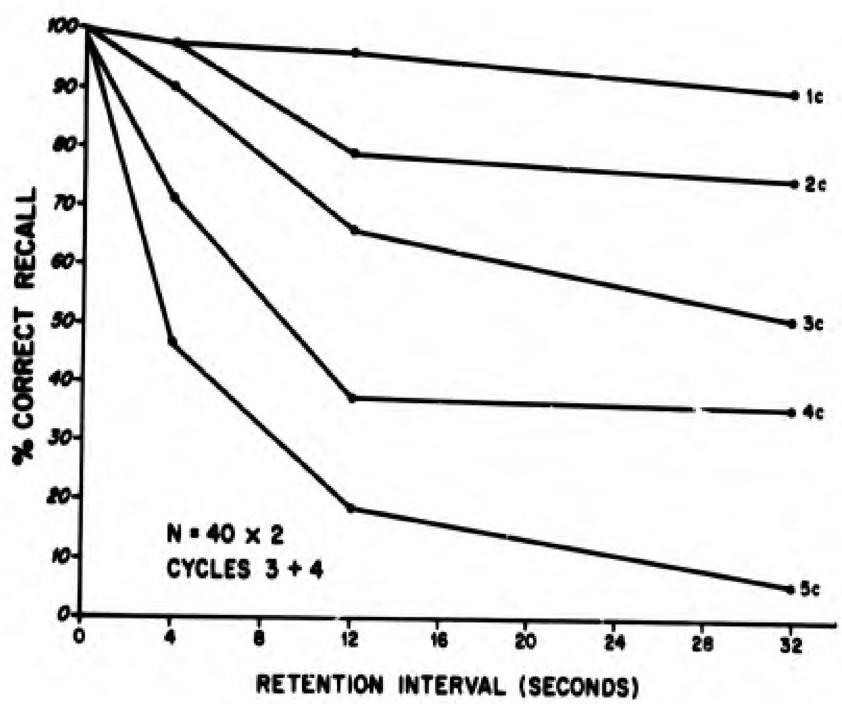

Fig. 4. Percent completely correct recall of units of 1 to 5 consonants (Melton, 1963).

Spaced practice. -- In addition to considering a single presentation of a stimulus followed by interfering activity and attempted recall, one could expand the paradigm to consider multiple presentations of the stimulus separated by interfering activity. A variation of the basic procedure, the double-presentation condition, is a sequence defined by: presentation of the stimulus item, a period of interfering activity, a second presentation of the same stimulus item, a second period of interfering activity, and attempted recall. The most interesting question about the procedure is what effect is produced by increasing the time of first period of interfering activity while holding everything else constant. In other words, does spacing the presentations of the stimuli help or hurt performance? The only published study testing this question (using the B-P paradigm) was conducted by Peterson (1963). The second interference interval was held constant at 6 seconds in Peterson's experiment and the spacing between the first and second presentations were varied from 1 to 11 seconds. Performance (percent of CVC trigrams recalled) improved from 0.66 for a 1 second spacing interval to 0.77 for an 11 second spacing interval. The phenomenon of improved performance with increasing separation of presentations of interfering activity will be referred to as spaced practice improvement (SPI). There was one defect in Peterson's study, however. All the stimuli that Ss saw were presented twice. Therefore, it was possible that the Ss could have partially ignored the stimuli on their first presentation knowing that they would get a second chance. Although it is hard to see how this artifact could have caused the SPI, it is easy to see how it was possible for this artifact to cause an underestimation of the size of the effect.

Prior activity. -- Perhaps the only qualification needed to restrict the sweeping generality that rapid forgetting occurs over a few seconds in the B-P paradigm is the following: at least one trial containing similar material must precede the trial in question without a substantial rest interval between the trials in order for rapid forgetting to take place. On the first trial, or on a trial preceded by a long rest period (2-3 minutes), forgetting is rarely more than 10% even after retention intervals of 20 seconds. (Peterson & Gentile, 1965; Loess & Waugh, 1967). In fact, there is occasionally no forgetting on such trials (Keppel & Underwood, 1962).

Proactive interference (PI) reaches a maximum quickly in the B-P paradigm. In most conditions, performance reaches asymptote in about two or three trials, and the maximum is about five or six trials. The number of trials it takes to reach a maximum is a function of the rest interval (the interval between the end of the recall period for one stimulus and the onset of the next). The general rule is that the longer the rest interval, the more quickly performance reaches a steady state. However, the longer the rest interval, the higher is the level of this steady state performance. It has been argued that PI is stronger and longer lasting than the data show, since there is a learning-to-learn effect (which increases as the number of trials increases) that is tending to cancel out PI (Melton, 1963). However, at the beginning of a later trial block (after a long rest), the learning-to-learn effect should be quite small and yet performance starts at almost 100% on the first trial and asymptotes at about trial 2 or 3 in most experiments (Loess & Waugh, 1967; Loess, 1967). This finding makes it unlikely that PI is being completely masked by a learning-to-learn effect.*

Similarity of the preceding items to the to-be-remembered item is an important variable. When the preceding stimulus items in a B-P experiment are from one semantic class and the to-be-remembered item is from another semantic class, performance on the to-be-remembered item is considerably higher than performance on an item preceded by items from the same semantic class. This phenomenon is "release from PI." Perhaps the most systematic study of "release from PI" has been done by Loess (1967). Loess investigated the recall of word trigrams and found that a change in the taxonomic class (e.g., trees, birds, presidents, etc.) from which the words were drawn produce dramatic release from PI. To give some idea of the magnitude of the effect obtained, in one experiment, performance (percent of word trigrams recalled) was about 30% in the control group on the 13th trial of the experiment while in the group whose category had been changed on the same trial, performance was about 90% (the retention interval was 9 seconds in both cases).

In another experiment by Loess, there was a comparable recovery of performance following a second and third category change in another experiment. Within a block of trials in which the category stayed the same, performance decreased to a level of 30-40% word trigrams recalled by the second or third trial within the block. When the word trigrams from different categories were presented alternately (e.g., Trigram 1 - birds, Trigram 2 - countries, Trigram 3 - presidents, Trigram 4 - trees, Trigram 5 - birds,

---

*It should be pointed out that what is being referred to as proactive interference is generally referred to as proactive inhibition. Since the term, proactive inhibition, has strong theoretical connotations, the more neutral term, proactive interference, will be used to describe the observed effect with no theoretical rationale implied.

Trigram 6 - countries, etc.) performance was about the same on the first presentation from each category as for the first presentation in the experiment, but that on the second presentation performance went down to an asymptotic level that was a bit higher (about 10%) than when the categories were blocked.

The following generalizations summarize the data presented above and are approximately true. First, performance starts out on trial 1 at about 80% - 100% (depending mainly on the retention interval, and depending slightly on the nature of the materials). Second, performance seems to reach an asymptotic level at about the second or third item from the same class with almost all the change occurring between the first and second items. Third, the number of items needed to reach asymptote and the level of asymptote depend on the amount of time between the items of the same class.

Another question that might be asked is what range of situations will produce "release from PI." This has been studied by several workers (Wickens, Born, & Allen, 1963; Loess, 1968; Wickens & Clark, 1968; Bennett, 1969). In the original study of "release from PI", Wickens, et al. (1963) used letters as one semantic class and digits as the other and got large effects comparable to Loess's results discussed above. Making the stimulus perceptually distinctive (e.g., making the stimulus a different color, the background a different color, the stimulus a different typeface, etc.) is another possible manipulation to produce "release from PI." However, there has been little or no release from PI (at most about a 10% difference in performance) with this manipulation, which suggests that at most a small part of "release from PI" can be accounted for by perceptual, attentional, or motivational variables. (Hofer, 1965; Loess, personal communication).

There have been other manipulations that have produced release from PI as well. Bennett (1969) had Ss learn two category labels, each for an arbitrary set of consonants. The labels were well-learned by extensive pre-training on a choice reaction time experiments. In a subsequent Brown-Peterson task, a change of category from one arbitrary set of consonants to the other produced "release from PI." Another interesting manipulation defined the categories in terms of similarity as measured by the semantic differential (Wickens & Clark, 1969). "Release from PI" was also found in this experiment although the difference in performance was only about 10-15%. A complete statement of the conditions necessary for "release from PI" is clearly not possible without still more data.

A dependent variable that has been studied in many of the preceding studies is the percent of intrusions (errors that are incorrect responses rather than omissions). Intrusions were originally computed as a measure of interference from previous items and some researchers have suggested that all errors are caused by intrusions (Keppel & Underwood, 1962). Although the data to be reviewed below cast doubt on that claim, they are inconclusive for two conflicting reasons. In the first place, there could be "covert intrusions", items that intrude but for one of several reasons are not overtly recalled by the S (he "knows it is wrong", or the item is too weak to be coherently recalled). On the other hand, a number of overt intrusions could simply be guesses; that is, the guessed item in no way interferes with the "correct" item, but when the S can not recall the correct item, some previous item is a "good guess". Thus, the intrusion data reported below are unlikely to be a reliable measure of the intruding properties of previous stimulus items.

In the original study of intrusions in STM, Keppel and Underwood (1962) found that the percent of intra-experimental intrusions (IEIs) increased as the retention interval increased. However, the percent of IEIs did not increase in absolute value as much as did the percent of total errors. This is the usual finding in B-P experiments. A somewhat contradictory finding with supra-span materials (Conrad, 1960) was that IEIs increased as the intertrial interval decreased while the number of total errors remained constant.

A more stable finding with intrusion data is that the bulk of intra-experimental intrusions come from the stimulus immediately preceding, and almost all the rest are from the item preceding that item. Although the scores vary from experiment to experiment, a rough characterization of the percent <u>of total intrusions</u> is: 1 back - 60-68%, 2 back - 10-20%, 3 back - 5-10%.

In his "release from PI" studies, Loess (1968) has found that most (80-90%) of the intrusions come from the same category as the to-be-remembered items, and of those, 60-80% come from the most recent item of that category. In addition, the percent of intrusions that come from the most recent item of the same category decreases significantly (10-15%) if there are items between it and the to-be-remembered item.

<u>Nature of intervening activity</u>. -- One experiment studying the similarity of the intervening activity materials to the stimulus items used

consonants (C) and digits (D) as the two classes of materials (Reid, 1967). In the experiment, Ss were presented three stimuli (either DDD's or CCC's) and had to shadow (i.e., read at a fixed rate) either D's or C's as the intervening task. A 2 x 2 factorial experiment was conducted. The major finding was that the retention was poorer when the stimuli to be remembered and the intervening material were from the same semantic class. The groups that shadowed materials similar to the memory materials had about 10% lower scores than the corresponding groups that shadowed materials different from their memory materials. Thus, it seems that similarity of the intervening activity is a significant variable, but not nearly as powerful as the similarity of the preceding items where 50-60% improvements in performance can be found by changing the category of material. In the case of the intervening activity, the S is not making any attempt to learn the intervening material. This difference in intent-to-learn probably accounts for a large part of the above difference. Results similar to Reid's have been reported by Corman and Wickens (1968).

The above experiment also manipulated intervening task difficulty by varying the rate of shadowing (1 unit/second, 2 unit/second, 3 unit/second, and 4 unit/second). Retention decreased as the rate of shadowing increased and the curves appeared to reach different asymptotes.

A prior study similar to Reid's, but using a somewhat different paradigm was conducted by Wickelgren(1966). Similarity was varied by using two categories of letters: consonants with the vowel sound "ē" (e.g., B, C) and consonants with the vowel sound "ĕ" (e.g., F, L). The Ss copied letters that were presented to them aurally. The to-be-remembered item (a single letter) was separated from the previous items by a tone and pause that were a cue that the item was the one to be recalled later. Similarity was defined as the percentage of letters that have the same vowel sound as the critical item. The number and similarity to the to-be-remembered item of the letters both before and after the to-be-remembered item were factorially varied. Significant decrements in performance were obtained both by increasing the number and similarity to the recalled item of either the prior items or the items between the critical item and recall. As would be expected from other studies, the effect of numbers of intervening items was much stronger than the effect of similarity of those items. While increasing the number of prior items from zero to two affected performance, there was little consistent change in performance by adding more prior items. Effects due to prior-item similarity only occurred if there were no similar letters in the succeeding items. Similarly, the effects due to succeeding-item similarity were clear only when there were no similar prior items. Thus, the data suggested that the similarity effect was simply due to the number of confusing (similar) letters, and the effect reached asymptote at about 4 or 5 letters.

Reid (1967) also studied Posner's (1964) measure of task difficulty, information reduction, in the Brown-Peterson paradigm. He used two of Posner's information reduction tasks (adding pairs of digits -- information reduction of 2.7 bits; classifying pairs of digits as above or below 50 and as odd or even -- information reduction of 4.6 bits). He also varied the speed of presentation (1 digit pair/sec, 1 digit pair/0.75 sec). Reid found a linear decrement of performance as a function or rate of information reduction. This finding and all findings relating information reduction to task difficulty must be tempered with the knowledge that this measure can at best be applied to a limited class of tasks. For example, if the task is to multiply each digit pair by 163,724,859, there is no information reduction and yet the task is clearly more difficult than either of the above.

Interpolated task difficulty and the similarity of the interpolated task to the to-be-remembered item do appear to be significant variables in the B-P paradigm. In Reid's experiments, difficulty was the more potent variable. This may have been a function of the particular levels of similarity and difficulty that were selected.

Relation of Brown-Peterson Phenomena to Other STM Phenomena

The central concern of this paper is the possible implications of data from the B-P paradigm with respect to the mechanisms of memory. If regularities in B-P data are limited to that paradigm, then they are likely to result from particular strategies used by Ss in that task. On the other hand, if the regularities generalize in meaningful ways to other memory paradigms, then the B-P data are more likely to reveal basic memorial processes. In the following discussion the effects of retention interval, amont of practice, spacing of practice, and prior activity on STM will be discussed. As mentioned before, the two paradigms that will be discussed in this section are the paired-associate (PA) and free recall (FR) paradigms.

In the PA paradigm, there are two types of trials: study trials, on which the stimulus and response are both presented; and test trials, on which only the stimulus is presented and the S attempts to produce the response. The two major PA procedures are distinguished by the arrangement of study trials and test trials. In the anticipation procedure, a study trial for an item always follows a test trial for that item, while in the study-test procedure, there is no necessary contiguity of study and test trials for an item. In long-term memory research, Ss are typically required to learn a fixed list of paired-associates according to either the study-test or anticipation procedures. A relatively recent innovation designed to permit a clearer assessment of short-term effects than can be obtained from the list learning situation in the continuous PA variation. In the continuous antici-pation method, test-study pairs are presented continuously and the intervals between test-study trials of a given item are the major independent variables of interest. In the continuous study-test method, study and test trials of the various items are presented in a continuous flow, and interval between a study trial of an item and the ensuing test trial is the major independent variant of interest.

The standard FR paradigm is experimentally quite straightforward. The S is given a list of items at a constant rate (typically one item every 1 or 2 seconds) and at the end of the list, he is asked to recall as many of the items in whatever order he wishes.

Retention Interval. -- The intervening activity in the PA and FR para-digms is different in kind from the intervening activity in the B-P paradigm. In both the PA and FR paradigms, the interval between the presentation and attempted recall of any given item is filled with presentations and attempted recalls of other items. Also, since the rate of presentation of trials is usu-ally fixed in these two paradigms, the time of the retention interval is con-founded with the number of intervening items. In spite of these procedural differences, the forgetting curves as a function of the time of intervening activity look strikingly similar to those obtained in the B-P paradigm. An example from the PA literature is shown in Figure 5. The forgetting curve in Figure 5 is derived from a paired-associate experiment by Bjork (1966) and is obtained only from items on which the subject makes a subsequent error. Thus, the fact that the asymptote is at the "guessing level" rather than at the non-zero level observed in the B-P experiments is not inconsistent since the long term memory performance was subtracted out by looking only at items on which a subsequent error was made. The fact that performance goes to "guessing" when the correction is made is certainly not a logical necessity, however. As in the B-P paradigm, the curve is roughly exponen-

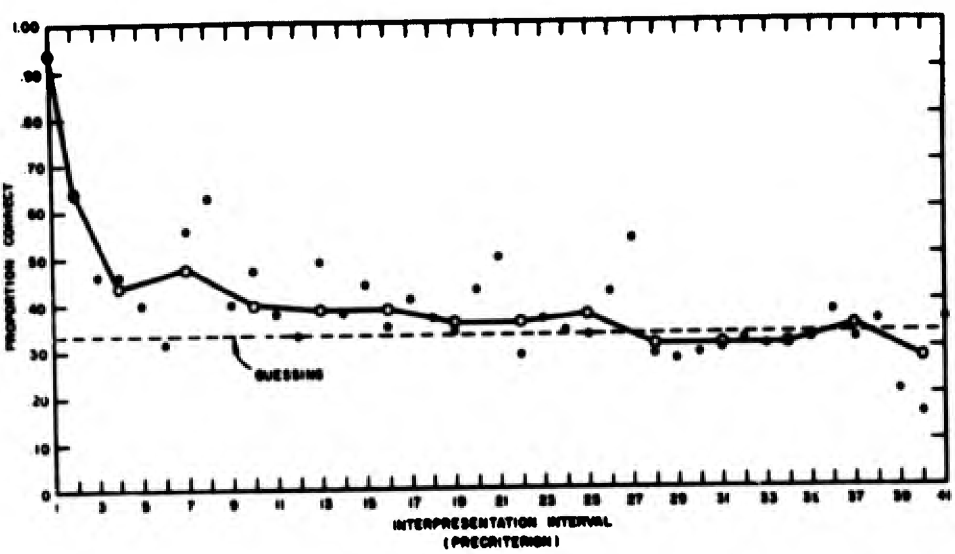

Figure 5. Proportions of Correct Responses Prior to the Last Error as a Function of Interpresentation Interval. (Each open circle plots the average of three successive points.) (Bjork, 1966).

tial and levels off to a very slow decay rate at about 30-40 seconds of interpolated trials.

Curves having approximately the same decay parameters and same general shape have been obtained in both the continuous study-test procedure (Peterson, Salzman, Hillner, & Land, 1962) and the probe procedure (Murdock, 1963). It is interesting to note that the decay times in the two paradigms (BP and PA) are approximately equivalent even though the tasks are quite different.

In free-recall learning, forgetting curves that are reasonably similar to B-P curves are found if we ignore performance for the early items in the list (see Figure 6).

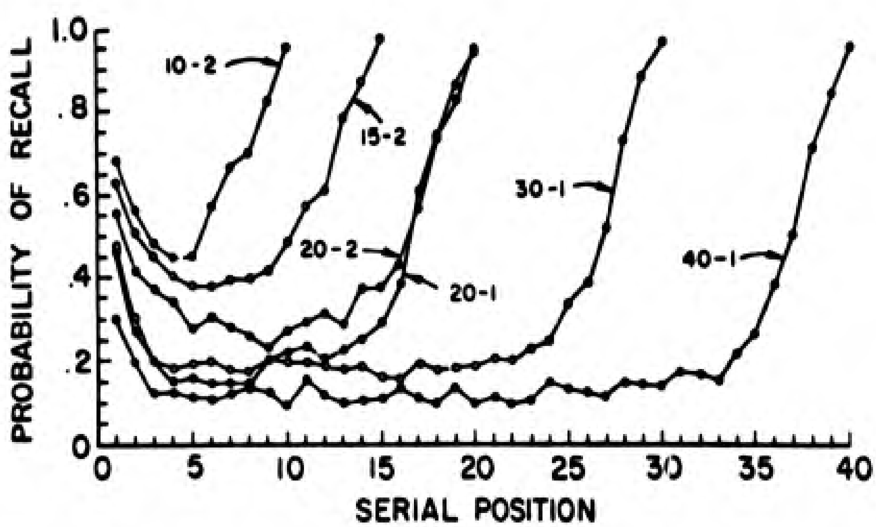

Figure 6. Serial position curves for free verbal recall (Murdock, 1963).

The four most stable findings from FR curves are: (a) performance at the end of the list decreases to an asymptotic level by about the 7th or 8th item from the end of the list (recency effect) and this performance is independent of the size of the list; (b) performance on the first 4 or 5 items is superior to those in the middle of the list (primacy effect); (c) the level of performance in the middle of the list is at a non-zero asymptote and the level lowers as the length of the list increases; (d) if an intervening task is placed between the end of the presentation of the list and the beginning of the recall task (e.g., counting backwards) the recency effect is destroyed in that performance on the last 7 or 8 items is approximately equal to that on items in the middle of the list, whereas performance on items at the beginning and middle of the list is approximately the same as if there was no intervening task (Postman & Phillips, 1965; Glanzer & Cunitz, 1966).

The forgetting curves for the most recent items differ from PA curves in that they are slightly more sigmoid, and they reach asymptote after intervals of 9 items or so (9 seconds in one experiment) between the presentation of the item and the end of the list. It must be remembered, however, that this is not the retention interval, since the retention interval is the interval between the presentation of the stimulus and <u>its</u> attempted recall. Since it has been shown in other experiments that, with no instructions to the contrary, the last few items are recalled first, it is likely that the average actual retention interval for the item nine removed from the end of the list is the 9 seconds quoted above plus several seconds of attempted recall of other words. Thus, the figure of 10-20 seconds for forgetting curves to reach asymptote would seem a reasonable guess for the FR paradigm.

In summary, we can say that forgetting curves in these two paradigms have the same general shape (except for primacy) and decay times of the same order of magnitude as those in the B-P paradigm.

Rehearsal. -- In the PA paradigm, the number of exposures of the S-R pair is typically manipulated rather than the time allowed for rehearsal after any one particular exposure. In such experiments, the asymptote of the forgetting curve increases significantly as the number of study trials increases (e.g., Atkinson & Shiffrin, 1967). Of course, since the exposures of the PA item are generally spaced in contrast to the massed exposures in Hellyer's (1962) B-P experiment, the experiment analogy is not as strong as one would wish.

In the FR paradigm, decreasing the rate of presentation is the experimental analogue of increasing rehearsal time in the B-P paradigm. The situation is complicated for although decreasing the rate of presentation increases the time available for rehearsal, it also increases the retention interval and the rate of processing during the retention interval. The probable trade-off between the latter two variables may explain why the number of interpolated items is a more reliable measure of the retention interval than is time elapsed.

In spite of the possible complex interactions which might result from manipulations of the rate of presentation, some significant findings have emerged. Glanzer (1969) and Murdock (1962) have shown that a slower presentation rate improved performance in all but the recency part of the FR serial position curve. Of further interest is the finding that slowing down the presentation rate improves performance much more when the S is allowed to use the extra exposure time in whatever way he sees fit than when he is forced to use the time to repeat the item just presented over and over again (Glanzer, 1968).

Thus, in PA and FR learning added presentation or rehearsal time improves recall performance. The improvement in performance does not seem to decrease as the retention interval increases, and in the FR case just discussed the improvement seems to increase as the retention interval increases.

Spacing of Practice. -- There has been considerable research on the effects of the spacing of presentations employing both the anticipation and study-test procedures of the PA paradigm. The following paragraphs summarize the major findings.

The most stable findings in all of these experiments is that for reasonably long retention intervals, if presentations of items (study trials) are massed, poorer performance is obtained than if the same number of presentations are spaced. (Peterson, Saltzman, Hillner, & Land, 1962; Peterson, Hillner, & Saltzman, 1963; Bjork, 1966; Young, 1966; Atkinson & Shiffrin, 1967; Rumelhart, 1968). In an extreme case, (Greeno, 1964) two presentations that were massed were not significantly better than a single presentation.

The standard procedure is to vary the number of items between two study trials (the spacing interval) and to hold constant the number of items between the second study trial and a test for recall (the retention interval). Although performance is worst when the spacing interval is zero intervening items, performance does not go up indefinitely as the spacing interval is increased. Three studys that employed long spacing intervals (Peterson, et al., 1963; Young, 1966; and Atkinson & Shiffrin, 1967) found that spaced practice improvement reached a maximum for spacing intervals in the range of about 30-60 seconds, and decreased from the maximum for longer spacing intervals. If performance is plotted as a function of spacing interval, the curve rises in a gradual but negatively accelerated way and then starts to decline (see Figure 7). The maximum improvement of spaced practice over massed practice is usually about 10-15% in PA experiments.

One qualification of the generality of the spacing effect in PA learning is that massed practice appears to be slightly better than spaced practice if the retention interval is one or two intervening items (Peterson et al., 1963). Rumelhart (1968) also found significant interactions between the spacing interval and the retention interval.

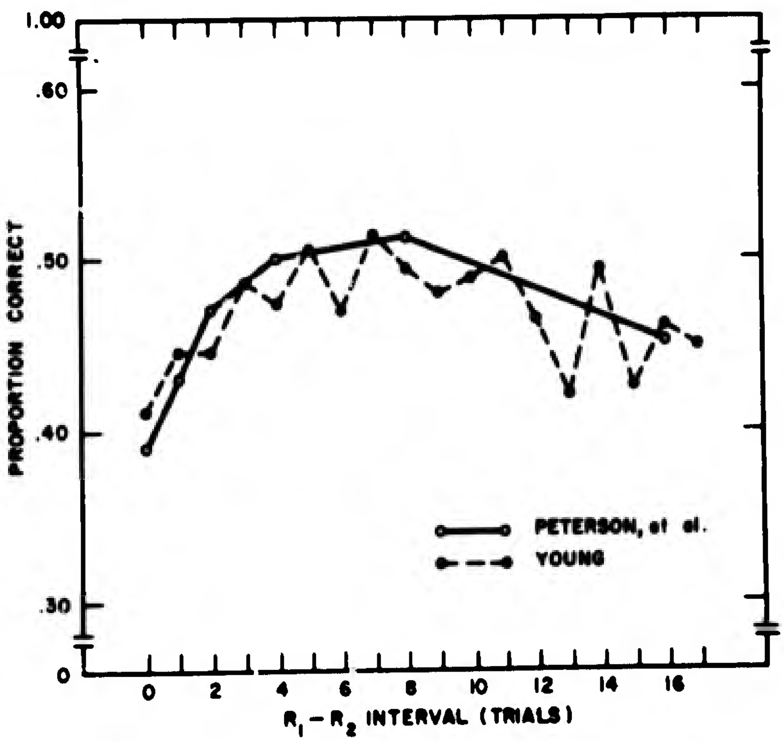

Figure 7. Proportions of correct responses as a function of the spacing of two study trials (Peterson, et al., 1963; Young, 1966). Reprinted from an article by Bjork (1966).

Bjork and Abramowitz (1969) studied the effects of spacing for more than two trials using the anticipation procedure. They held the total interval between trials 1 and 3 for a given item constant and varied the position of trial 2. The behavior of principal interest was performance on the test at trial 4. The interval between trials 1 and 2 will be referred to as the first interval and the interval between trials 2 and 3 will be referred to as the second interval. Performance was found to be best when the two intervals were equal and performance decreased the more the two intervals were unequal. Also, performance was the same when the first interval was equal to X and the second interval was equal to Y as when the first interval was equal to Y and the second interval was equal to X.

In free recall learning, the primacy and recency effects complicate the study of the effects of spacing. To counteract this difficulty, experiments have been designed such that the first and second presentations of the item of interest are in the middle of the list where performance is quite stable. The original study of spacing in the free recall paradigm (Waugh, 1963) found no effect of manipulating the spacing interval. However, an extensive series of experiments by Melton and Shulman (1967) and another experiment by Glanzer (1969) have consistently found a significant improvement due to spacing. Melton and Shulman were unable to isolate any variables accounting for the discrepancy between their results and Waugh's results. The order of magnitude of the spacing effect is about the same as in the PA data (about 10%-15%).

Prior Activity. -- A full discussion of proactive interference (PI) in long-term memory (LTM) can be found in Underwood (1957). One point that should be made is that proactive interference discussed in the context of the Brown-Peterson paradigm is quite different from PI in LTM experiments. In B-P experiments, PI reaches asymptote at 2-5 trials while in LTM experiments, PI does not reach asymptote after dozens or even hundreds of trials. Thus, in the following discussion, it should be remembered that PI in STM paradigms is not necessarily the same phenomenon as PI in LTM paired-associate paradigms.

The PA paradigm that is most suited for investigating PI over short retention intervals is the PA probe technique (Murdock, 1963). In this technique, each item of a PA list is presented once in a series of study trials, one stimulus item is selected as a test (probe) and the S is asked to supply the correct response item. In the original study, Murdock (1963) factorially varied the number of items prior to the critical item and the number of items following the critical item (the retention interval). He found that there was a decrement due increasing the number of prior items, but as in the B-P studies, the effect reached asymptote at 1 or 2 prior items. This result held for all retention intervals. Furthermore, there was no decrement in performance on later lists, suggesting that a prior list did not interfere with the current list. There were also almost no intrusions from prior lists in any of the experiments in Murdock's article. The evidence, therefore, suggests that PI in the probe PA experiments results from events occurring a short while before the to-be-remembered item.

An interesting variation on the PA probe experiment has been developed by Bjork (1969). In Bjork's paradigm, Ss were given a series of PA pairs as in Murdock's experiment, but at a certain point, the Ss were given a cue that they could forget all the items previous to that cue (the "forget" items). In this paradigm, one can independently vary the number of previous "forget" items, the number of items after the "forget" cue but prior to the critical item (the "remember" items), and the number of items succeeding the critical item (the retention interval). In one-fourth of the lists there was no "forget" instruction and since such lists were interspersed randomly among the others during the experimental session, Ss were forced to pay attention to the "forget" items until the instruction to forget. Proactive effects similar to those in the Murdock study discussed previously resulted from "remember" items prior to the critical item, but there was no proactive effect owing to the "forget" items. Thus, previous items seem to interfere only when the S is still trying to remember them in this paradigm. In the B-P paradigm, however, the S is never trying to remember items past the trial on which they are presented and tested and yet there is a strong PI effect.

The study of PI in FR is more complicated (as is the study of the other variables of interest in this review) since the paradigm has much less structure. The basic data bearing on PI in FR have been mentioned previously, namely the "primacy effect" and the lowering of performance in the middle of the list with increasing list length. The "primacy effect" (items at the beginning of the list are recalled better than those in the middle) might result from several factors. The first items of the list (and especially the first item, which has the most pronounced primacy effect) are probably encoded better because the S is less fatigued. Furthermore, at the time they are presented, he can give his individual attention to encoding them, which he probably does less efficiently with items in the middle of the list. In addition, the situation is set up from the S's standpoint such that the strategy of trying to remember the first few items probably serves as an anchor to help him recall subsequent items. Thus, the primacy effect cannot be cited unambiguously as evidence for specific interference from previous items (i.e., interference in the holding or recall stages of memory).

The lowering of performance in the middle of the list, on the other hand, does seem to be less ambiguous evidence for some holding or recall interference. Since performance in the middle of the list is asymptotic, the retention interval per se is relatively unimportant for these items. Therefore, there seems to be no plausible non-interference explanation of why the 10th item of a 20 item FR list should be recalled better than the 10th item or the 30th item of a 40 item list.

Summary

The following statements, though a bit oversimplified, summarize the main effects discussed above.

1) In several types of STM experiments, the observed forgetting curves reach non-zero asymptotes in the range of 15-30 seconds. These forgetting curves are usually exponential in shape though they are occasionally sigmoid shaped.

2) Increasing stimulus presentation time or the time available for rehearsal improves performance across all retention intervals in the B-P paradigm. Although increasing the number of study trials in PA experiments and slowing presentation rate in FR experiments are somewhat different manipulations, they produce the same qualitative effect.

3) Increasing the spacing of repetitions of a given item in all paradigms produces a monotonic improvement in performance for spacing intervals of 0 to 30 seconds. For spacing intervals of greater than 30 seconds, there is some attenuation of the above effect. The longer spacing intervals have been studied only in the PA paradigm. In addition, there is an interaction between spacing interval and retention interval (massed practice is better for short retention intervals and spaced practice is better for long retention intervals), although this too has been studied only in the PA paradigm.

4) Although retention decreases with increasing similarity between the stimulus and the intervening activity material, similarity between the interpolated material and the stimulus item is not necessary to produce significant forgetting in the B-P paradigm. Retention also decreases as the difficulty of the intervening task increases. The effects produced by manipulating the nature of the intervening items have not been systematically studied in PA and FR paradigms.

5) The presentation of prior similar materials drastically worsens performance in the B-P paradigm. In fact, with no prior similar items, there is little forgetting. However, the effect reaches asymptote after very few trials (usually 2 or 3). A similar effect occurs in PA probe experiments, although the decrement in performance is not nearly as dramatic. In FR learning, the decrement has not reached asymptote even after 10 to 15 previous items. In the PA probe paradigm, previous items that the S has been told he can forget have no apparent effect on retention of later items.

6) The number of units or "chunks" in the stimulus item to be recalled in the B-P paradigm is a significant determinant of performance. The asymptotes of retention curves are lower, the larger the number of units is in the stimulus item. (It is not obvious what the exact analogue of this manipulation would be in the other two paradigms.)

7) The more errors that are made, the larger the absolute number of intra-experimental intrusions in the B-P paradigm. (However, there is an experiment Conrad, 1960, in which intrusions increased while errors remained constant.) There is no consistent finding with respect to the percent of all errors that are intrusions and there are some theoretical reasons to believe that intrusions are not a particularly reliable measure of interference.

The preceding summary statements, hopefully, make reasonably clear what we do and do not know about short-term memory. They cover most of the major regularities in STM. What we lack is any kind of systematic knowledge of more detailed interactions of the data and, on another level, a theory to predict such details adequately.

The Present Experiments

The experiments to be reported focus on three variables, rehearsal, interference, and spacing of practice. They exclude from consideration variables that pertain to the content of materials -- similarity of materials to each other, difficulty of intervening task, and number of chunks in the stimulus.

More specifically, the experiments are designed to study: (a) the interaction of rehearsal time with interference time, in a design nearly identical to that of two studies discussed earlier (Peterson & Peterson, 1959; Hellyer, 1962); (b) the interaction of rehearsal time with spacing and retention intervals; (c) the effect of rehearsal on the likelihood of a correct response on a second recall test given a correct response on a first test; (d) the interaction of rehearsal time of an interfering item with rehearsal time of the to-be-remembered item and the retention and spacing intervals.

CHAPTER II

THE EXPERIMENTAL PROCEDURE

Two experiments were designed and conducted to assess the interactions of rehearsal, intervening activity, and prior activity. Both experiments employed variations of the basic Brown-Peterson paradigm outlined in Chapter I. The experiments are described below in turn following a description of the experimental materials and the experimental apparatus common to both experiments.

Materials

The to-be-remembered items in both experiments were trigrams constructed from three common four-letter nouns. The 324 words used in the experiment were selected (on an intuitive basis) to be the most common four-letter non-homophonic words. The sets of three words were selected by intuition to have as little relationship by meaning as possible. Some of the trigrams were taken from a set of such trigrams collected by Noyd (1965). The items were also constrained so that (a) no two words in a trigram began with the same letter, (b) no two words in the same trigram or in two adjacent items rhymed; (c) there were no obvious associations among words in adjacent tri-trigrams, and (d) no word was ever used in more than one trigram. The actual items are listed in the Appendix.

Apparatus

The apparatus was a high-speed (change time was less than .05 seconds) memory drum. The stimulus words, interference numbers, and recall, ready and rest instructions appeared in the same window. The counting was paced by a buzzer attached to the memory drum apparatus. The timing sequences in the experiment were controlled automatically by a high-speed tape reader reading a prepunched tape. Each advance of the memory drum could be clearly heard by the S and supplied an additional cue for the advance of the drum. The time that the memory drum took to advance was negligible compared to the times used in the experiment. A fan in the room provided a constant background noise in order to minimize random disturbances.

Method: Experiment I

Subjects

The S were 72 undergraduate and graduate students at the University of Michigan. They were drawn from a pool of Ss who had expressed interest in being paid Ss in psychology experiments. Although some of the Ss had been in verbal learning experiments, none had been in an experiment in which groups of words had been used as stimuli. Therefore, all Ss were naive with respect to the items, and no S had been in a Brown-Peterson paradigm experiment.

Experimental Design

There were three tasks used in Experiment I. In all tasks, the ready period was simply the word, "READY", exposed for 2 seconds. The presentation period was a 2 second exposure of a word trigram. The rehearsal periods following a stimulus presentation were signalled by a drum advance to a blank space appearing in the window. Interference periods were marked by the appearance of the three-digit number for 1 second followed by a rotation of the drum to a blank space. Ss were required to count from the number presented and to continue counting while the window remained blank.

During this whole temporal interval, his counting was paced by a buzzer that sounded at a rate of one beat per 3/4 second. The S was expected to count at the rate of one three-digit number per beat. All recall periods were 8 seconds; during the first 2 seconds, an appropriate cue word for recall was exposed on the drum, and during the last 6 seconds the drum advanced to a blank space. The rest period was the word, "REST", exposed for 5 seconds.

The three tasks can be described by the following sequences of activity.

Task I (Simple): Ready, Presentation, Rehearse, Interference, Recall, Rest.

Task II (Double-Presentation): Ready, $Presentation_1$, $Rehearse_1$, $Interference_1$, $Presentation_2$, $Rehearse_2$, $Interference_2$, Recall, Rest.

Task III (Double-Test): Ready, Presentation, Rehearse, $Interference_1$, $Recall_1$, $Interference_2$, $Recall_2$, Rest.

In the Double-Presentation (DP) Task the trigram was presented twice. In the double-test (DT) task the trigram was tested twice. The words, "RECALL ONE", were used as a cue for the first recall interval in the DT task. All other recall periods in the three Tasks were indicated by the word, "RECALL". The distinctive cue, "RECALL ONE", was used to inform the S in the DT task that a second test was forthcoming.

In all three tasks, the rehearsal and interference intervals were factorially varied, to wit:

Task I (S): Rehearsal: 0, 3, 6, and 9 seconds; interference: 1, 4, 7, 10, 16, and 22 seconds. Thus, there were 4 x 6 = 24 conditions in Task I.

Task II (DP): $Rehearsal_1$: 0, 3, and 6 seconds; $interference_1$: 7 and 22 seconds; $rehearsal_2$: 0, 3 and 6 seconds; $interference_2$: 10 and 22 seconds. Thus there were 3 x 2 x 3 x 2 = 36 conditions in Task II.

Task III (DT): Rehearsal: 0, 3 and 6 seconds; $interference_1$: 7 and 22 seconds; $interference_2$: 10 and 22 seconds. Thus there were 3 x 2 x 2 = 12 conditions in Task III.

The 72 conditions of the three tasks were run in a randomized design in which the S was not given any cue as to which task he was about to do. Thus, the S did not know which task he was doing until he finished the first interference period at which time he saw either "RECALL", or a repetition of the stimulus item, or "RECALL ONE", which identified Task S, Task DP, or Task DT, respectively. This precedure was employed (rather than blocking the conditions by task) in order to (a) insure that the S was doing the same thing during the first rehearsal period of Task DP that he was doing during the rehearsal periods of Tasks S and DT, and (b) equate in a simple way for average level of practice in all conditions.

The exact design can best be described by explaining how it was constructed. Seventy-two sets of stimulus words were assigned randomly to 72 ordinal positions honoring the constraints noted in the materials section. The 72 conditions were randomly assigned to the same 72 ordinal positions with the constraint that the probability that a DP condition followed any condition was 1/2.

The first S was presented with the sequence of trials generated by the above randomization procedure. The presentation sequences for the other Ss were derived from the basic initial randomization by keeping the stimulus items and the numbers used for interference in the same ordinal positions and changing the ordinal position of the conditions. To get the ordering for S 2, the conditions were cyclically permuted within tasks, with the 36 DP conditions being permuted in the opposite order from the 24 S conditions and the 12 DT conditions (i.e., the nth DP condition to appear for S 1 was moved to the ordinal position of the n-1 DP condition for S 2, while the S and DT conditions were moved back to the ordinal position of the succeeding condition of the same task). This procedure was repeated to generate the trial sequence for all 72 Ss with the following exception: at the 25th ordering 12 of the S conditions were interchanged with the 12 DT conditions, and at the 49th ordering the remaining 12 S conditions were interchanged with the 12 DT conditions.

The results of the above procedure can be summarized as follows: (1) the ordinal position was completely confounded with stimulus differences and number differences; (2) ordinal position, stimuli, and interference numbers were completely counterbalanced across all conditions with the following restriction: each of the 36 DP conditions employed each of 36 of the 72 stimulus items twice across the 72 Ss and the 36 S and DT conditions employed each of the other 36 stimulus items twice across the 72 Ss.

Procedure

The S was read instructions which detailed the memory tasks he would face in the experiment. After the instructions were read, the S was practiced at counting backwards. If it appeared that the S could not attain the desired counting rate, he was told to count as fast as possible. In the instructions, it was emphasized that the recall period was over when "REST" appeared and that they should not continue trying to recall thereafter. Ss were encouraged to rehearse in their own most natural manner during the rehearsal periods in the experiment.

In order to familiarize the Ss with the three memory tasks, they were given three practice trials consisting of a single exemplar of each of the three tasks. Following the three practice trials, Ss were told the experiment was about to begin and they were given three more trials including a single instance of each task and asked if there were any questions (there were none). The experiment proper proceeded immediately. The items used in the practice trials were three-word items drawn from the same pool as the regular items. All the rehearsal times and interference times employed during the experiment proper appeared at least once in the total "warm-up" session.

The 72 experimental trials were run continuously: a ready period immediately followed the rest period of the preceding trial except for a 2 minute break taken between trials 36 and 37. The total experimental session lasted about one hour and 15 minutes. At the end of the session, the S was asked a few questions about what he thought he was doing during the experimental session.

Method: Experiment II

Subjects

The Ss were 72 undergraduate and graduate students taken from the same population used in Experiment I. No S participated in both experiments.

Experimental Design

The experimental design was identical to that in Experiment I except for the nature of the three tasks.

Task I (Simple): Identical to Task I in Experiment I -- Ready, Presentation, Rehearsal, Interference, Recall, Rest.

Task IV (Forget): Ready, $Presentation_1$, $Rehearsal_1$, $Interference_1$, $Presentation_2$, $Rehearsal_2$, $Interference_2$, Recall, Rest.

Task V (Control): Ready, $Interference_1$, Presentation, Rehearsal, $Interference_2$, Recall, Rest.

The difference between Task IV (Forget) in this experiment and Task II (Double-Presentation) in Experiment I was that in Task IV the two trigrams were different. The S's instructions were to remember the last trigram. Thus, at the point at which he saw the second trigram, he knew he could forget the first item and remember the second.

In Task V(C), the first interference period consisted of seeing one three-digit number, counting backwards for 5 seconds, seeing another three-digit number, and counting backwards for a varied number of seconds. The 5 second initial counting period was chosen to be equal to the average first presentation and rehearsal periods in Tasks II (DP) and IV (F). This task served primarily as a control condition for the Forget Task, and secondarily as a control for the Double-Presentation Task.

The times used for the presentation, recall, rest, and ready periods were the same as in Experiment I. The other times were varied as follows:

Task I (S): Rehearsal: 0, 3, 6 and 9 seconds; interference: 1, 4, 7, 10, 16 and 22 seconds. Thus there were 4 x 6 = 24 conditions in Task I.

Task IV (F): $Rehearsal_1$: 0, 3, and 6 seconds; $interference_1$: 7 and 22 seconds; $rehearsal_2$: 0, 3, and 6 seconds; $interference_2$: 10 and 22 seconds. Thus there were 3 x 2 x 3 x 2 = 36 conditions in Task IV.

Task V (C): Interference$_1$ = 7 and 22 seconds (not including the constant 5 second initial interference time); rehearsal = 0, 3 and 6 seconds; interference$_2$ = 10 and 22 seconds.

A randomization procedure identical to that used in Experiment I was carried out with F conditions substituted for DP conditions and C conditions substituted for DT conditions. The same stimulus items appeared in the same order as in Experiment I, and 36 new trigrams were constructed and used as the Presentation$_1$ items (or items to be forgotten) in F conditions.

It should be pointed out that a combined rather than a blocked design was even more mandatory in Experiment II than in Experiment I. That is, Task I (S) trials were needed as "catch" trials for Task IV (F) trials. Since the S did not know whether a trial was an S condition or an F condition until after the first interference period (when either "RECALL" or the second set of words appeared) the S could not ignore the first trigram in F conditions until the second trigram appeared.

Procedure

The procedure was identical to that used in Experiment I.

CHAPTER III

RESULTS

In this chapter the recall data from Experiments I and II are discussed by task rather than by experiment. At the end of the chapter, the intrusion data from all tasks are presented.

Task I (Simple)

The recall data from Experiments I and II were analyzed both in terms of the proportion of constituent words recalled correctly, word retention data, and in terms of the proportion of entire word trigrams recalled correctly, trigram retention data. In the former case, free recall scoring was used, i.e., if a word was recalled, it was scored as correct independent of whether it was recalled in the same serial position as it was presented. In the latter case, a trigram was counted as correct only if all three words in a trigram were recalled but not in the correct order.

In Figure 8, the trigram retention data from the simple conditions are shown pooled across both experiments since the data from the two experiments were in good agreement. Retention in Experiment I was a bit higher ($.01 < p < .05$), but the interactions between experiment and rehearsal and/or interference were non-significant.

The difference in performance level between the two experiments could be explained by an occasional tendency on the part of Ss in Experiment II to guess that the item before him was an item to be forgotten (i.e., from an F condition) and therefore to pay minimal attention to it. The fact that this difference is as small as it is suggests that such strategies were not employed frequently, if at all. The Ss in Experiment I and the Ss in Experiment II

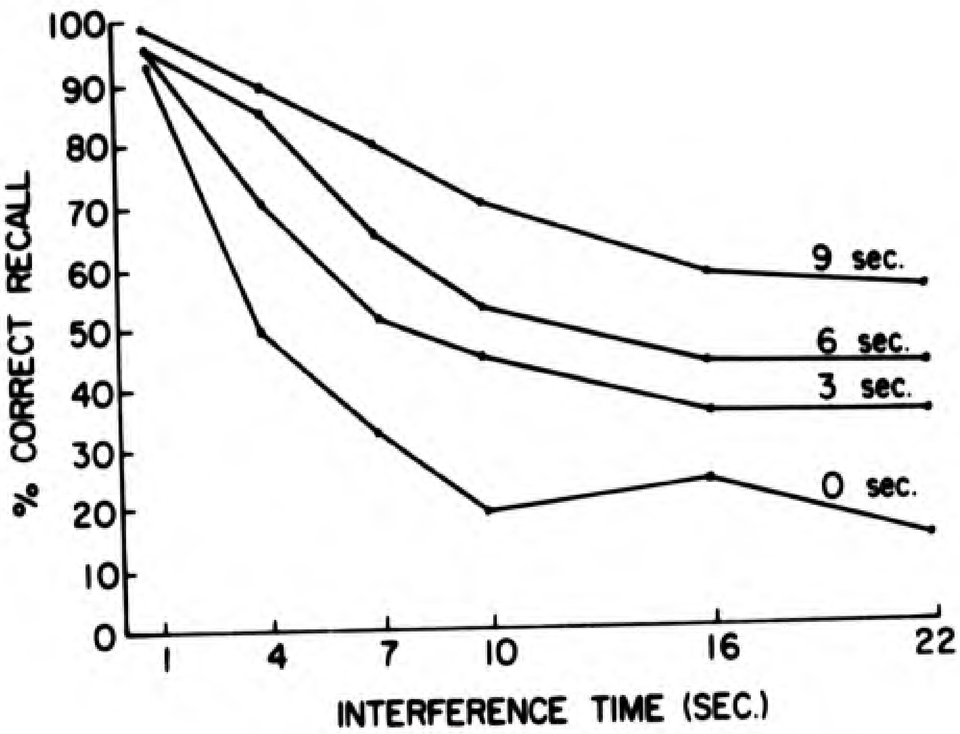

Figure 8. Percent completely correct recall of 3-word units as a function of covert rehearsal.

were intentionally equated for approximate university grade-point average in order to minimize any differences across experiments owing to S differences.

The main trends in the trigram retention data shown in Figure 8 are quite clear. There are large and significant differences due to changes in both rehearsal and interference times. The forgetting curves for each rehearsal time interval appear to reach an asymptote with increasing retention interval. The curves are clearly more sigmoid in shape the greater the rehearsal time.

One surprising result is that the curves in Figure 8 for the 9 second and 6 second rehearsal conditions are quite different from each other, whereas the curves in Figure 8 for the 6 second and 3 second rehearsal conditions are quite close to each other. One might have expected that rehearsal would have marginally decreasing effectiveness. Here again, an artifact may have played a role. Since there was no 9 second first rehearsal time with the Double-Presentation and Forget Tasks, the Ss may have been rehearsing a bit harder after 6 seconds in the unlikely event that they discovered this flaw in the design and had reasonably accurate internal time clocks. In a previous study (Pollack, 1961) using supra-span materials, marginally decreasing effectiveness of rehearsal time was also not the rule, so that the results of the present experiments need not have been artifactual.

The recall data scored in terms of the percent of individual words recalled, Figure 9, have the same general flavor as the trigram retention data shown in Figure 8. The curves in Figure 9 are possibly a bit more linear than the curves in Figure 8 and the performance asymptotes are a bit less marked. As in Figure 8, the curves become increasingly sigmoid in shape with increasing rehearsal times.

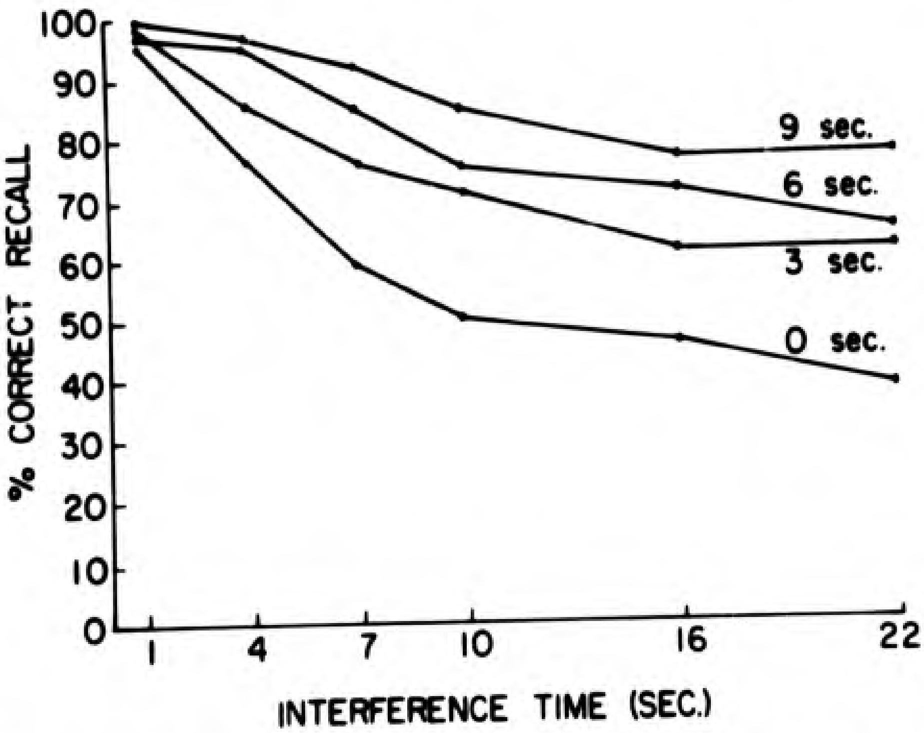

Figure 9. Percent words recalled as a function of covert rehearsal.

Task II (Double-Presentation)

The most striking feature of the trigram retention data (Table 1) is the improvement in performance with spacing of practice. Perhaps the magnitude of the difference can best be captured by the following statistic: the ratio of errors in the conditions where the first interference interval is 22 seconds (spaced practice) to the errors in conditions where it is 7 seconds (massed practice) is three to five.

The effects on retention of the first rehearsal interval, the second rehearsal interval, and the retention interval (the second interference interval) are all large and significant. The beneficial effect of additional rehearsal time is of about the same magnitude whether it is added to the first rehearsal period or to the second rehearsal period. However, retention is somewhat better when the additional rehearsal time was in the first rehearsal period. This contrast is significant ($p < .05$).

There appears to be no meaningful interaction between the two interference times -- spaced presentations are more effective than massed presentations regardless of whether the retention interval is 10 or 22 seconds. There is no exact comparison between Tasks DP and S which would equate the total stimulus presentation plus rehearsal time. Perhaps the best approximation is a comparison of the 9 second rehearsal condition (11 seconds stimulus presentation plus rehearsal time) in Task S with the conditions in Task DP that allow 6 seconds total rehearsal time (10 seconds stimulus plus rehearsal time). Retention in DP massed conditions is 15% better and retention in DP spaced practice conditions is 25% better than the comparable S conditions (the figures are differences in per cent trigrams recalled correctly). The difference in retention between spaced and massed practice appears to be relatively independent of both rehearsal times.

TABLE 1

TASK II (DP) RETENTION DATA

Second Rehearsal and Interference Intervals (sec.)		First Rehearsal and Interference Intervals (sec.)					
		$R_1=0$		$R_1=3$		$R_1=6$	
		$I_1=7$	$I_1=22$	$I_1=7$	$I_1=22$	$I_1=7$	$I_1=22$
		Percent of Trigrams Correctly Recalled					
$R_2=0$	$I_2=10$	68.1	80.6	84.7	86.1	88.9	94.4
	$I_2=22$	58.3	68.1	72.2	83.3	73.6	87.5
$R_2=3$	$I_2=10$	80.6	90.3	83.3	90.3	86.1	97.2
	$I_2=22$	58.3	79.2	73.6	81.9	73.6	88.9
$R_2=6$	$I_2=10$	88.9	93.0	90.3	91.7	88.9	94.4
	$I_2=22$	68.1	76.4	75.0	93.0	80.6	87.5
		Percent of Individual Words Correctly Recalled					
$R_2=0$	$I_2=10$	86.1	93.1	93.5	96.8	94.4	98.1
	$I_2=22$	75.5	81.9	81.5	92.1	87.0	95.8
$R_2=3$	$I_2=10$	92.6	96.8	92.6	96.8	95.8	99.1
	$I_2=22$	81.0	90.3	87.5	91.7	85.2	95.4
$R_2=6$	$I_2=10$	94.9	97.7	94.9	97.7	95.4	98.1
	$I_2=22$	84.7	92.1	86.1	96.8	88.9	94.9

The word retention data is quite similar to the trigram retention data (see Table 1) except that there appears to be more noise in the word data. The improvement in performance due to increased spacing is even more marked for the word retention data than for the trigram retention data as there are over twice as many errors in massed conditions as in spaced conditions. The only major difference is that there appears to be less difference in the effectiveness of the two rehearsal intervals in the retention data. (The difference is in the same direction as in the trigram retention data -- the first rehearsal produces a bigger change -- but it was not significant.)

Task III (Double-Test)

Both the trigram retention data and word retention data from the first recall test in Task DT agrees quite closely with the corresponding S conditions (see Table 2). This is to be hoped, since the only difference between the Task S recall period and the first recall period of Task DT was that "RECALL" appeared as the recall cue for Task S and "RECALL ONE" appeared as the recall cue for Task DT. There is some forgetting between the two tests for the trigram retention data, but the amount is small (7% not recalled on test 2 given correct recall on test 1) and insignificant, and there is a lesser amount (5%) of "spontaneous recovery"(i.e., correct recall on trial 2 given a correct recall on trial 1). For longer second interference intervals, there is a bit more forgetting, but the total amount of forgetting is so small that the difference is insignificant. In fact, there is actually slightly better overall performance on the second recall than on the first for the word retention data.

TABLE 2

TASK III (DT) RETENTION DATA

Rehearsal Interval (sec.)	Interference Intervals (sec.) $I_1=7$ $I_2=10$	$I_1=7$ $I_2=22$	$I_1=22$ $I_2=10$	$I_1=22$ $I_2=22$
Recall 1	Percent of Trigrams Correctly Recalled			
0	26.4	22.2	20.8	18.1
3	44.4	50.0	27.8	34.7
6	66.7	63.9	45.8	41.7
Recall 2	Percent of Trigrams Correctly Recalled			
0	26.4	19.4	22.2	18.1
3	51.4	48.6	29.2	37.5
6	66.7	59.7	47.2	41.7
Recall 1	Percent of Individual Words Correctly Recalled			
0	56.5	49.5	43.5	40.7
3	65.7	72.2	57.4	56.9
6	82.4	80.6	62.0	64.4
Recall 2	Percent of Individual Words Correctly Recalled			
0	59.7	48.1	49.1	41.7
3	67.1	70.4	57.4	63.0
6	83.8	75.0	63.0	66.2

To summarize: there is little or no forgetting from test 1 to test 2 in this experiment, and performance levels on the first test agree well with the comparable conditions in Task S.

Task V (Control)

Since the data from Task C form a baseline against which to compare the data from Task F, the Task C data will be presented first.

The data from Task C can be summarized fairly succinctly (see Table 3). The overall retention (both for trigrams and words) in this task is comparable to that in the comparable S condition of either experiment except that it is higher. That is, the effects of rehearsal and the retention interval seem to be the same as in Task S. (The difference in per cent of trigrams recalled is 12.3% between Task C and the equivalent conditions in Task S, Experiment I and 15.8% between Task C and Task S, Experiment II.

TABLE 3

TASK V (C) RETENTION DATA

Rehearsal Interval (sec.)	Interference Intervals (sec.) I_1=12 I_2=10	I_1=12 I_2=22	I_1=27 I_2=10	I_1=27 I_2=22
	Percent of Trigrams Correctly Recalled			
0	31.9	34.7	41.7	25.0
3	50.0	29.2	62.5	51.4
6	79.2	50.0	75.0	55.6
	Percent of Individual Words Correctly Recalled			
0	65.7	54.2	70.8	49.1
3	73.1	56.5	81.9	69.4
6	90.7	67.1	85.2	73.6

Comparing performance on the two tasks is a bit tricky, though, since Ss in C conditions knew that they were going to get only one presentation of the item and they would have to remember that item. For S conditions in Experiment I, Ss had some expectancy that the items will be presented again, and in Experiment II, there was some chance that they would have to forget the item later. Therefore, the difference in performance between S and C conditions could not necessarily be attributed to the initial interference activity. On the other hand, Ss in the C conditions were likely to be a bit fatigued after their initial counting period and in addition, were not given a ready signal before the stimulus to help them attend to the stimulus. Thus there were factors working against the observed difference as well as for it.

However, the comparison between the short (12 seconds) and the long (27 seconds) prior interference intervals is not subject to the above complication and is more meaningful. The average retention in the long prior interval conditions is somewhat better than in the short prior interval conditions, although almost all the difference is accounted for by the 3 second rehearsal conditions. This latter fact is the reason that the rehearsal x prior interference interaction effect in the analysis of variance is larger than the prior interference main effect. Although the overall difference is reasonably large (6.1% for items, 3.7% for words), the main effect was not significant for the trigram data. It was significant ($p < .05$) for the word data. The rehearsal x prior interference interaction is significant ($p < .05$) for both words and items. There is no interaction ($F < 1$) between the prior interference interval and the retention interval.

Task IV (Forget)

The most striking aspect of the data from Task F is that effects associated with the to-be-forgotten item (first rehearsal time and interference time after the first item) are all quite small (see Table 4). Overall retention in Task F was about the same as in the comparable conditions of Task S for Experiment I and a bit higher than in Task S for Experiment II. For reasons similar to the ones outlined in the previous section, F conditions are not quite comparable to S conditions, since in F conditions the S knew he would be tested on the second item from the moment he saw it. Thus the interfering effects of the first item could be nullified by extra incentive to work. However, Tasks F and C are comparable, since the only difference between them is in the presence of the first, to-be-forgotten, item of Task F.

There are several effects in the Task F data that merit comment. As mentioned before, the effects of second rehearsal and interference intervals are approximately the same as the effects of rehearsal and interference in Task S. The only effect that is significant for the first rehearsal and interference intervals ($p < .05$) is that the per cent of words recalled when the first interference period was long is greater (2.5%) than when the first interference period was short. The improvement in performance with increasing first interference interval shows up in the trigram retention data as well as the word retention data but it is small (1.5%) and not significant ($.10 > p > .05$).

The effect of increasing first rehearsal, though small, is to improve performance and the effect shows up in both trigram and word retention data, but it is stronger for the trigram retention data. There is a 3.5% dif-

TABLE 4

TASK IV (F) RETENTION DATA

Second Rehearsal and Interference Intervals (sec.)		First Rehearsal and Interference Intervals (sec.)					
		$R_1=0$		$R_1=3$		$R_1=6$	
		$I_1=7$	$I_1=22$	$I_1=7$	$I_1=22$	$I_1=7$	$I_1=22$
		Percent of Trigrams Correctly Recalled					
$R_2=0$	$I_2=10$	22.2	26.4	27.8	33.3	30.6	23.6
	$I_2=22$	12.5	12.5	12.5	16.7	11.1	23.6
$R_2=3$	I_2-10	41.7	38.9	54.2	43.1	48.6	59.7
	$I_2=22$	31.9	27.8	29.2	36.1	36.1	38.8
$R_2=6$	$I_2=10$	54.2	62.5	52.8	56.9	61.1	54.2
	$I_2=22$	47.2	44.4	41.7	45.8	43.1	41.7
		Percent of Individual Words Correctly Recalled					
$R_2=0$	$I_2=10$	53.1	57.9	56.0	58.8	56.9	55.9
	$I_2=22$	33.3	35.2	37.0	37.5	34.3	42.1
$R_2=3$	$I_2=10$	67.1	70.8	75.5	74.1	67.1	73.6
	$I_2=22$	50.0	53.7	54.6	57.5	58.7	56.9
$R_2=6$	$I_2=10$	76.9	81.9	74.1	83.8	84.3	80.1
	$I_2=22$	68.1	63.0	62.0	69.9	63.0	64.8

ference between the 0 and 6 second rehearsal conditions for trigrams and a 2.2% difference for words, $.10 > p > .05$ in both cases. Any theory that postulates that rehearsal increases the potency of an item and therefore increases its interfering effect would predict the opposite effect. Thus, not only is the expected detrimental effect of rehearsal of the forget item not significant, it is not observed. There are no interactions whose F ratio approached significance involving the first rehearsal or interference interval.

Retention in F conditions is definitely worse than in the comparable C conditions (11.6% for trigrams, 9.0% for words). Of additional interest is the fact that the effect of increasing the first interference interval is smaller for Task F (words 2.5%, trigrams 1.5%) than for Task C (words 3.7%, trigrams 6.1%).

Intrusion Data

The intrusion data are at the same time rich and elusive. They are rich because there are many possible aspects to look at and elusive because the effects are difficult if not impossible to summarize and interpret. In what follows each single word intruded is counted as an intrusion, regardless of whether it was put in the same serial position as it originally appeared. Only overt errors in which the word comes from a trigram previously presented in the experiment are counted as intrusions.

One aspect of intrusion data that has been investigated fairly widely is the per cent of all intrusions that come from the trial one back, the trial two back, etc. In this aspect, the intrusion data from all tasks look similar in general outline (see Table 5). The immediately preceding trigram provides the most intrusions, the trigram before that provides the

TABLE 5

FREQUENCY OF INTRUSIONS AS A FUNCTION

OF POSITION OF INTRUDING ITEM

Experimental Task	Number of Trials Intruding Item is before Critical Item									
	1	2	3	4	>4	0*[a]	1*	2*	3*	>3*
Percent of Total Word Responses that Are Overt Intrusions										
Task I(S), Exp. 1	3.97	0.48	0.10	0.12	0.12	-[b]	-	-	-	-
Task II (DP)	1.26	0.36	0.12	0.05	0.23	-	-	-	-	-
Task III(DT)Recall 1	6.71	1.27	0.12	0.00	0.46	-	-	-	-	-
Task III(DT)Recall 2	6.87	1.08	0.15	0.04	0.54	-	-	-	-	-
Task I(S), Expt. 2	2.06	0.27	0.14	0.08	0.39	-	0.25	0.02	0.02	0.10
Task IV(F)	1.00	0.37	0.26	0.04	0.53	4.33	0.09	0.05	0.00	0.12
Task V(C)	0.81	0.42	0.46	0.12	0.46	-	0.19	0.00	0.00	0.15
Percent of Total Word Intrusions from Tasks That Are Intruded by Position of Previous Item										
Task I(S), Expt. 1	81.4	9.9	2.0	2.4	4.3	-[b]	-	-	-	-
Task II(DP)	62.4	17.8	5.7	2.5	11.5	-	-	-	-	-
Task III(DT)Recall 1	78.4	14.9	1.4	0.0	5.4	-	-	-	-	-
Task III(DT)Recall 2	79.1	12.4	1.8	0.4	6.2	-	-	-	-	-
Task I(S),Expt. 2	62.2	8.1	4.1	2.3	23.3	-	7.6	0.6	0.6	2.9
Task IV(F)	13.0	4.8	3.3	0.5	6.8	56.2	1.2	0.7	0.0	1.5
Task V(C)	30.9	16.2	17.6	4.4	17.6	-	7.4	0.0	0.0	5.9

[a]Starred items are "forget" items, that is, items that were never tested for recall. Thus, a 0*item is an intrusion from the forget item on that trial, a 1*item is from a forget item on the previous trial, etc. It should be remembered that since only half the trials in Experiment 2 are Task IV (Forget), there are only half as many forget items as regular items.

[b]Since Experiment 1 has no forget items, the starred columns are blank for these tasks.

next most and the third trigram before that provides most of the rest. The bulk of the intrusions come from the three preceding trials (93.3% for Task S, Experiment I; 85.9% for Task DP, 94.7% for Task DT test 1, 93.3% for Task DT, test 2, 77.9% for Task S, Experiment II; 70.4% for Task F, 54.5% for Task C). The comparison across experiments is a bit difficult because there were 50% more total trigrams presented in the second experiment than in the first. (See Table 5 for the details of the scoring.)

The major difference between the tasks is the per cent of all intrusions that come from the immediately preceding trigram. For Experiment I, the figures for Tasks S, DP and DT (tests 1 and 2) respectively, are 81.4%, 62.4%, 78.5% and 79.1%; for Experiment II, the figures for Tasks S, F and C are 62.2%, 56.2%, and 30.9%. The comparison between Task F and the other tasks is a bit difficult since in Task F the immediately preceding trigram is a "forget" trigram, while in all other tasks, the immediately preceding trigram is one which they might have recalled previously.

While comparison across tasks is a bit hazardous, the following effects appear to be significant: the number of intrusions from the immediately preceding item is greater in Task S, Experiment I than for Task S, Experiment II; intrusion rates for the trigram immediately preceding Task C items are lower than in all other conditions. Comparison among the various tasks on the instructions from trigrams more than one removed is also a bit hazardous since the intrusion rates are small for all but the trigram one back (with the possible exception of Task F).

The second question that can be asked of the intrusion data is how the intrusion data relate to the error data. The hypothesis that a constant per cent of all errors are intrusions is a reasonable generalization of the data

TABLE 6

Per Cent of Errors That Are

Overt Intrusions - Task I (S)

Rehearsal interval (sec.)	Interference interval (sec.) 1	4	7	10	16	22
			Experiment 1			
0	60.0	23.5	23.5	17.8	24.4	21.2
3	25.0	33.3	20.7	24.5	23.8	18.9
6	66.7	60.0	12.5	9.1	15.9	20.5
9	-[a]	0.0	10.5	14.7	23.4	30.6
			Experiment 2			
0	11.1	16.1	16.3	11.1	8.8	14.1
3	0.0	13.5	15.5	17.3	13.8	9.2
6	11.1	7.1	15.9	10.3	10.0	10.1
9	50.0	22.2	11.5	5.9	9.3	16.9

[a]There were no errors in this condition.

of these experiments especially within a given task (again with the possible exception of Task F). The intrusion data from Task S shown in Table 6 are typical.

The overall intrusion rates (per cent of all errors that are intrusions) for the tasks are: Task S, Experiment I - 17.4%; Task S, Experiment II - 12.7%; Task DP - 24.5%; Task DT, test 1 - 21.9%; Task DT, Test 2 - 21.7%; Task F - 19.7%; Task C - 8.7%. The following comparisons between tasks appear to be meaningful: the per cent of errors that are intrusions is lower for Task S, Experiment II than for Task S, Experiment I, and lower for Task C than Task F. The percentages are a bit higher for the easy Task DP and difficult Task DT than for the average S condition. These data are not unrelated to the previous intrusion data, since intrusion rates of items more than one removed are relatively low, and most of the differences in per cent of errors that are intrusions are from the trigram immediately preceding.

Given the relatively small amount of intrusion data from each condition, it seems quite impossible to ask more complex questions, such as how the relative contributions to the intrusion data from the items n-removed change as one goes from condition to condition in each task.

To summarize the last few paragraphs, the percentages of errors that are intrusions are relatively constant across conditions within a task. In comparing across tasks, the per cent of errors that are intrusions does not relate monotonically to the absolute number of errors for Task S, DP, and DT in Experiment I. Thus, the explanation of these data, if attempted, would be relatively complex.

Perhaps the information most desired in a comparison of Tasks F and C is the relation of the absolute differences of number of intrusions in various conditions to differences in number of words recalled, and more specifically, how the number of intrusions from the item that should have been forgotten relates to differences in performance (see Table 7). As mentioned before, all differences are rather small, so that a coherent (and hopefully meaningful) pattern emerges only after averaging over some conditions. The comparison of F conditions averaged over first rehearsal and

TABLE 7

Intrusion and Retention Data from

Task IV (F) and Task V (C) Compared

Condition	Second Rehearsal and Interference Intervals (sec.)					
	$R_2=0$		$R_2=3$		$R_2=6$	
	$I_2=10$	$I_2=22$	$I_2=10$	$I_2=22$	$I_2=10$	$I_2=22$
	Percent of Individual Words Recalled Correctly					
Task IV (F)	56.4	36.6	71.4	55.2	80.2	65.1
Task V (C)	68.3	51.7	77.5	63.0	88.0	70.4
Difference	11.9	15.1	6.1	7.8	7.8	5.3
	Percent of Individual Words That Are Intrusions					
Task IV (F)	6.5	10.8	5.3	7.9	3.3	7.0
Task V (C)	4.4	3.7	2.1	3.9	0.7	1.0
Difference	2.1	7.1	3.2	4.0	2.6	6.0
	Percent of Individual Words Coming From "Forget" Item Of That Trial					
Task IV(F)	4.3	7.0	3.1	5.5	1.8	4.4

interference interval with the comparable C conditions averaged over first interference interval is suggestive. In most of these averages the difference in errors between the F and C conditions is about double the difference in number of total intrusions. (The two exceptions are the 0 second rehearsal 10 second interference conditions and the 6 second rehearsal 22 second interference conditions. In the former, the difference in intrusions is only 18% of the difference in errors while in the latter, it is about 113%.) The overall difference in intrusions is 46.3%. Remember that the absolute number of errors that are intrusions for Tasks F and C are 19.7% and 8.7% respectively.

The number of intrusions from the to-be-forgotten item of the trial also accounts for about 40-50% of all the errors in each comparison of averages. In fact, these data more closely parallel the total error differences than the overall intrusion differences do. The overall per cent of the difference in errors between Tasks F and C that is accounted for by intrusions from the "forget" item is 48.3%.

Although the differences compared above are small, the effect is reasonably lawful and should not be dismissed lightly. It appears here that overt intrusion data are quite directly linked with differences in performance. The fact that such small differences in performance are so dramatically related to intrusion data, while the large differences in performance across tasks and between conditions in Task S are not clearly related to intrusion differences suggest that interference from previous items can not be the only cause of forgetting. Using this editorial comment as a transition, the discussion section will begin with exactly that issue and issues related to it.

CHAPTER IV

THEORETICAL ANALYSIS

The data reviewed in Chapter I and the results of Experiments I and II can be combined to suggest a theoretical framework of memory. In this chapter, some general constraints on an adequate theory of memory that seem necessary in view of the above data are stated, a theoretical framework consistent with those constraints is postulated, and some of the more fine-grained features of the data of the present experiments are analyzed with respect to that theoretical framework in search of some more specific constraints on an adequate theory.

Some General Constraints on an Adequate Theory

A Multiplicity of Stores

Much of the data discussed in Chapter I argue for at least two kinds of memory stores -- stores that can be distinguished from each other by functional properties as well as by differential decay rates. The work of Sperling (1960) and Averbach & Coriell (1961) has clearly demonstrated the existence of a visual store with a decay time of less than 1 second. The work of Treisman (1964) suggests that there is a similar auditory store, and thus it is plausible that there could be sensory stores for each modality. The visual store has too short a decay time to be likely to play an important role in the body of data reviewed in Chapter I and the other sensory modalities with the exception of the auditory modality are likely to be irrelevant. The acoustic store, however, merits a more detailed discussion.

In Treisman's work on dichotic listening, the same message (in prose) was sent into both ears with the delay between messages systematically varied. The S was required to shadow one of the messages. When the shadowed message preceded the other message, the Ss perceived that the messages were in fact identical for relatively long delay times, but when the shadowed message followed the other message the identity could be perceived for delay times only up to about 2 seconds. These data suggest that a sensory image of the input is maintained for a period of 2 seconds in the nonattended auditory channel before enough decay takes place so that the trace is unrecognizable. It seems reasonable that the material in this store is coded acoustically since there is no opportunity for conscious recoding.

There are other data which suggest that loss of information from an acoustic store plays an important part in short-term memory experiments. Conrad and his coworkers (Conrad, 1964; Conrad, 1965; Baddeley, 1966a; Baddeley, 1966b) demonstrated that errors in short-term memory were related to acoustic confusions. They also demonstrated that the order of acoustically similar material was remembered less well than the order of unrelated materials while the order of semantically related materials was remembered as well as the order of unrelated materials. They further demonstrated that for higher levels of learning and for longer retention intervals, semantic similarity produced a significant decrement in performance while acoustic similarity produced none. Thus, these data show that acoustic information is lost in short periods of time and that there is little acoustic information lost in long-term memory. These data do not add any information as to the time course of the loss of acoustic information except to say that it is relatively short.

The initial rapid decay in the retention curves in all three paradigms (B-P, PA, and FR) followed by a non-zero asymptote in retention (or at least a much slower decay rate) strongly argues for two memory stores with markedly different decay rates. Further evidence for two stores comes from the observation that only the recency portion of the retention curve is lowered when 30 seconds or so of interpolated activity comes between the end of a FR list and the recall period (Postman and Phillips, 1965; Glanzer and Cunitz, 1966). If all items were stored in a single store, then one would expect appreciable (though not necessarily uniform) reduction in performance throughout the FR serial position curve.

The evidence is reasonably conclusive that there is a short-term acoustic store and that there is a short-term store of some kind with a decay time of about 10-30 seconds. A central question is whether this acoustic store, which decays in 2 seconds in Treisman's (1964) experiments, plays any important role in the forgetting curves of the usual short-term memory experiments. Since the acoustic system is probably unique in that the S can help to maintain the trace by producing more sensory input (rehearsal), the decay time of the store could easily be much more than 2 seconds for tasks other than Treisman's.

It was pointed out in Chapter I that decay times for the FR paradigms ranged from about 10-30 seconds. Decay times in PA learning were about 30 seconds. These data would certainly make sense if the forgetting curves in the paradigms are largely due to loss from such an acoustic store, since the S is continually rehearsing some items in the usual paired-associate learning experiment, while in free-recall learning, the stimuli are usually presented at such a rate that the encoding of new stimuli demands a reasonably large part of the S's attention. In the B-P paradigm, the first 10 seconds is the rapidly decaying portion of the forgetting curve. If this were taken to be an estimate of the decay time in the acoustic store, it would be compatible with the above data under the assumption that the S is doing some covert rehearsal during the interfering activity (from verbal report data), but not nearly as much as in the PA paradigm.

The argument identifying the acoustic store with the store which decays in about 10-30 seconds is on much shakier grounds than the argument for the existence of either store. Hopefully, the above argument makes the identification plausible and reasonable, and for the sake of parsimony the acoustic store and the short-term (decay time 10-30 seconds) store will be thought of as synonomous.

Response Competition

Although a pure decay theory can predict that incorrect recalls will tend to be similar to the correct item, it can not easily predict proactive interference (PI) effects. If one postulates that the trace of an item decays in pieces, then it is easy to explain the former effect with a decay theory. However, some additional mechanism is needed to explain why more errors are made when similar material precedes to the to-be-remembered item than when dissimilar material precedes the to-be-remembered item. Since retention on trials after a long rest period or on Trial 1 of a B-P experiment is quite close to 100% even after long retention intervals, in contrast to retention of around 10-20% after several massed trials, the previous stimuli might be "interfering" in some way with the to-be-remembered item.

Two likely sources of proactive interference are: a) response competition -- the recall of the to-be-remembered item depends not only on its absolute strength but on its strength relative to other competing items; and b) storage interference -- the more similar two items are, the more each will cause the other to decay. While there is nothing in the data that demands either interpretation (in fact, both processes could occur), the response competition mechanism seems more parsimonious.

Although a response interference explanation is not required to explain that many of the errors are intrusions from previous stimuli, some sort of response rule is required which will predict that recent items are likely to be recalled instead of the to-be-remembered item. Since the most natural response rules of that class are response interference mechanisms, one would have to try hard to select a response rule that wouldn't predict response interference. Since one is going to use a response rule that predicts PI effects anyway, a storage interference mechanism seems to be superfluous until the data require it.

Decay of Traces

The assumption that all traces decay with time is both prima facie more reasonable than the assumption that some traces spontaneously recover over time and is also sufficient to explain the observed phenomenon of "spontaneous recovery." The assumption that a given trace gets stronger or more potent over time after extinction would appear to defy the law of entropy. This is not to say that whatever is left of the trace could not be becoming more and more resistant to further decay. It makes more sense to assume that information, strength, or energy is spontaneously lost or dissipated rather than spontaneously recovering.

The following simple response rule will be used as an example of a large class of response rules that could be used: the probability of making a response is the strength of the trace of that item divided by the sum of the strengths of all the traces connected to the stimulus in question (Luce, 1959). This rule will be referred to as the ratio response rule. For wide classes of decay functions, the ratio response rule will predict that the probability of making an intrusion response will increase as the retention interval increases ("spontaneous recovery" of the intrusion response). It can be shown that given appropriate decay functions the same response rule can be made consistent with many of the "spontaneous recovery" phenomena of long-term memory.

Differential Access to Material in the Long-Term Store

"Release from PI" data dramatically demonstrate that there are occasions when material from LTS is retrieved with no difficulty while there are other occasions when material from LTS is very difficult to retrieve. One way to view this phenomenon is: if the S has a unique code for a stimulus item then the item can be retrieved from LTS perfectly, whereas if he has a general code, such as a general semantic label, then he will have difficulty retrieving the item from LTS if there are other recent items that have been similarly coded. In terms of the response rule introduced in the previous section, we can think of a unique code as a code which designates only one trace with appreciable strength, while a general code designates many traces with appreciable strength. Since semantic labels can function as unique codes and produce "release from PI" if other exemplars of the semantic class have not appeared recently in the experiment (as demonstrated in many experiments reported in Chapter 1), the traces in question are very likely to be in a store with semantic organization and hence in LTS. Since performance is close to 100% for long retention intervals in "release from PI" data, there is almost always some representation of the item in LTS in all B-P experimental trials.

STS and LTS Representations not Unique

In many quantitative treatments of memory, representations of an item in STS and LTS are each conceived of as single-valued -- LTS is a store in which an item can be stored indefinitely and from which items can be retrieved perfectly, while STS is a store in which an item has a fixed probability at any moment of entering the LTS and a fixed probability of being lost from the STS and from which items also can be retrieved perfectly.

It has been argued in the previous sections that the above conception of LTS is not tenable if it is to explain PI data. It is clear from the previous section that items in LTS are not always perfectly retrievable even for retention intervals of several seconds. Furthermore, since the interference effects show a time course -- the time of rest intervals and the number of intervening items between two items from the same semantic class are both significant variables in retention -- some notion of a decaying LTS trace is mandatory.

One result that argues against a simple all-or-none trace loss in STS is the change in shape of the decay curve in the present experiments with increasing rehearsal. If all that increasing rehearsal did was to put more items into LTS, then the retention curves for the various rehearsal periods would have the same shape, but different asymptotes. The change in shape can alternately be interpreted as an LTS phenomenon, but a very complex and unlikely LTS decay function would be needed to explain the data.

Another STM result that argues for a complex view of STS is the observed interaction of the spacing effect with retention interval in PA learning. Although a definitive mechanism for the interaction has not been proposed, the superiority of massed practice for short retention interval does suggest that massed practice produces a stronger STS representation after the second stimulus presentation than spaced practice does.

Two mechanisms that have been proposed to supplant the notion of a single all-or-none STS trace are (a) a continuously decaying trace or (b) multiple copies of all-or-none traces for a single item. For the former, rehearsal strengthens the trace and for the latter, it adds more copies of the trace. The two ideas are hard to separate, and since it is somewhat easier to talk about an item either being in the STS or not being in the STS, the multiple copy conception will be adopted. It can be shown that a small number of copies (five cr so) is sufficient to produce the observed changes in shape of the Task S retention curves of the present experiments.

A Conceptualization of Memory

The General Framework

In this section, a framework consistent with the general constraints of the previous section is proposed. The framework reflects the author's tastes and biases and is certainly not the only such framework consistent with the available data.

Three entities are conceived as playing a significant role in the usual STM experiments: STS, LTS, and working memory (WM). As mentioned before, LTS is conceived as being a semantically organized store in which items gradually decay and in which response competition often plays a significant role in retrieval. STS is conceived of as being an acoustically organized store in which items may be multiply represented, but in which each representation decays in an all-or-none fashion. Although response competition could play some part in STS, we will assume that it is negligible for the purposes of most experiments and that any information that is in STS can be retrieved.

Working memory, which is intended to correspond closely to an attentional central processing mechanism, is conceived as a limited capacity storage region of approximately constant size which overlaps STS and LTS and therefore has structures common to both stores. The conception of WM here is analogous to an amoeba which is divided between the two regions of LTS and STS can send out "psuedopods" into either region. Its "psuedopods" probably can extend into almost all of STS, but there are probably regions of LTS to which it has only indirect access.

The retrieval process is conceived as an address search. The contents of WM at any given time are conceived of as being a collection of memory or thought units (e.g., words) which have addresses (<u>codes</u>) of other traces in LTS attached to them. Retrieval is accomplished by simply going to an address and selecting a stimulus trace at the address. If there is more than one trace at an address, one of them is selected according to a retrieval rule which depends on the strengths of all traces at the address, such as the ratio response rule suggested earlier. If there is more than one code attached to a memory unit in the WM, then another retrieval problem exists. A set of addresses can be conceived of as presenting much the same problem as does a number of traces at a specific address. It is probably not worth trying to distinguish between the two sources of competition -- competition between addresses and competition between traces at a given address -- so they will be lumped together.

Memory units can remain in WM as long as they are attended to. In a typical STM experiment the S must either attend to a few items (keep his WM in the same parts of STS and LTS) or try to switch his attention back and forth from memory unit to memory unit (move the "psuedopods" of his WM around). Since WM has a limited capacity, some memory units can fall out of WM if the S is forced to attend to other things. In the B-P paradigm, it will be assumed that the interfering task is such that (as a rough approximation) the S can keep a one word unit with one code attached in his WM during the interference interval.

When a S attempts to retrieve a memory unit, he fails to retrieve it only if it is neither in WM, STS, nor retrievable from LTS. If these processes were independent, then the probability of not retrieving a memory unit would be:

$$pr(\text{no retrieval}) = pr(\text{not in WM})pr(\text{not in STS})pr(\text{not retrieved from LTS}) \quad (1)$$

Although the independence hypothesis is probably wrong, it will serve as a convenient tool to help analyze the data of the present experiments. If the basic assumption that recall fails only when all stores fail, is correct, then the product rule (Equation 1) should be a reasonable approximation to the true combination rules unless there are some large correlations between the several encoding, storage, and retrieval processes. Later discussion will show that the correlations are not likely to be large.

Some Specifications

The role of rehearsal. -- Within the above schema, rehearsal and additional presentations of an item have at least four likely functions: (a) adding more copies to the STS representation of the item; (b) strengthening the LTS representation of the item; (c) serving as an interference period for previously presented stimuli; (d) allowing the S to recode an item so that it has a "better" representation in LTS. Since the first three functions are reasonably straightforward, only the fourth function needs amplification.

Recoding to achieve a "better" representation is conceived of as producing a more unique code (UC) for the item in LTS. Examples of unique codes that Ss seem to use in these experiments (from verbal reports) are visual images that contain the objects represented by the three words in the trigram, a phrase or short sentence containing the three words, or occasionally, a rhythmic pattern that forms a cohesive whole. Although the previous discussion strongly suggests that uniqueness of an LTS code is not all-or-none, to simplify discussion, codes will be referred to as either unique codes, codes with which the item in LTS can be perfectly retrieved, or general codes, codes to which some trace of all previous items are connected. An example of a general code (GC) is the address, "short words."

Our discussion so far has prudently avoided the question of what the units of memory are. For the purposes of the present discussion, we will assume that items that are uniquely coded in LTS are largely three-word units, although they could easily be two-word units or single words (e.g., the visual image probably contains all three words if the S has a long enough time to think about it, but he might only get one or two of the words in his picture if he doesn't have sufficient rehearsal time). If the words are not uniquely coded, then they are thought of as being represented by three traces generally coded independently in LTS. Items in STS are thought of as being stored as an ordered set of acoustic patterns. An exact specification of the units of decay in STS is not important for any further discussion. The units are probably small -- at the level of the phoneme or distinctive feature -- and the decaying STS trace can be thought of as a piece of recording tape in which various sections get erased but the record that something was there remains.

As mentioned before, the increasingly sigmoid shape of the retention curves in this experiment with longer rehearsal and the interaction between the retention and spacing intervals in several PA experiments suggest that the STS representation becomes more potent with rehearsal or massed repetitions. Since shorter stimulus units can be rehearsed faster, one would also expect more sigmoid retention curves the shorter the stimulus unit (see Figure 4). Since a shorter unit is also more likely to be uniquely codable, one would also expect a higher asymptote in the retention curve the shorter the stimulus unit. Since other data of the present experiments do not bear critically on STS, the fine structure of STS will be ignored subsequently.

A key question about the effect of rehearsal on long-term storage is the relative importance of the following three functions: (a) strengthening the LTS trace of the to-be-remembered item; (b) weakening other LTS traces, and (c) finding a more unique code for the to-be-remembered item. The analysis in the next section of the Task F and Task C data from Experiment II points to some conclusions.

Another key question about the effects of rehearsal on memory is whether some experimental conditions provide better conditions for unique coding of items than others (e.g., longer interference intervals between presentations of the stimulus allow the second stimulus presented to be coded more effectively). An analysis of the Task DP and Task C data from Experiments I and II will help to decide between the above explanation of spaced practice effects and others.

The role of interference. -- The hypothesized role of interfering activity in the BP paradigm is to cause the STS traces units to drop out in an all-or-none fashion and to cause all LTS traces to decay gradually. These decay processes are conceived of as being relatively autonomous processes except for the amount of furtive rehearsal and/or attention the S can achieve during this counting or shadowing activity. Most of the experiments which demonstrate that the rate of the interfering activity significantly affects retention use rates in which the S must be doing a fair amount of rehearsal (e.g., one three-digit number per second). Thus, it is not clear that manipulation of the difficulty of the intervening activity is doing anything other than determining the amount of covert rehearsal that the S can achieve during the interference period.

A major issue in STM theories is whether consolidation as well as decay occurs during interfering activity. Mechanisms of consolidation are often a bit obscure. One well defined mechanism of consolidation that might be assumed in the present system is a transfer of material from STS to LTS during interference as long as material is in STS (Atkinson and Shiffrin, 1968; Landauer, 1969). This and other conceptions of consolidation are evaluated in the next section as explanations of the spacing data in these experiments.

The role of recall. -- The role of recall in STM is not well understood and there is little critical data on it. Intuition would lead one to think that if an item is recalled, all traces associated with the item must be strengthened, but little solid data exist even on this point. The result from the DT Task of Experiment I that an item was almost certain to be recalled on the second recall test given that it was recalled on the first recall test, supports the above assertion but is hardly conclusive. Since the time allowed for recall was 8 seconds, the S had time to rehearse the item after recalling it. Furthermore, it is hard to separate the following two hypotheses: a) recall strengthens items so that they are very likely to be recalled a second time; b) strong items are likely to be recalled the first time and therefore, because they are strong, are likely to be recalled a second time given that they were recalled a first time.

A more serious lacuna, though, is the lack of any details of the retrieval mechanism. Even if the simple level of analysis of the theory presented here in which all sorts of complexities are reduced to response strengths is accepted, and a relatively simple response rule such as the ratio rule is also accepted, there are still problems. The major problem is that there are likely to be important sequential dependencies between recall of a unit and recall of another unit. Even if it is assumed that all the positive dependencies are accounted for by assuming that units can be organized into bigger units, intuition leads to the conclusion that recall of one item must also interfere with subsequent retrieval of other items just as interfering activity interferes with subsequent retrieval of items.

A subsidiary problem is that the period allowed for recall is restricted and the latency of retrieval of an item (especially from LTS) must be a relatively important variable determining retention performance measures. Even if latency of retrieval were a simple function of response strengths, one would have to specify the function, at least approximately, to seriously attempt a quantitative fit of retention data. The obvious non-independence of the experimentally defined unit (words) makes the problem of achieving a meaningful quantitative fit even more difficult.

Another feature that must be built into any meaningful theory of recall is a matching process. From introspection, it is clear that traces are dredged out of what seems to be LTS and compared against either the acoustic traces left in STS, semantic information in WM, or other semantic information dredged up from LTS. An example of a check against semantic information is searching for the appropriate words in a crossword puzzle where associates of the cue word are pulled out and rejected until the correct one is found. Thus, a reasonable retrieval rule will probably have to be further complicated by a sequential search process.

For the purposes of subsequent discussion, no serious attempt will be made to compute response probabilities from trace strengths. It will simply be assumed that the S either has a unique code or he doesn't and that:

pr(no retrieval from LTS) = pr(no UC)pr(can not retrieve trace with GC)   (2)

The probability of not having a unique code is a function only of encoding conditions (i.e., once he gets it, he will keep it for at least 20 seconds) and the probability of not getting a trace of LTS using the general code is a function of trace strengths and random events at the time of recall. Thus it is not unreasonable to think that the probabilities on the right hand side of Equation 2 are independent and therefore multiply to give the probability of retrieval failure from LTS. The theory can certainly be used to yield ordinal predictions of changes in the probability of retrieval resulting from some experimental manipulations.

Some General STM Trends: An Evaluation of the Theory

With all the difficulties mentioned, a quantitative fit of data will involve specifications of many unknowns: the LTS decay function, the STS decay function, the LTS retrieval rule, coding rates, and some specification what the subject is doing during interference. Clearly, the subject is doing different things in the PA and FR paradigms (e.g., trying to encode other items), than in the B-P paradigm. A qualitative assessment of the theory seems feasible, however, and can serve to put the theory in some context. Of course, the theory will "predict" much of the data, since it was constructed with much of the STM data in mind (especially the general constraints discussed before).

In free recall, the recency part of the curve is thought to be basically a STS phenomenon while the words recalled in the rest of the list are thought to come from LTS. When 30 seconds of interfering activity is interspersed between the end of the list and recall, the chance of any items being in STS becomes essentially zero so that the recency effect disappears. The stability of the earlier part of the free recall serial position curve can be explained by assuming the ratio response rule and exponential decay of LTS traces, since the LTS traces of all the items are decaying and their relative strengths should remain about the same. The lowering of performance in the middle of the list with increasing list size can easily be explained by the ratio response rule or just about any response competition model. As mentioned in Chapter I, the primacy effect could have many causes.

In PA learning, the duration of the rapid decay portion of the forgetting curve is somewhat longer than in the other two paradigms. This can be explained by the fact that the S has more opportunity to keep rehearsing an item while succeeding items are being presented. Thus the rapid part of the forgetting curve consists of both STS decay and items being "bumped out" of WM or, to use Atkinson and Shiffrin's (1968) term, the rehearsal buffer. In the B-P paradigm, this "bump out" process presumably occurs very rapidly at the beginning of the interference period and FR rates are rapid enough that the S's working memory is taxed by perceiving each item as it enters. For slower FR presentation rates, the duration of the recency portion presumably should approach that of the rapid decay portion of the PA forgetting curves.

In many continuous PA experiments, the data seem to be consistent with the notion of a LTS store in which all items are perfectly retrievable. This may be because the standard strategy of the S is to keep a few items in WM and process them until he gets a unique code. If the S is using the recall time for rehearsal as well as retrieval, then he might not bother trying to retrieve an item with a general code since the time allowed for recall is brief. In other PA experiments, the data demand some sort of increasing difficulty in retrieving items from LTS with increasing retention interval. In these studies, the Ss get at most two or three presentations of a particular S-R pair and huge numbers of S-R pairs so that the above strategy becomes less likely and many items are probably not worked on intensively

(Atkinson and Shiffrin, 1968). Since the S can not rely on unique codes, he will have to fish around with his general codes if he wants to do well in the experiment.

The PI data in the B-P paradigm are relatively easy to explain on a qualitative level assuming something like the ratio response rule and exponentially decaying LTS traces. Specifically, one can explain the following regularities in PI data mentioned in Chapter 1: (a) retention of an item after a long rest interval or retention of an item constructed from a new category is quite high -- this prediction follows because the other traces attached to the general code are weak or non-existent; (b) retention on a particular trial in experiments in which several categories are used depends largely on the number of trials separating the to-be-remembered item from the last item of the same category -- this follows because the further back are the interfering items attached to the same code, the weaker they are; (c) retention asymptotes at a smaller number of trials the longer the intertrial interval -- this follows because the longer the intertrial interval is, the weaker each competing trace is compared to the one succeeding it so that the longer the rest interval, the fewer trials it takes a competing trace to be negligible in strength; (d) intrusions come largely from the same semantic category -- this follows because intrusions come from retrieving the wrong items using the appropriate general code; (e) the more items that separate the last item of the same category from the to-be-remembered item, the fewer the intrusions from that item -- this also follows because the traces of the intruding item decay more the longer they have to decay.

In the next section, the PI effects and the effects of the spacing of presentations in the present experiments are analyzed in an attempt to test the general theory proposed here and to see if more constraints are implied.

Analysis of the Present Experiments

An Analysis of Interference

To clarify discussion in this section, the STS will be largely ignored. It will be assumed that subjects are recalling a bit from STS at the 10 second retention interval and essentially nothing from it at the 22 second retention interval (the two retention intervals used in the Control and Forget Tasks). It will also be assumed to begin with that all events going on before the presentation of the stimulus that is to be remembered change the S's ability to form a unique code for it very little and, therefore, that all effects owing to manipulation of these events result from differential ease of retrieval of the to-be-remembered item using a general code. It will further be assumed that the content of the S's working memory at the end of a reasonably long retention interval is one memory unit (word) plus one code.

One problem in this discussion is to specify the proper unit of analysis. Since retrieval using general code is the focus of interest and the units of retrieval from general codes are assumed to be individual words, the units in the subsequent discussion will be words. It doesn't matter that the words are not likely to be the functional memory units for uniquely coded items. All that does matter is that a certain proportion of the total words will either be in working memory or uniquely coded, and for the present, this proportion is assumed to be a function only of the length of the rehearsal period of the to-be-remembered item.

Rehearsal of the "Forget" item. -- Before attempting a semi-quantitative estimation of decay rates from some of the interference data, consideration of the effect of rehearsing the to-be-forgotten item in Task F helps limit the number of possible mechanisms to consider. As mentioned in the previous section, rehearsal of an item could (a) increase the LTS strength of an item or at least to prevent its decay, (b) allow for the creation of a new unique code for an item, (c) allow generally coded traces of the item to decay when its unique code is rehearsed and (d) cause decay of prior items. Since the data from Task F indicate that increasing rehearsal decreases interference, effects c and d probably at least offset effect a. Since effects c and d are likely to be small, effect a must be close to zero. Although the effect of increasing rehearsal is not significant, one can with some confidence reject the assertion that rehearsal of the forget item increases the total amount of interference.

It is surprising that additional rehearsal doesn't add to the total interference in the light of the observation in Chapter I that prior items which are simply read or shadowed interfere much less than those that the subject has, at one time, tried to remember. Perhaps well-learned stimuli (four-letter words) reach a maximum strength in LTS after about two seconds of exposure so that adding rehearsal time to the presentation time does not increase the strength of the LTS generally coded traces appreciably in these experiments. However, if there were an appreciably shorter presentation time, the added rehearsal time might add to the strength of a generally coded LTS trace.

Regardless of what might happen with shorter presentation intervals, the data of the two experiments indicate that additional rehearsal of the to-be-remembered item does improve retention vastly for the 2 second presentation interval employed, so that rehearsal must be having a significant effect on the LTS representation. Thus, for the presentation time and rehearsal times of these experiments, the overwhelming effect of rehearsal must be to allow for more unique codes (UCs) to be formed for LTS traces rather than to strengthen LTS traces.

Tasks F and C compared. -- The two other major regularities of the Task F and Task C data -- retention in a Task F condition is worse than in the comparable Task C condition, and increasing the first interference interval in both Tasks F and C improves performance -- imply that response interference is producing significant effects in the present experiments. Retention is poorer in the F conditions than in the C conditions because everything is the same except that another interfering trace is present. The longer first interference interval produces better retention in both tasks because all interfering stimuli have decayed more.

One can obtain an order of magnitude estimate of LTS decay rates and the proportion of words recalled from general codes (GCs) using the product retrieval rule, the ratio response rule, and some simplified decay assumptions. If we assume that for a given condition that the probability of not retrieving an item is:

$$\text{pr(no retrieval)} = \text{pr}\begin{pmatrix}\text{not retrieving a}\\ \text{word from a GC}\end{pmatrix}\text{pr}\begin{pmatrix}\text{a word not in WM}\\ \text{or having a UC}\end{pmatrix} \qquad (3)$$

and if we furthermore assume that the difference in retention between an F condition and a comparable C condition is solely due to differential difficulty of retrieval using a general code, then:

$$\frac{\text{pr(word not recalled in Task F)}}{\text{pr(word not recalled in Task C)}} = \frac{\text{pr(word not retrieved from a GC in Task F)}}{\text{pr(word not retrieved from a GC in Task C)}} \quad (4)$$

Since the quantities on the left hand side of Equation 4 are observables, then the quantities on the right hand side can be estimated if some simplifying assumptions about decay and retrieval are made. Since it has been claimed that there is almost always one code in WM, then it is likely that either the S retrieves the whole 3-word item using his WM or he has the trace of one word in his WM and he starts using a GC to get the other two (we are assuming his STS is not contributing anything to retrieval). Thus, if we assume that the subject first hunts around for a unique code, and, when he can't find it, starts to try to retrieve items using his general code, then a retrieval rule which says that he overtly recalls the first two items he retrieves (using the ratio response rule for each retrieval) might be a reasonable gross approximation to his retrieval process from LTS. The desired quantities can be estimated with one more restriction, namely, that the ratio of the strength of the traces of the "forget" words to the strength of all the previous traces is a little over two to one (the estimate is derived from the ratio of the intrusion rates).

The estimated ratio of the strength of the traces of the to-be-remembered words to the strength of the traces of the "forget" words is about two to one. Since the temporal interval between the "forget" item and the to-be-remembered item is approximately equal to the temporal interval between the "forget" item and the preceeding item, the two to one ratio estimated from the calculation is in reasonable agreement with the two to one ratio estimated from the intrusion data. The suggestion is that LTS traces decay approximately exponentially and have a "half-life" of about 15-20 seconds.

One can also calculate the approximate values of the probabilities on the right-hand-side of Equation 4. The estimated average for Forget conditions is 0.87 and the estimated average for Control conditions is 0.74. All the above estimates were calculated using data averaged over all conditions with a rehearsal interval of 3 seconds for the to-be-remembered item. The data for other rehearsal intervals give similar estimates.

Since the recall probabilities from Task F are of the same order of magnitude as those in the comparable S and DT conditions, we can get estimates of unique coding rates from Task F that will be reasonable guesses for the other two tasks. For 0 seconds rehearsal, there is about a 30% probability that a word is encoded either in WM or UC, for 3 seconds rehearsal the probability is about 50%, and for 6 seconds rehearsal the probability is about 60%. Thus, about 80-90% of the words that are recalled at long retention intervals in Tasks S, DT, and F are coming from WM or a UC.

Probably the most glaring omission in the above model of retrieval is the absence of any matching processes in WM. The restricted time allowed for recall and the very appreciable time needed for retrieving each item are likewise ignored. It is clear from a typical S's behavior that he retrieves some items and rejects them on the basis of some other information in WM. Furthermore, it is reasonably likely that the S might not overtly recall some items he has retrieved because he is busy trying to retrieve other items and the recall period ends before he can say them.

While these factors are likely to produce much greater variability in recall than the retrieval model used in this section predicts, the model still could be a reasonable approximation for the purposes of this discussion. The estimates obtained give an idea of what the LTS decay rates are if the theory is correct and also give estimates of the proportion of the words recalled that are coming from sources not subject to response interference.

One trend in the data that does give some difficulty to the present theory is a relatively greater improvement in retention due to increasing the initial interference interval in Task C than in Task F. Since the "forget" item is considerably stronger than the previous items, reducing its interfering effect in addition to reducing the interfering effects of previous items should yield the prediction that the improvement in Task F should be markedly greater than that in Task C.

One possible escape from this difficulty is to postulate that covert rehearsal of an item during interference helps to maintain its LTS strength. According to this explanation the "forget" item does not decrease in strength as rapidly during the first interfering activity as the traces of the previous items, and therefore the total interference does not decrease as much in the Task F condition as in the Task C conditions.

A second possible escape is to invoke the concept of fatigue. If it is significantly more fatiguing to count backwards while trying to remember something than when not trying to remember something, then the subject will have a greater increment of fatigue from 15 extra seconds of counting in F conditions than in C conditions. If differential fatigue results in differential coding ability at the time of the second presentation, then the effect of fatigue could outbalance the otherwise expected result. If differential fatigue is invoked as a concept, however, then to whatever extent it is fatiguing to do a lot of counting, we are underestimating the other effects of spacing of presentations in Task F, C, and DP. Perhaps, since the differential effect of spacing is on the average small for Task F vs. Task C (2.5% words recalled vs. 3.7% words recalled, respectively) further discussion of the effect is not warranted.

While the previous discussion has dealt with the word data, all of the interference effects are also observed in the trigram retention data. To predict the magnitude of the effects in the trigram data from the magnitude of the effects in the word data requires some reasonably detailed assumptions about the conditional probabilities of recalling one word of a trigram given that you have already recalled another. Since recall of one word given a UC is highly correlated with recall of the other words of the item, while recall of words from a GC is probably negatively correlated (recall of one word interferes with recall of others), the predictions would be quite complicated. However, one would expect intuitively that the interference effects should be somewhat smaller for trigram retention since complete units are more likely to come from unique codes than are individual words. For the same reason, one would expect even a higher proportion of trigrams correctly recalled to come from unique codes than the 80-90% estimated for individual words.

Intrusion Data. -- If one applied the retrieval model of the previous section as is, then one would expect essentially all errors to be intra-experimental intrusions. One would expect a few errors due to acoustic distortion in STS, but no errors of omission. That errors of omission are reasonably common argues strongly for some mechanism for suppressing overt responses. As mentioned in the previous section, it is likely that a comparison is made between retrieved items and other stored information and, if a match is not made, the response is suppressed. Such information could be of various types: bits of acoustic information, bits of semantic information (e.g., the last word began with a "C"), and other cues. One that some subjects appeared to use (from verbal reports) was "the feeling of having recalled the item before." If such a cue were used, then the ratio of intrusions from "forget" items to the number of errors caused by the presence of the forget item should be higher than the overall ratio of intrusions to errors since the "forget" items have not been recalled and do not have this additional rejection cue. In fact, the ratio of intrusions from the "forget" item to differences in errors (word data) between C and F conditions is about 50% while the overall ratio of intrusions to error ranges from about 15-33% (see Tables 6 and 7). Other explanations of this phenomenon, however, are possible (see below).

An overall analysis of the intrusion data is further complicated by a scoring problem. Since many four letter words used in the experiment were acoustically similar to others, it is hard to tell an acoustic distortion from an intrusion. In the scoring system used in the present experiments, all errors that could be counted as intrusions were, with the probable result that there is a significant overestimate of intrusions.

The above complications make any precise analysis of intrusion data fairly difficult. However, some of the intrusion data do seem sensible if we assume: (a) that the latency of retrieval (probably a function of relative and absolute strengths of traces) is an important variable determining the percent of errors that are intrusions as opposed to omissions; (b) that the decay function is such that the ratio of any two strengths becomes more equal as time passes (i.e., the decay function is "faster" than exponential decay). The former predicts that the total number of intrusions should be low for Task C (and also Task DP) and the latter predicts that the percent of intrusions coming from the most recent item should be lower in both those conditions than in the other three conditions (see Table 5).

Within a task, the analysis is probably more difficult. The loss of acoustic units with increasing retention interval probably produces a loss of information against which to check his retrieved words and a lowering of response suppression. On the other hand, the probable increase in the latency of retrieval with increasing retention interval would produce an increase of response suppression. Acoustic distortions probably add another source of variance. The intrusion data within a task in fact show no clear trends, and the argument above suggests that they are not likely to be simply explained (see Table 6).

An Analysis of Spaced Practice Improvement. -- To facilitate the discussion of spaced practice in the Double-Presentation (DP) Task, a brief resume of the theory will help. The probability of not recalling an item correctly is given by the product rule:

$$\text{pr(error)} = \text{pr}\begin{pmatrix}\text{item not}\\ \text{in STS}\end{pmatrix}\text{pr}\begin{pmatrix}\text{item}\\ \text{not UC}\end{pmatrix}\text{pr}\begin{pmatrix}\text{item not re-}\\ \text{trieved with GC}\end{pmatrix}\text{pr}\begin{pmatrix}\text{item not}\\ \text{in WM}\end{pmatrix} \qquad (2)$$

The STS decay has been assumed to produce the first rapid loss of memory. Increasing the rehearsal time increases the probability of the item being uniquely coded, increases the potency of the STS representation, and does not change the strength of the trace of the generally coded traces very much (let us assume not at all for simplicity). During interference, LTS and STS each decay and both also decay during rehearsal of other items. The rate of

STS decay probably depends on how much "rehearsal" the subject can get in during the interfering activity. Presumably furtive rehearsal during the interference period will also help to prevent the LTS trace from decaying somewhat. For simplicity, let us say that the amount of covert rehearsal is the only factor which influences either STS or LTS trace decay rate. Recall might have some more interesting properties than the reinstatement of STS and LTS traces to some approximation of their original strength, but we don't have the relevant data to decide the issue. Items "interfere" if they are designated by the same code as the item to be recalled. If several items are designated by the same code, the probability of recalling an item is given by something like the ratio response rule.

Within this universe of discourse, there seem to be three classes of explanation for spaced practice improvement (SPI) in the DP Task. The three categories of explanation are: (a) reduction of interference from previous items, (b) "consolidation" of the trace of the to-be-remembered item (i.e., some increasing resistance to extinction that the trace acquires with time in the memory system) and (c) differential coding (greater ease in forming unique codes on the second presentation, the further removed the first presentation of the stimulus). The ensuing discussion attempts to show that the first two explanations are likely to account for only a small part of the spaced practice effect and therefore some form of differential coding explanation is a likely cause of SPI.

Reduction of Interference from Previous Items. -- Fortunately, Task C allows an estimate to be made of the magnitude of the reduction of interference from previous items for Task DP data. The primary assumption of the following discussion is that the interference effect will be approximately multiplicative (the product rule). If we assume that the LTS decay rate for previous items (which the S is not attempting to retain) is the same whether the activity is counting backwards for an interval of time (Task C) or trying to encode and retain the word while counting backwards for the same period of time (Task DP), then a fairly straightforward estimate may be obtained of what portion of the spaced practice effect can be accounted for by the reduction in strength of interfering LTS traces. Since the two retention intervals used in both tasks are relatively long, we will assume that the contribution of STS to the final retention is not important.

In this analysis, the statistic of interest for both Task C and Task DP is the ratio of the errors (proportion of words recalled incorrectly) in conditions with 22 seconds of initial interference (spaced conditions) to the errors in conditions with 7 seconds of initial interference. The ratio of observed errors in spaced vs. massed conditions should be equal to the ratio of retrieval probabilities from a GC, since we are assuming that spacing has no effect other than weakening interfering traces and therefore no effect on the likelihood that an item has a UC or is in WM (see Equation 3).

Since the relative strengths of the generally coded traces of the to-be-remembered item and those of the competing items at the time of recall should be the same in a DP condition as in a comparable C condition the probability of retrieval from a GC is a function only of those strengths. Thus, the ratio of errors in spaced conditions to errors in massed conditions should be the same for Task DP as for Task C. The generally coded traces of an item in a DP condition should, at the time of recall, be at about the same strength as those in a comparable C condition since it was shown in the previous section that additional rehearsal time has little effect on the strength of generally coded traces. Thus in both tasks, the LTS trace will be at some maximal strength at the end of the last (and for Task C, the only) rehearsal interval. Both traces should decay some amount depending on the length of the retention interval.

The observed ratio of errors in massed practice vs. spaced practice for Task C (averaged over all variables other than the initial interference interval) is 1.05 to 1. The analogous ratio for Task DP errors for the 3 second first rehearsal conditions (and thus equated for time interval with task C) is 2.26 to 1. The ratio for Task DP averaged over all first rehearsal conditions is 2.13 to 1. Since only 5% more errors are observed in massed practice than in spaced practice for Task C while 120% more errors are observed in massed practice than in spaced practice for Task DP, it is clear that the basic assumptions of the argument would have to be off wildly in order for the interference explanation to have a possibility of being true. The key assumptions of the foregoing comparison are: (a) the product rule, (b) prior items decay at the same rate during a counting interval as during an interval in which the S is trying both to retain an item and count, and (c) the second presentation simply pushes the GC trace strength back up to some maximum. It seems intuitively unlikely that any of these assumptions can be grossly in error and therefore reduction of interference from previous items only explain a small fraction of the spaced practice improvement (SPI) in Task DP.

Consolidation of Memory Trace. -- Consolidation refers to some kind of process by which the trace becomes stronger, more accessible, or more permanent with increasing time in storage. Thus, in this view, SPI is a bit of a paradox. Given that everything else is held equal in the comparison between DP conditions except for the spacing interval, consolidation must explain SPI by postulating that the stimulus trace, on the average, is (in some sense) getting more potent during the interference period. But when performance is tested at the end of the first interference interval (i.e., as in Task S), we find that performance worsens as the interference interval gets longer. Although the above phrasing of the problem suggests a logical contradiction, one can think of several possible mechanisms for consolidation that could account for SPI.

The basic idea behind most consolidation explanations of SPI is that although the probability of recall at the end of the shorter spacing interval would be higher, the probability of a uniquely coded trace existing in LTS is higher at the end of the longer spacing interval. If the contributions to performance at the end of the first interference interval other than that of the uniquely coded items are largely irrelevant to final performance, then consolidation can possibly explain SPI.

What are possible mechanisms for consolidation within the present schema? The most likely three would be either that: (a) the trace of the acoustic pattern can somehow become a relatively permanent and easily retrievable entity; (b) somehow isolated units (words) can form themselves into a larger retrieval unit simply by sitting in the LTS or; (c) the covert rehearsal during interference can produce more uniquely coded items. Although the three mechanisms seem a bit implausible, they could occur, and the next step is to estimate the maximum possible magnitude of these effects.

The discussion is hampered, of course, by the fact that there is no direct measure of the percentage of the time that items are retrievable from LTS. There is an additional problem in that recall data does not give a complete record of what can be retrieved (e.g., smaller fragments than words). Nonetheless, it is possible to make some estimates of how much could possibly be consolidated during the first interference task.

The basic problem is to compute how many more uniquely coded items could be formed during the extra 15 seconds of interference time in the spaced conditions of Task DP. At the end of the 22 seconds retention interval, the STS is going to be quite unimportant and, hence:

$$\text{pr(error)} = \text{pr(no UC)pr(unable to retrieve from a GC)pr(not in WM)} \quad (2')$$

The equation is probably not too bad an approximation for the 10 seconds retention interval, either. If the second term on the right hand side of the equation varies in the relatively minor way argued in the previous section, and the third term is relatively constant as has been assumed throughout, we can then argue that the probability of error is related (more or less) by a multiplicative constant to pr(no UC).

$$\frac{\text{pr(error in massed practice)}}{\text{pr(error in spaced practice)}} = \frac{\text{pr(no UC formed in massed practice)}}{\text{pr(no UC formed in spaced practice)}} \quad (5)$$

Furthermore, since we are assuming that SPI is caused by added consolidation of the trace during the extra 15 seconds of spacing, the effect of the second presentation-rehearsal-interference ($P_2R_2I_2$) sequence on pr(no UC) should be the same in spaced and massed practice. In both spaced and massed conditions, the probability of no UC being formed during the $P_2R_2I_2$ sequence should be the same (depending on only the R_2 and I_2 intervals) given that no UC is formed by the end of the first presentation-rehearsal-interference ($P_1R_1I_1$) sequence. Thus, the ratio of errors in massed practice (MP) to spaced practice (SP) should equal the ratio of unique codes that were not formed during the first presentation-rehearsal-interference sequence.

$$\frac{\text{pr(error in MP)}}{\text{pr(error in SP)}} = \frac{\text{pr(no UC formed in MP)}}{\text{pr(no UC formed in SP)}} =$$

$$\frac{\text{pr(no UC formed in } P_1R_1I_1 \text{ of MP)pr(no UC formed in } P_2R_2I_2 \text{ of MP)}}{\text{pr(no UC formed in } P_1R_1I_1 \text{ of SP)pr(no UC formed in } P_2R_2I_2 \text{ of SP)}} =$$

$$\frac{\text{pr(no UC formed in } P_1R_1I_1 \text{ of MP)}}{\text{pr(no UC formed in } P_1R_1I_1 \text{ of SP)}}$$

$$(P_2R_2I_2 \text{ of MP} = P_2R_2I_2 \text{ of SP}) \qquad (6)$$

The key problem is to estimate the values of the probabilities in the last ratio of Equation 6. The minimum that pr(no UC formed in $P_1R_1I_1$) can be is the error rate in the appropriate Task S condition. The Task S error rate will be an underestimate since there will also be some correct responses coming from STS and from generally coded LTS traces. These two factors will thus tend to cause a slight overestimate the absolute difference between MP and SP in the number of unique codes formed in the first intervening time period.

There is no particularly good way to separate out the effects of the P_1R_1 intervals from those of the I_1 interval, however. Furthermore, the unit of consolidation is unclear although the most natural interpretation of the three consolidation mechanisms stated earlier is that the units of consolidation are entire trigrams.

If the sole level of consolidation is the trigram unit, it can be shown, in spite of the above difficulties, that consolidation must be a negligible part of SPI. At 0 seconds of rehearsal and 22 seconds of interference (Task S) there are about 15% correct responses (trigrams). Hence, the maximum proportions of trigrams that are uniquely coded at the end of I_1 is about 0.15. If we assumed that none of the unique codes are formed during the presentation period and if the rate of consolidation were constant during the interference time, then approximately 0.05 would have become uniquely coded during the first 10 seconds and 0.15 would have become uniquely coded during the total 22 seconds. Even under these extreme assumptions,

$$\frac{\text{pr(no UC formed in } P_1R_1I_1 \text{ for MP)}}{\text{pr(no UC formed in } P_1R_1I_1 \text{ for SP)}} = \frac{0.95}{0.85} = 1.12 \qquad (7)$$

Thus, consolidation predicts a ratio of 1.12 for the errors in massed practice to the errors in spaced practice which is a far cry from the observed ratio of 1.72. In the previous estimate, the constant consolidation rate probably overestimates the extent of consolidation, since to whatever extent the mechanism of consolidation is transfer from STS to LTS, there isn't much in STS after 10 seconds to consolidate. Of course, assuming that no items were uniquely coded during P_1 also leads to considerable overestimation of consolidation so that the ratio estimated should probably be something under 1.05 instead of the estimated 1.12.

However, the units of consolidation could conceivably be individual words and not necessarily trigrams and therefore the above argument might not necessarily rule out consolidation as a major determinent of SPI. For the word data, performance for 0 seconds rehearsal and 22 seconds interference is about 40% (Task S). Since 3 seconds rehearsal boosts performance to about 60%, it would seem to be a reasonable guess that not more than 25% of the total words could go into LTS during the interference period. Perhaps the most reasonable first approximation to the rate of consolidation is that as long as the item is in the STS, it has a constant probability of being consolidated (this is a version of the Atkinson and Shiffrin (1968) "buffer" model). If there is a reasonably constant rate of decay in STS, then the proportion of items "consolidated" by 22 seconds is no more than the eventual total proportion of items consolidated which is (rate of consolidation/ rate of decay). The proportion of items consolidated in the last 12 seconds can be shown to equal:

$$\text{pr(recall at 7 sec.-pr recall at 22 sec.)} \times \frac{\text{rate of consolidation}}{\text{rate of decay}}$$

$$= (.60-.40)(.25) = 0.05. \qquad (8)$$

Therefore, if 40% of the items are uniquely tagged after 22 seconds, only 35% are uniquely tagged after 7 seconds. The ratio of (40/35) = 1.14 is more sizable than the estimate from the trigram data but a long way from the observed ratio of errors for the word data (2.13). Even if the over-generous estimate of a constant consolidation rate were used, the predicted LTS values would be 40% and 27% for 22 seconds spacing and 7 seconds spacing, respectively, and the ratio of about 1.5 is still quite far off the observed error ratio.

Thus, it has been demonstrated that even with the most generous reasonable estimates the process of consolidation, if it exists, can explain at most half of SPI and probably only about 5-10% of it. There are two possible ways to salvage the consolidation theory explanation of the observed magnitude of SPI. The first is to postulate that more information is consolidated than is revealed than in our word data; i.e., bits and pieces of words are stored. This explanation can't help too much if the bits and pieces are coming largely from STS since most of the STS traces have decayed in the first 7 seconds. Furthermore, it is hard to see that these bits and pieces can help the final trigram or word retention score much since they could not have helped recall very much (i.e., performance was poor) in the 0 second rehearsal 22 second interference S condition.

An alternative argument to rescue a consolidation theory explanation of SPI is to postulate the following mechanism: traces are being consolidated into units, some codes are simultaneously being lost from working memory during I_1, but the codes are recovered during R_2. Although this kind of explanation could succeed since almost all the processes involved are unmeasurable in the present experiments, there are several problems. The first is that the small amount of forgetting for one word stimuli (Murdock, 1960) indicates that code loss is probably not too high (at most 10-15% during the 22 second period). The second is that the extremely high recall in the second test of the DT Task given recall on the first test excludes appreciable code loss from WM after a test. There is no obvious mechanism to explain why code loss should occur to an appreciable extent only after the first presentation of a stimulus and not after either the second presentation or a test.

Although it is possible that the total galaxy of consolidation theories has not been covered in this section, it seems reasonable to conclude that the notion of consolidation is unlikely to account for more than a small portion (10%) of SPI. It is not even clear that the last notion of consolidation that we discussed, as vague as it is, can explain the magnitude of SPI observed in Experiment I, it is only not shown that it can't. Almost by process of elimination, we are led to the third class of explanation, differential coding.

Differential coding of items. -- The mechanism of differential coding can have at least three different phenomenological rationales: (a) if the S thinks he knows the item well, then he doesn't bother to work as hard during the second presentation-rehearsal period; (b) if the S doesn't know which words he knows well and which he doesn't know well, then he doesn't know where to direct his effort; (c) if the S already has a code which is "bad" (i.e., not unique enough), then he is less likely to be able to think of a new one because the old one is interfering. In all three versions the basic idea is that traces other than uniquely coded traces remaining after short spacing intervals (easily retrievable "bad" codes or STS traces) will probably inhibit learning during the second presentation-recall period but not add anything appreciable themselves to the final recall.

A major problem in trying to estimate the reasonableness of these notions is to identify what is meant by the phrase "the S thinks he knows the item well." It should be clear that the phrase is not necessarily synonomous with a state of being able to recall the item if asked to do so. Upon a second presentation of a stimulus, since there are likely to be occasions when one could "recognize" a stimulus upon a second presentation and think therefore he could have recalled it when in fact he would not have. For the moment, however, let us see what would happen if it is assumed that thinking one knows an item well and being able to recall it are synonomous. Separate evaluations of the three rationales will be postponed until it is decided that some explanation of this type can explain the magnitude of the observed SPI.

As a simplification, let us assume that if the item would not have been recalled at the end of the first interference period there is a parameter, λ_a, which when multiplied by the unique code acquisition rate of the first presentation-rehearsal period gives the unique code acquisition rate for items during the second presentation-rehearsal period. Similarly, λ_b represents the multiplier that gives the acquisition rate for items during the second recall period if the item would have been recalled at the end of the first interference period. Since the overall effect of the second rehearsal period is observed to be a bit less than the first, then if P_a is the proportion of items that would not have been recalled and P_b the proportion that would have been recalled, then even though λ_a can be more than 1,

$$P_a \lambda_a + P_b \lambda_b \leq 1. \qquad (9)$$

The most severe test of differential coding theory, as of consolidation theory, is to predict performance in the 0 second rehearsal conditions. This is because there are more constraints on what can be learned the shorter the encoding time. In Task S, for 0 seconds rehearsal, the percent of trigrams recalled for 7 and 22 seconds retention interval are about 33% and 15% respectively. If 15% is approximately equal to the proportion of uniquely-coded items formed for a 0 second rehearsal period, then the number of additional codes formed for mass practice during the second presentation is:

$$\text{(proportion of items recallable but not UC)(coding rate}_b\text{)} + \text{(proportion of items not recallable)(coding rate}_a\text{)} \quad (10)$$

$$= (.18)(.15\lambda_b) + (.67)(.15\lambda_a)$$

For spaced practice, since there are essentially no recallable trigrams at the end of I_1 that are not uniquely coded, the number of additional codes formed during the second presentation period is (.85) ($.15\lambda_a$). Therefore, the extra number of items coded in the spaced condition during P_2 is $(.18)(.15)(\lambda_a - \lambda_b)$ subject to the constraint of Equation 9. If λ_b is set approximately equal to 0, then λ_a can be at most about equal to 1.27. For these values of λ_a and λ_b most favorable to differential coding theory the improvement in performance in the spaced practice condition predicted is about 3.6% which is quite a bit short of the observed difference of 9.8%.

However, the above calculation need not throw out differential coding theory. It is conceivable that the appropriate units of memory could be words instead of three word units. If words are the appropriate units of analysis, then the above calculation can be repeated using the word data instead of the trigram data. The observed difference in word retention between spaced and massed practice is 6.4%, and $\lambda_a = 1.2$ and $\lambda_b = 0.4$ are values sufficient to predict the observed retention difference and satisfy the constraint equation (Equation 9). The main problem with thinking of words as the units that are either well-coded or poorly-coded is that the general schema was built with the implicit idea that unique codes chiefly applied to three-word units. If it is essential to the formulation of the theory that unique codes are formed primarily for three-word units, then the previous calculations indicated that differential coding could explain only a fraction of SPI. There are two possible avenues of escape around this difficulty.

The first is that λ_a does not necessarily have to be constrained to be small. The data from Task DP on rehearsal suggested only that on the average, extra time (rehearsal) was a bit less effective after P_1 than after P_2. It is conceivable and even likely that, in spite of the above, the 2 seconds of P_2 are much more effective than the 2 seconds of P_1. One way to phrase this is that λ_a and λ_b are not necessarily fixed over the P_2R_2 interval. It is plausible that λ_a could be quite high at the beginning, corresponding to a feeling of urgency from knowing that the item has been forgotten, but could decrease rapidly. With such a λ_a function, the extent of SPI can be explained. For example, if $\lambda_a = 4$ and $\lambda_b = 0.4$ for P_2 in the previous trigram data calculations, then we can predict the observed error ratio for trigrams, 1.7.

Since it is likely that "thinking you know the item well" does not correspond to being able to recall it, there is a second way to rescue differential coding theory. It is probable that some weaker measure of retention (a bit more like recognition) would correspond to an impression of "thinking you know the item well." It is doubtful whether recognition is exactly the right measure either, since one can often recognize something as being the right stimuli while knowing that he couldn't have recalled it. Whatever the exact measure of "thinking that you know the item well" is, it is quite plausible that the difference between the proportion of times that "one thinks one knows the stimulus well" upon seeing it after 7 seconds interference vs. seeing it after 22 seconds interference is bigger than the corresponding difference in proportions in the Task S trigram recall data.

It should be pointed out that there has been a bit of an overestimate of SPI in assuming that all items that are recalled at 22 seconds retention interval are coming from uniquely coded items since general codes were probably contributing a bit to performance. It does not seem likely that the difference this would produce in our estimates of "proportion of well-coded items" (especially for three-word items) would be great enough to change the order of magnitude of percentages involved and could be compensated for by small changes in λ_a. Thus, either of the two loopholes above can easily allow us to explain the magnitude of SPI with differential coding theory.

One interesting piece of supporting evidence for a differential coding theory is that the difference between spaced and massed practice at each retention interval and the difference between Task S performance at 7 and 22 seconds are both relatively unchanged with varying R_1 intervals. On the other hand, neither R_1 nor R_2 interacts significantly with the spacing interval (I_1). This non-interaction is one piece of data which casts some shadow on a differential coding theory since one would expect that if spacing made P_2 potent, it would make R_2 more potent as well. The explanation offered before was that the differential coding effect wears off by the end of the 2 second presentation interval. A possible test of differential coding is the existence of an interaction of R_2 and I_1 for shorter stimulus presentation periods.

To outline the argument of the last two sections, the key piece of data for both consolidation theory and differential coding theory to explain is the large difference between massed and spaced practice when both R_1 and R_2 are 0 seconds. Both theories rely on using the moderate difference between what the subject has retained after 7 seconds and what he has retained after 22 seconds in Task S and the fairly low level of performance after 22 seconds in Task S to explain SPI. There are two key points on which the discussion turned. The first is that retention measures from Task S probably gave a reasonable measure of what could possibly have been consolidated but only gave rough estimates of the state of "thinking that you know an item well." The second is that the low retention score after 22 seconds of interference in Task S puts greater constraint on consolidation theory than on differential coding theory since consolidation is a process whose differential effect in spaced vs. massed practice takes place before the end of the first interference period (i.e., during the span of Task S).

It has not been specified which of the three phenomenological interpretations mentioned at the beginning of this section should be placed on differential coding, and the data do not allow those effects to be distinguished easily. In fact, it is not clear that they are easily separated conceptually. One possible way to test whether mechanism (b) (i.e., he knows which part of the stimulus to work on) is a significant part of differential coding would be to see whether commonly missed words in a single presentation condition are not commonly missed in a spaced double presentation condition. Here, positive results would be interesting, negative results would not be.

Relation to other spacing data. -- Perhaps the discussion of spaced practice is best concluded with some speculation on the relationship of spaced practice effects observed in the present experiments to other observed spacing effects. In the present experiments, the process of consolidation is an unlikely determinant of a major part of SPI. Even though some covert rehearsal is probably going on during the interference, it is intuitively fairly unlikely that much of it can be conscious recoding of the stimulus; it is just too difficult to do that and count backwards too. The conscious processing going on during interference in the B-P paradigm is probably in the form of attending to the stimulus trace as it was coded during the presentation-rehearsal period.

On the other hand, in PA learning, it is reasonably likely that the subject takes several PA pairs and keeps on rehearsing them (even while others are presented) until a UC is formed for those pairs. To whatever extent this process is going on, one would expect "consolidation" (i.e., formation of a UC) going on during a period that is nominally an interference period in the experiment. Thus, it is plausible that "consolidation" could play an important role in spacing effects in PA learning (Greeno, 1968; Atkinson & Shiffrin, 1967) and not play any significant role in spacing effects in the B-P paradigm.

The declining effectiveness of spaced practice in PA learning beyond a certain optimal spacing interval has a natural consolidation interpretation: at some point, the effects of LTS decay start to outweigh the effects of consolidation. However, there is also a differential coding interpretation: the second presentation will have optimal effect if the trace has decayed enough so that enough of the item is lost for the S to know what has not been well-coded but not decayed so much that the whole item is so difficult to retrieve that the subject does not know what is relatively well-coded. (In PA experiments, the dichotomy between completely unique and general codes is probably unrealistic.)

So far, no inverted "U"-shaped curve of retention as a function of spacing interval has been found in B-P or FR experiments. However, no really long spacing intervals have been employed in those two paradigms. If it is found that retention continues to improve or levels off as the spacing interval is made longer in B-P or FR experiments, then further evidence would exist that the processes underlying spacing in the PA paradigm are somewhat different than in the other two paradigms. Much work needs to be done to discover exactly what is causing spacing effects and to discover whether the spacing effects observed in the three paradigms have more than surface similarity.

CHAPTER V

SUMMARY AND CONCLUSIONS

The major experimental findings of the present experiments were the following. (a) Covert rehearsal greatly improved retention for all retention intervals studied (up to 22 seconds). This result contradicted an earlier finding of Peterson and Peterson (1959) that only overt rehearsal improved retention. (b) Increasing the amount of rehearsal allowed produced progressively more sigmoid forgetting curves in Task S. (c) In Task DT recall on the second test was essentially identical to the recall on the first test. (d) In Task DP there were about twice as many errors when the intervening interference period was 7 seconds (massed) as when it was 22 seconds (spaced). (e) In Task F, there was a slight improvement in performance when more time was allowed for rehearsal of the first item, and also when the two items were separated by 22 seconds of interference (spaced) rather than by 7 seconds (massed). (Both these effects bordered on significance.) (f) In Task C, there was a significant improvement in retention when the initial interference period was 27 seconds as opposed to 12 seconds. (g) Retention in Task C was significantly better than in comparable conditions of Task F.

The data were analyzed within a conceptual framework that distinguished between four types of storage: working memory (WM), a short-term acoustic store (STS), items in a long term store (LTS) that could be easily retrieved (UC), and items in LTS that suffered from response interference (GC). Within this framework, it was concluded that the primary effect of rehearsal is to allow for more unique codings in LTS. This conclusion was deduced from the large improvement in performance resulting from rehearsal of the to-be-remembered item as contrasted with the lack of decrement in performance (in fact a slight improvement) resulting from rehearsal of an interfering item. The increasingly sigmoid retention curves with increasing rehearsal (in Task S) suggested that rehearsal was fortifying the STS trace somewhat.

From the discussion on spacing in Task DP, it was decided that (at least for the spacing and retention intervals studied here) consolidation was an unimportant process, and therefore that there was no reason to believe that interference had any effect other than to produce decay. Since reduction of interference from competing memory traces was also shown to be an inadequate explanation of spacing, it was concluded that the major cause of spaced practice improvement in Task DP was that the longer spacing interval allowed for more effective encoding of the stimulus at the second presentation.

The results of the present experiments thus offered general confirmation for a multistore memory system and also supported the existence of a response competition process in short-term memory. The present experiments also provided evidence that the primary function of rehearsal is to allow for recoding of the stimulus, with strengthening of stimulus traces a process of secondary importance.

APPENDIX

Stimuli

The same 72 word trigrams were used as the to-be-remembered items in Experiments I and II. They are listed in the order that they were presented in both the experiments.

1)	FORK	DECK	MILL	25)	SALT	RING	DUST	49)	PAGE	BEAN	VEST
2)	DOME	CROW	PLOT	26)	DEED	FROG	GOAL	50)	BAND	RISK	LARK
3)	CASE	MAST	BOSS	27)	BULK	LUNG	CROP	51)	BOOM	JAIL	COAT
4)	MONK	CODE	FUND	28)	VIEW	FISH	DEBT	52)	NOTE	HEAP	TWIN
5)	RICE	GAZE	FOOT	29)	CAGE	SEED	DUKE	53)	GRIN	TENT	PART
6)	TANK	MASS	CARD	30)	FOAM	STEM	COLD	54)	TERM	FIRE	ROOT
7)	ACHE	WOOL	LOCK	31)	SPOT	BARN	LIST	55)	MASK	PINE	BOOK
8)	LACE	CREW	TREE	32)	FOOD	NAIL	POND	56)	YARD	DRUM	SOAP
9)	SONG	GULF	DOOR	33)	CHIP	DOVE	REAR	57)	LIMB	CAPE	GANG
10)	MOSS	SASH	BOAT	34)	RULE	POST	COMB	58)	TOOL	MILE	FATE
11)	WEST	BACK	WAGE	35)	BOND	RACE	WEED	59)	WISH	GOAT	HARP
12)	LAKE	SKIN	DAWN	36)	SEAT	WOLF	PILE	60)	LOOM	MULE	BROW
13)	BONE	TRIP	SACK	37)	BATH	CAVE	FLAG	61)	BAIT	ROOF	MESS
14)	STAR	CHIN	TUBE	38)	BIRD	HOME	PLUM	62)	MOLD	EAST	FILE
15)	GALE	FOOL	BENT	39)	GRIP	WALL	INCH	63)	MIND	SWAN	WHIP
16)	PIPE	SELF	EDGE	40)	RACK	POOL	GIFT	64)	TOUR	CARE	LAMP
17)	SCAR	VINE	HOOF	41)	DUCK	LINE	PACE	65)	HOOK	LAND	PUMP
18)	LEAF	FUEL	WORD	42)	TILE	SHIP	LUMP	66)	HILL	BELT	SEAL
19)	TIME	CLAY	BOLT	43)	GLUE	TONE	SPAN	67)	MODE	BEAK	LAWN
20)	NOON	ARCH	PLAN	44)	TRAY	ZEAL	DOCK	68)	SHOE	DISH	ROSE
21)	FORM	DOLL	BANK	45)	YEAR	CORD	STEP	69)	DOOM	BARK	FARM
22)	MILK	FIST	GOWN	46)	PALM	LOAF	WING	70)	CART	MOON	SOUP
23)	HAWK	RUST	LIFE	47)	CLUB	BEAD	LUCK	71)	ZONE	SHED	TRAP
24)	TOWN	BEAM	COLT	48)	SAGE	LOSS	BOOT	72)	SILK	GOLD	COOK

The 36 word trigrams used for the "forget" items in Experiment II are listed in the order that they were presented.

1	WORM	BOMB	DUNE	13	RIND	BELL	FLAW	25	STUB	BULL	TEST
2	FUSE	DIME	CORK	14	SAND	NEWS	TASK	26	GERM	VASE	DATE
3	DRUG	SNOW	COIN	15	TUNE	PILL	MOLE	27	WARD	DESK	HEAT
4	SUIT	LOOP	HOSE	16	BRIM	SLOT	MOTH	28	FERN	WASP	PULP
5	WAND	RAMP	FAIR	17	DOSE	REEF	LASH	29	WEST	LARD	COIL
6	NAME	HALL	MEAL	18	HAND	CRAB	OATH	30	CUBE	KING	LENS
7	GEAR	DIKE	PACT	19	MARE	TART	SHIN	31	SLED	HERB	TOMB
8	VERB	TURF	MOOD	20	GAME	COST	LASH	32	FOLK	SLAB	NOUN
9	WIFE	RAFT	NECK	21	MARK	PORT	HOST	33	FAME	DISK	HIVE
10	DIRT	SHOP	FACT	22	LOIN	CRIB	RINK	34	SOIL	ROPE	DOWN
11	ROOM	WIND	KNEE	23	CORN	RAIL	JOKE	35	ROCK	GIRL	LIME
12	SOOT	MINK	ZEST	24	FLAP	MINE	CAKE	36	FORT	GRID	CALF

The 8 word trigrams used in the practice trials are listed below. The first 6 were used in both experiments and the last two were used only in Experiment II.

1)	HARM	CANE	BILL	4)	TIRE	FACE	BEEF	7)	SAFE	GRAM	MALT
2)	PATH	FILM	WIRE	5)	HEAD	GLOW	YARN	8)	WICK	LORD	STEW
3)	BUSH	HORN	CAMP	6)	BITE	TWIG	ROBE				

BIBLIOGRAPHY

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 2, New York, Academic Press, 1968.

Averbach, E., & Coriell, A. S. Short-term memory in vision. Bell Sys.Tech. J., 1961, 40, 309-328.

Baddeley, A. D. Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. Quart. J. Exp. Psychol., 1966a, 18, 362-265.

Baddeley, A. D. The influence of acoustic and semantic similarity on long-term memory for word sequences. Quart. J. exp. Psychol., 1966b, 18, 302-309.

Bennett, R. The release of proactive inhibition as a function of prior discriminative training. Unpublished Master's thesis, 1969, University of Michigan.

Bjork, R. A. Learning and short-term retention of paired-associates in relation to specific sequences of interpresentation intervals. Inst. for math. stud. in soc. sci., Tech. Rep. No. 196. Stanford Univ., Stanford, California, 1966.

Bjork, R. A. The effect of instructions to selectively forget during short-term memory. Memorandum Report, 1969, Human Performance Center, University of Michigan (in press).

Bjork, R. A., & Abramowitz, R. L. The optimality and communitivity of successive interpresentation intervals in short-term memory. Paper presented at the Midwest Psychological Association Convention, Chicago, 1968.

Broadbent, D. E. Perception and Communication. New York: Pergamon, 1968.

Brown, J. Some tests of the decay theory of immediate memory. Quart. J. exp. Psychol., 1958, 10, 12-21.

Corman, C. D., & Wickens, D. D. Retroactive inhibition in short-term memory. J. verb. Learn. verb. Behav., 1968, 7, 16-19.

Conrad, R. Serial order intrusions in immediate memory. Brit. J. Psychol., 1960, 51, 45-48.

Conrad, R. Acoustic confusions and memory span for words. Nature, 1963, 197, 1029-1030.

Conrad, R. Acoustic confusion and immediate memory. Brit. J. Psychol., 1964, 55, 75-84.

Glanzer, M. Storage mechanisms in free recall. Trans. N. Y. Acad. Sci., 1968, Series II, Vol. 30, 1120-1129.

Glanzer, M. Distance between related words in free recall: Trace of the STS. J. verb. Learn. verb. Behav., 1969 (in press).

Glanzer, M., & Cuntiz, A. R. Two storage mechanisms in free recall. J. verb. Learn. verb. Behav., 1966, 5, 351-360.

Greeno, J. G. Paired-associate learning with massed and distributed repetitions of items. J. exp. Psychol., 1964, 64, 286-295.

Greeno, J. G. Paired-associate learning with short-term retention: Mathematical analysis and data regarding identification of parameters. J. Math. Psychol., 1967, 4, 430-472.

Hellyer, S. Supplementary report: Frequently of stimulus presentation and short-term decrement in recall. J. exp. Psychol., 1962, 64, 650.

Hofer, R. Intertrial proactive inhibition in short-term memory. Psychol. Rept., 1965, 17, 755-760.

Keppel, G., & Underwood, B. J. Proactive inhibition in short-term retention of single items. J. verb. Learn. verb. Behav., 1962, 1, 153-161.

Landauer, T. K. Reinforcement as consolidation, Psychol. Rev., 1969, 76, 83-96.

Loess, H. Short-term memory, word class, and sequence of items. J. exp. Psychol., 1967, 74, 556-561.

Loess, H. Short-term memory and item similarity. J. verb. Learn. verb. Behav., 1968, 7, 87-92.

Loess, H., & Waugh, N. C. Short-term memory and intertrial interval. J. verb. Learn. verb. Behav., 1967, 6, 455-460.

Luce, R. D. Individual Choice Behavior. New York: Wiley, 1959.

Melton, A. W. Implications of short-term memory for a general theory of memory. J. verb. Learn. verb. Behav., 1963, 2, 1-21.

Melton, A. W., & Shulman, H. C. Further studies of a distributed practice effect on probability of recall in free recall. Paper read at Psychonomic Society Meeting, September, 1967.

Miller, G. A. The magical number seven, plus or minus two: Some limits of our capacity for processing information. Psychol. Rev., 1956, 63, 81-97.

Murdock, B. B. The retention of individual items. J. exp. Psychol., 1961, 62, 618-625.

Murdock, B. B. The serial position effect of free recall. J. exp. Psychol., 1962, 62, 482-488.

Murdock, B. B. Short-term memory of single paired associates. J. exp. Psychol., 1963, 65, 433-443.

Neisser, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

Noyd, D. Proactive and intrastimulus interference in short-term memory for two-, three-, and five-word stimuli. Paper presented at the Annual Meeting of the Western Psychological Association. Honolulu, Hawaii, 1965.

Peterson, L. R. Immediate memory: Data and theory. In C. H. Cofer & B. S. Musgrave (eds.), Verbal Behavior and Learning: Problems and Processes. New York: McGraw-Hill, 1963.

Peterson, L. R., & Gentile, A. Proactive interference as a function of time between tests. J. exp. Psychol., 1965, 70, 473-479.

Peterson, L. R., Hillner, & Saltzman, D. Supplementary report: Time between pairings and short-term retention. J. exp. Psychol., 1963, 64, 550-551.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. J. exp. Psychol., 1959, 58, 193-198.

Peterson, L. R., Saltzman, D., Hillner, K., & Land, V. Recency and frequency in paired-associate learning. J. exp. Psychol., 1962, 63, 396-403.

Pollack, I. Interference, rehearsal, and short-term retention of digits. Canad. J. Psychol., 1963, 17, 380-392.

Posner, M. I. Information reduction in the analysis of sequential tasks. Psychol. Rev., 1964, 71, 491-504.

Posner, M. I. Components of skilled performance. Science, 1966, 152, 1712-1718.

Postman, L., & Phillips, L. W. Short-term temporal changes in free recall. Quart. J. exp. Psychol., 1965, 17, 132-138.

Reid, L. S. Information processing and short-term memory. Charlottesville, Va., University of Virginia, 1967. Progress Report No. 1, NSF Grant GB-4069.

Rumelhart, D. E., The effects of interpresentation intervals on performance in a continuous paired-associate task. Inst. for math. stud. in soc. sci., Tech. Rep. No. 116, Stanford University, Stanford, Calif., 1967.

Sperling, G. The information available in brief visual presentation. Psychol. Monog., 1960, 74, No. 11.

Treisman, A. M. Monitoring and storage of irrelevant messages in selective attention. J. verb. Learn. verb. Behav., 1964, 3, 449-459.

Underwood, B. J. Interference and forgetting. Psychol. Rev., 1957, 64, 49-60.

Waugh, N. C. Immediate memory as a function of repetition. J. verb. Learn. verb. Behav., 1963, 2, 107-112.

Wickelgren, W. A. Phonemic similarity and interference in short-term memory for single letters. J. exp. Psychol., 1966, 71, 396-404.

Wickens, D. D., Born, D. G., & Allen, C. K. Proactive inhibition and item similarity in short-term memory. J. verb. Learn. verb. Behav., 1963, 2, 440-445.

Wickens, D. D., & Clark, S. Osgood dimensions as an encoding class in short-term memory. J. exp. Psychol., 1968, 78, 580-584.

Young, J. L. Effects of intervals between reinforcements and test trials in paired-associate learning. Inst. for math. stud. in soc. sci., Tech. Rept. No. 101, Stanford University, Stanford, California, 1966.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

University of Michigan, Human Performance Center
Department of Psychology, Ann Arbor, Michigan

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

Rehearsal, Interference, and Spacing of Practice in Short-Term Memory

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Scientific Interim

5. AUTHOR(S) *(First name, middle initial, last name)*

Alexander Warren Pollatsek

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July, 1969	112 + ix	54

8a. CONTRACT OR GRANT NO. AF 49(638)1235 (ARPA)
~~AF 49(638)1736 (ARPA)~~

b. PROJECT NO. 5002

c. 61101D

d. 681313

9a. ORIGINATOR'S REPORT NUMBER(S)

Technical Report No. 16
08773-41-T

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

AFOSR 69-2798TR

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES

TECH, OTHER

12. SPONSORING MILITARY ACTIVITY

Air Force Office of Scientific Research
1400 Wilson Boulevard (SRLB)
Arlington, Virginia 22209

13. ABSTRACT

Short-term memory as exhibited in a variety of experimental paradigms is heavily influenced by variations in the time allotted for rehearsal, in the time allotted for rehearsal-preventing interfering activity, and in the interval separating successive presentations of an item. This dissertation was designed to investigate systematically any interactions among rehearsal, interference, and spacing of practice in an effort to derive constraints on an adequate theory of short-term memory beyond those imposed by prior research.

Two experiments were conducted, each of which employed variations on the Brown-Peterson experimental paradigm. The results of Experiments I and II together with the results of other experiments suggest the following constraints on an adequate theory of memory: (a) there are at least two memory stores; (b) there is response competition in the long-term store (LTS); (c) some items in LTS are much more retrievable than others.

The data from Experiments I and II were analyzed in detail with respect to a specific theory designed to be consistent with the above constraints. Two major conclusions were drawn from the anlaysis: (a) rehearsal provided opportunity for more unique coding of LTS traces but did not necessarily strengthen them; (b) neither reduction of proactive interference nor consolidation could explain the effects of spacing in the Double-Presentation condition, but an explanation in terms of increased coding effectiveness with increased spacing seemed plausible.

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
1. Short-Term Memory						
2. Rehearsal						
3. Spacing of Practice						
4. Verbal Encoding						
5. Consolidation						

Unclassified
Security Classification

TECHNICAL REPORTS

1. Phillips, L. D. Some components of probabilistic inference. January 1966.
2. Egeth, H. E. Parallel versus serial processes in multidimensional stimulus discrimination. January 1966.
3. Biederman, I. Human performance in contingent information processing tasks. October 1966.
4. Peterson, C. R. & Beach, L. R. Man as an intuitive statistician. November 1966.
5. Martin, E. & Melton, A. W. Recognition memory for CCC and CVC trigrams of various association values. January 1967.
6. Smith, E. E. Choice reaction time: An analysis of the major theoretical positions. January 1967.
7. Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. February 1968.
8. Ligon, E. The effects of similarity on very-short-term memory under conditions of maximal information processing demands. May 1968.
9. Triggs, T. J. Capacity sharing and speeded reactions to successive signals. August 1968.
10. Tversky, B. G. Pictorial and verbal encoding in short-term memory. October 1968.
11. Lively, B. L. The Von Restorff effect in very-short-term memory. December 1968.
12. Greeno, J. G. How associations are memorized. December 1968.
13. Swensson, R. G. The elusive tradeoff: Speed versus accuracy in choice reaction tasks with continuous cost for time. December 1968.
14. Bjork, R. A. Repetition and rehearsal mechanisms in models for short-term memory. May 1969.
15. Kamlet, A. S. Processing of sequentially presented signals in information-combining tasks. June 1969.

MEMORANDUM REPORTS

1. Fitts, P. M. Cognitive factors in information processing. February 1967.
2. Melton, A. W., Sameroff, A., & Schubot, E. D. Short-term recognition memory. May 1967.
3. Martin, E. Responses to stimuli in verbal learning. October 1967.
4. Melton, A. W. First annual report: Human information handling processes. June 1968.
5. Jahnke, J. C. The Ranschburg paradox. July 1968.
6. Greeno, J. G. Theory of graphs on sets with applications to problem solving and understanding. October 1968.
7. Edwards, W. A bibliography of research on behavioral decision processes to 1968. January 1969.