INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in “sectioning” the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND
WRIGHT, JON RODNEY
PROCESSING STRATEGIES AND REPETITION EFFECTS
IN FREE RECALL.

RICE UNIVERSITY, PH.D., 1979
RICE UNIVERSITY

Processing Strategies and Repetition Effects in Free Recall

by

Jon R. Wright

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

John W. Breisford, Chairman, Thesis Committee
Professor, Psychology Department

William C. Howell, Professor
Psychology Department

David A. Schum, Professor
Psychology Department

Donald I. Wood, Professor
Education Department

HOUSTON, TEXAS
May, 1979
PROCESSING STRATEGIES AND REPETITION EFFECTS

IN FREE RECALL

Jon R. Wright

Abstract

A broad range of research on basic memory processes has established that memory performance is dependent on the spacing of stimulus repetitions (cf., Hintzman, 1974, 1976). Typically, the function relating memory performance to the spacing of repetition is monotonically increasing and negatively accelerated. The spacing effect is theoretically interesting for several reasons. First, the form of the spacing function suggests that traditional assumptions about learning as an accumulative process may be in error. Secondly, the spacing effect affords a convenient vehicle to ask general questions about the mechanisms underlying stimulus repetition effects in memory. These questions are addressed within a theoretical framework which emphasizes an information processing approach to human learning. This approach views human learning in terms of active rather than passive processes. In particular, the course of human learning is considered to be governed by control processes which function to guide the flow of information within the learning environment. Within this context, it is useful to consider the way in which human subjects adjust their control processes to changing task demands.

The role of the subject-controlled processing strategies in the spacing effect was investigated in three related experiments. The first two experiments manipulated processing strategies by introducing items of differential value into a free recall task. The results of these experiments indicated that subjects adjust their processing
strategies to give extra processing to highly valued items, and that the form of the spacing function is sensitive to this adjustment in strategy. Post hoc analyses indicated that effective processing strategies serve to introduce to-be-learned material into a variety of stimulus contexts. These results were interpreted in terms of the encoding variability theory of spacing effects proposed by Glenberg (1976).

The third experiment extended the interpretation of Experiments I and II to the incidental learning paradigm. In this experiment, contextual variability was manipulated in a concreteness judgment task. The results supported the theoretical interpretation of the spacing effect in terms of encoding variability.
ACKNOWLEDGMENTS

In many ways, the completion of a doctoral education represents something which exceeds the efforts of a single individual. It is attributed more appropriately to a community of faculty and graduate students who represent the true resources and value of higher education. The author wishes to express his sincere gratitude to that community for its understanding and encouragement during his doctoral research and throughout his graduate career. In addition, special appreciation is due to William C. Howell, David A. Schum, and Donald Wood who, as members of the author's dissertation committee, provided timely advice and perceptive criticism of the research which follows.

The author especially wishes to acknowledge his indebtedness to John W. Brelsford who served in the multiple roles of teacher, advisor, collaborator and friend in all phases of the author's graduate career. Above all, a graduate student wishes for a research advisor whom he can respect and trust both professionally and personally. The author feels particularly fortunate in this regard.

Finally, the author wishes to record the contribution of his wife, Lynda, who understood the extra dedication, uncertainties, and special hardships necessarily a part of graduate student life. Her presence made the difficult times easier and turned the good times into truly happy experiences.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Methodological Considerations and Descriptions of Spacing Phenomena</td>
<td>9</td>
</tr>
<tr>
<td>Theories of the Spacing Effect</td>
<td>16</td>
</tr>
<tr>
<td>Deficient Processing Theories: Experimental Evidence</td>
<td>41</td>
</tr>
<tr>
<td>Experiment I</td>
<td>57</td>
</tr>
<tr>
<td>Experiment II</td>
<td>69</td>
</tr>
<tr>
<td>Experiment III</td>
<td>87</td>
</tr>
<tr>
<td>General Discussion</td>
<td>97</td>
</tr>
<tr>
<td>References</td>
<td>105</td>
</tr>
</tbody>
</table>
INTRODUCTION

Improvement in performance with repetition is fundamental to virtually every task that has been studied by experimental psychologists. Historically, such dynamic changes in behavior have been the domain of learning theorists. In recent years, however, the study of learning, especially the study of human verbal learning, has been largely replaced by the study of memory. It is currently fashionable to speak of stimulus repetition effects in memory when one is addressing questions that formerly would have been considered to be the exclusive domain of the learning theorist (cf. Crowder, 1976, chapter 9; Hintzman, 1976). To the uninitiated, the change in vocabulary from that of the learning theorist to that of the memory theorists may seem to be something of an academic word game—there is a very real continuity in the basic questions that are asked and the techniques that are employed to answer them. However, this change is symptomatic of a rather profound difference in the way the experimental psychologist views human behavior. To state it simply, human beings are no longer seen as the passive products of environmental forces, but as active seekers, manipulators, and users of environmental information. Human beings, thus, are being viewed as processors of information.

Within the information processing framework, the ways in which people learn about their environment are no less important than in the past, although research hypotheses tend to be stated in somewhat different terms. Traditional reinforcement variables, for example,
are no longer thought to have the automatic, stamping-in properties they were once thought to have. Instead, such variables are analyzed in terms of their informational content and incentive value (Estes, 1969). Particular attention is paid to the ways in which basic information processing abilities are brought to bear on the task structure set up by the experimenter. From this point of view, the function of stimulus repetition may be considered to be one of providing opportunities to elaborate the nature of the stored representations of information in memory.

Despite the change in theoretical orientation, the basis for improvement in memory due to stimulus repetition is still not well understood. Traditional theories have treated learning as the accumulation of some underlying theoretical construct (e.g., habit strength). This theoretical approach is perhaps most apparent in the abstract characterizations of the learning process as proposed by mathematical learning theorists. For example, the linear model (Bush and Mosteller, 1955) describes the trial to trial increments in performance during a learning experiment as a proportion θ of the difference between the strength of the response on trial n and some performance asymptote. In the most familiar case, where the learning process is assumed to asymptote at Pr(correct) = 1.0, the difference equation describing trial to trial changes in performance is

\[ p_{n+1} = p_n + \theta(1 - p_n). \]

Linear models are widely descriptive of many simple learning situations.
It is, however, quite clear that such models, based on the idea of accumulation of response strength, are not capable of handling certain types of distributional and sequential phenomena that consistently appear in human learning data. In fact, certain of these distributional phenomena directly violate basic assumptions of these models. It is thus worthwhile to examine some of these more subtle aspects of learning data to see how they affect learning theories.

The effects of spacing or distributing stimulus repetitions are currently well known and have received a considerable amount of attention over the past decade (cf. Crowder, 1976, chapter 9). In the most typical finding, performance has been shown to be an increasing monotonic function of the spacing or distribution of repetitions. The effects are quite general, appearing in a large variety of experimental paradigms and with many different types of stimulus material (cf. Hintzman, 1974, 1976).

These spacing effects reflect an interesting paradox between the hypothetical strength of an item at the time stimulus repetition occurs and the probability of correct responding when its status in memory is assessed at the conclusion of an experiment. Figure 1 will help to elucidate this relationship. For purposes of clarity, the learning situation of reduced to its lowest common denominator, a single stimulus presentation (denoted P1), a single repetition (denoted P2), and a test.

Several features of this figure require comment. A fundamental property of human memory for simple, well-integrated stimuli is that
Figure 1. Left panel represents the predictions of strength theory. The spacing effect suggests the relationships shown in the right panel.
immediately following the presentation of a to-be-remembered (TBR) stimulus, the item is "learned" in the sense that the probability of a correct response is near 1.0. However, this "learned" state is highly transient. Under most conditions, large amounts of forgetting take place within a few seconds of stimulus presentation. Under some conditions, the probability of a correct response can fall near to zero within 15 seconds (Peterson and Peterson, 1959). With stimulus repetitions, performance shows a similar decline in the interval following stimulus presentation, but appears to asymptote at successively higher levels (Hellyer, 1962).

In accordance with these basic results, Figure 1 indicates that "memory strength" immediately following P1 and P2 is at a maximum and then declines rapidly with time. Increased probability of retrieval resulting from stimulus repetition is reflected in the asymptotic performance following P2. The predictions of traditional strength theory are shown in the left panel. Because the strength of the TBR item is higher for massed than for spaced repetition, simple strength theories based on some accumulation process predict that massed presentation should be superior to spaced presentation. In actual fact, however, the data suggest the relationships depicted in the right panel. The effectiveness of stimulus repetition appears to depend on how much forgetting has taken place in the P1-P2 interval. The more forgetting, the more effective is stimulus repetition.

As noted by Bjork (1970), these relationships would seem to eliminate any model which preserves the order of the relative strengths
of items over repetitions. That is, any model that always transforms the memories of two items having the recall probabilities $P(i)<P(j)$ at the time of repetitions into new memories with the recall probabilities $P'(i)<P'(j)$ cannot account for the effects of spacing repetitions. Many well researched models based on the primitive notion of response or memory strength, including the linear model cited earlier, have precisely this character. For this reason, the spacing effect has been cited as a refutation of strength theories of the memory trace (Crowder, 1976, p. 276).

Figure 1 illustrates the theoretical advantage gained by employing the concepts of the memory theorist. At least in the simple kinds of tasks that have been examined in the laboratory, learning is best understood in terms of the interplay between forgetting and acquisition processes. All learning theories must come to account for the fact that TBR information appears to enter some transient state of heightened accessibility before it becomes permanently acquired.

Since learning is viewed as a tradeoff between trial-to-trial forgetting and acquisition, it is no accident that learning theorists have become interested in experimental designs which make forgetting a salient part of the behavior being studied. Such designs employ variations of the simple events depicted in Figure 1.

An experiment by Bjork and Allen (1970) is an excellent illustration and provides an empirical demonstration of the analysis suggested by Figure 1. In a modified Brown-Peterson paradigm, subjects received two presentations of noun trigrams. During the spacing interval be-
between P1 and P2, and the retention interval between P2 and test, subjects were asked to perform a shadowing task in which they had to call out the names of numbers as they appeared in the viewslot of a memory drum. The function of the shadowing task was to exercise control over the study periods and to prevent rehearsal. The difficulty of the shadowing task during the spacing interval and the length of the spacing interval were manipulated factorially. An important feature of the experiment was that on some trials a recall test was inserted into the procedure at the time when P2 normally occurred. Performance on the single presentation trials could then be used to infer the status of the noun trigrams in memory when P2 occurred on the double presentation trials.

The Bjork and Allen data are reproduced in Table 1. As might be expected, a clear effect of shadowing task difficulty was found on the single presentation trials. Recall was lower for the difficult than for the easy task. For the double presentation trials, a spacing effect was obtained for both the easy and difficult distractor tasks. But, more importantly, there was a higher absolute level of recall plus an enhancement of the spacing effect for the difficult task. This is precisely the condition were performance on the single presentation trials indicated that item strength was lowest. These data suggest that the effectiveness of stimulus repetition is somehow related to the amount of forgetting taking place between presentations. Bjork and Allen manipulated forgetting in the spacing interval in two ways—first, by increasing the difficulty of the distractor task, and second,
Single Presentation Trials

<table>
<thead>
<tr>
<th>Task</th>
<th>4.5 sec</th>
<th>18 sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
<td>54</td>
<td>45</td>
</tr>
<tr>
<td>Hard</td>
<td>39</td>
<td>32</td>
</tr>
</tbody>
</table>

Double Presentation Trials

<table>
<thead>
<tr>
<th>Task</th>
<th>4.5 sec</th>
<th>18 sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
<td>61</td>
<td>70</td>
</tr>
<tr>
<td>Hard</td>
<td>63</td>
<td>77</td>
</tr>
</tbody>
</table>

Table 1. Percent recall as a function of interval length, task difficulty, and trial type in the Bjork and Allen experiment (from Bjork and Allen, 1970).
by increasing the length of the spacing interval itself. Apparently both variables exercised some control over the effectiveness of the repetition at P2.

Research on spacing phenomena seems capable of restricting the domain of theories that account for repetition effects. As has been pointed out, the simplest types of mechanisms based upon primitive notions of memory or trace strength can be rejected. There are, in addition, some interesting features of the spacing effect data that may serve to restrict theories of repetition even further. The Bjork and Allen study indicates that the events at P2 are particularly important for the understanding of repetition effects. Because the effectiveness of stimulus repetitions appears to be somewhat dependent on the amount of forgetting between presentations, it seems natural to expect that an adequate theory will stress some kind of compensatory process which adjusts to the accessibility of information present at P2. This point of view will subsequently be developed in more detail.

Having established the importance of the spacing effect as it relates to the learning process, it is appropriate to examine certain considerations related directly to the effect.

METHODOLOGICAL CONSIDERATIONS AND DESCRIPTIONS OF SPACING PHENOMENA

From an historical perspective, several distinct phases of research on the spacing or distribution of stimulus repetitions in human learning can be identified. During the 1940's and 50's, most of the research on
this question was conducted within the framework of the paired-associates learning and transfer paradigms which dominated the research on human learning and memory during that period. Such techniques involved presenting subjects with short lists of paired stimulus-response items for memorization and subsequently asking for the appropriate response to each stimulus item. The dependent measure was typically some global measure of performance such as trials to a criterion of list mastery. A trial consisted of a randomized presentation of an entire list of paired associates. The basic empirical question concerned whether the introduction of a rest period between trials (distributed practice) facilitated or inhibited the acquisition of the appropriate responses. In general, these efforts produced a highly conflicting and disappointing body of research. In an important review paper, Underwood (1961) concisely summarized the research in this area and concluded that the distribution of paired-associates trials facilitated acquisition only under a limited set of conditions, e.g., in those cases where substantial amounts of responses integration were necessary.

The interest generated in the one-trial vs. incremental learning controversy by the drop-out experiment of Rock (1957) and Estes' miniature experiments (Estes, 1961) rather dramatically changed the kinds of procedures that were used to study paired-associates learning. These new procedures analyzed acquisition performance on a more molecular level than had been the case previously. Primary attention was directed at the "fate" of individual items rather than at the acquisition of entire lists of paired associates. The result was that experimental
psychologists turned their attention to manipulation of items within lists as opposed to the manipulation of lists of paired associates. In this context, the effects of distribution repetitions are found to be large and robust across a large variety of experimental procedures and stimulus material. In paired-associates learning, these effects were first reported by Peterson, Hillner and Saltzman (1962). The basic results have since been extended to free recall (Melton, 1967), the Brown-Peterson distractor task (Koeppel and Underwood, 1963), verbal discrimination learning (Ciccone, 1974), recognition memory (Ciccone and Brelsford, 1974), a composite of the Brown-Peterson task and free recall (Whitten and Bjork, 1977), frequency judgments (Hintzman and Block, 1970), and to various types of incidental learning tasks (Rowe and Rose, 1976; Rose and Rowe, 1977; Shaughnessy, 1976).

There are several reasons why the techniques employed prior to 1961 failed to uncover any consistent effects of spacing. The reasons are straightforward but require a short discussion of the procedures currently employed in laboratory research on repetition effects. The typical sequence of events in an experiment on the spacing effect is illustrated in Figure 2. In many experiments, a large number of items is presented sequentially to the subject, and both the spacing interval and the retention interval are filled with the presentation of additional TBR items. Thus, critical items of various sorts are interleaved throughout the stimulus presentation. Although spacing can be indexed by the amount of time elapsing between repetitions, many studies make use of an indexing variable, interpresentation lag, which is defined as
Figure 2. Sequence of events in an experiment on the spacing effect. See text for explanation.
the number of items or events intervening in the spacing or retention interval.

When the presentation order of a paired-associates list is randomized before repeating the stimulus presentations, lag is a confounded variable. This was the common procedure prior to 1961. The distribution of lag values under these conditions is triangular with a mode of list length minus one. For example, with a list length of ten paired-associate items, the distribution of lag values has a mode of nine with few lag values in the range zero to five. On the other hand, the spacing function for paired associates learning is negatively accelerated with most of the increase due to spacing taking place over this same range of lag values. Since it peaks rather quickly, the spacing function is easily obscured by randomization procedures which allow lag to vary over a large range from trial to trial. The effects were further obscured by the fact that global measures of performance such as trials to criterion were typically used. In multi-trial experiments, the effects of spacing repetitions are only apparent if one observes sequences of correct and incorrect responses for individual items.

Although initial increases in performance over small values of spacing are reported in virtually all memory tasks, there are some differences in performance from task to task as spacing is increased to large values. In free recall, performance sometimes continues to increase for very long values of spacing (Madigan, 1969; Melton, 1967). Some nonmonotonic effects have been reported in paired-associates
learning (Breisford, Shiffrin and Atkinson, 1968; Young, 1971). This latter result appears to be related to the length of the retention interval employed (i.e., the interval between P2 and test). Glenberg (1976) was able to show that, with short retention intervals, the spacing function appears as an inverted U-shaped curve with declining retention around an optimal spacing interval of somewhere between four and eight items. With longer retention intervals, recall is a monotonically increasing function of interpresentation lag.

Plotted in Figure 3 are data from a replication of Melton's original experiment in free recall (Melton, 1967). There are several features of importance in the figure. First, recall for zero lag items is only slightly above that for single presentation items. Second, recall rises sharply as spacing increases from zero to small values of lag. Note that the largest increment on the entire curve is between lags zero and two and that the function is monotonic. These findings are consistent across many experiments in free recall and other paradigms. Consequently, most experimental and theoretical efforts have concentrated on explaining these figures of spacing phenomena.

Despite a considerable amount of research over the past decade or so, the spacing effect is still not well understood. The effect has sufficient generality to suggest that it represents some general feature of the human information processing system. Furthermore, the effect appears to be a convenient vehicle for examining the basis of stimulus repetition effects in general. A glance at Figure 3 suggests why this might be the case. Spacing effects provide a situation in which stimulus
Figure 3. Percent recall as a function of interpresentation lag in free recall. Data are from a replication of Melton’s original experiment (Melton, 1967) reported by Madigan (1963).
repetition ranges from being highly ineffective to being highly effective. Understanding why spacing produces this range of performance differences is likely to provide some clues as to the processes which underlie stimulus repetition effects.

It seems clear that the spacing effect is a well-established empirical phenomenon in a variety of experimental contexts. We turn, then, to the theoretical notions that have been proposed as explanations for the effect.

THEORIES OF THE SPACING EFFECT

The fact that spacing effects have been obtained in a large variety of experimental tasks has greatly influenced the theoretical work on the problem. The predominant assumption has been that the effects of spacing repetitions in each of these paradigms has a common, underlying cause. Theories of spacing phenomena have therefore tended to be stated at a level which can encompass what is known about performance in each of the various tasks. This is somewhat out of step with modern trends in theory development which have stressed detailed and precise theorizing within a narrow range of tasks. However, the task similarities are compelling and the assumption of a unique cause is certainly reasonable as an initial working hypothesis.

There are essentially two classes of explanations that have influenced the majority of experimental work on spacing phenomena. The first is based on expansions of the encoding variability concept originated by Martin (1968). The second is a rather wide range of
theories variously called "inattention" theory (Underwood, 1969), voluntary attention theory (Hintzman, 1974) or deficient processing theory (Hintzman, 1976). These two classes of theory and the available evidence relating to each are discussed in this section.

**Encoding variability theory**

Although the encoding variability concept was implicit in earlier writing, specifically that of Estes (1955), Underwood (1963) and Lawrence (1963), it was developed in a more useful form and applied to human learning and memory by Martin (1968). The basic notion is, in fact, a simple extension of the idea of encoding itself. Once it is accepted that an encoding process intervenes between the presentation of a stimulus and the working form of that stimulus in memory, it is a simple step to formulate the idea that the same nominal stimulus can be stored in different forms depending on variations in the encoding process. Thus, the relationship between a physical stimulus and its memory code is conceived to be of the one-to-many variety. The problem for the experimental psychologist is to determine the characteristics of functional stimuli and to discover the variables which govern the course of the encoding process.

The thrust of Martin's original article was to explain certain of the effects of stimulus meaningfulness on paired-associates learning and transfer. His important insight was to show how associative learning preceded through a mediational stage in which associative bonds were formed between **encoded** forms of stimuli and responses rather than
between the nominal stimuli and responses that were presented by the experimenter. The encoding process was described as a probability distribution over the possible encoded forms of each nominal stimulus. The probability distribution was considered to have greater variability for low meaningfulness stimuli than for high meaningfulness stimuli. According to the theory, subjects must first learn to encode the stimulus in a consistent way and then learn specific stimulus-response pairings. When the nominal stimulus is encoded in one way on the first trial, and in different ways on subsequent trials, it is difficult for hypothetical associative bonds between the encoded forms of stimuli and responses to gain strength and become associated. The difficulty of learning low meaningfulness stimulus-response items is thus a direct result of high variability in the encoding process for such items. Martin developed some interesting experimental tests of his hypothesis. A review can be found in Martin (1972).

It is possible, however, to derive different predictions from the application of encoding variability. Consider a repetition experiment in which associative strength is built up between all of the encoded representations of a stimulus and the appropriate response. In such a case one might assume that there are multiple retrieval routes to the appropriate response established in memory. One would expect that performance based on such memory structures would be high in comparison to situations in which only a single retrieval route is established.

With these basic ideas at hand, the spacing effect can be accounted for by assuming that the encoding process is influenced by
contextual changes which spontaneously take place over the course of stimulus presentations. Thus, stimuli which occur at very short P1-P2 intervals are likely to be encoded similarly on each presentation. Stimuli are, on the other hand, likely to be encoded in increasingly different forms as the P1-P2 interval increases. The increment in memory performance due to the spacing of repetitions is therefore a simple consequence of the contextual changes necessarily involved in the sequential presentation of TBR stimuli.

There are several testable predictions regarding this characterization of spacing phenomena. For example, consider a situation in which the encoding distribution for a certain nominal stimulus is degenerate around a single encoded form in memory. Thus, there is no variability inherent in the encoding process and no basis to predict an effect of spacing. This leads to an important prediction of encoding variability theory—that the spacing effect should be dependent on the nature of the stimulus material being employed. For highly fractionable and complex stimuli, there should be large and robust spacing effects due to potential multiple encodings. On the other hand, well-learned and integrated stimuli should generate small spacing effects. In the extreme case, we should expect the spacing of repetitions to have no effects on recall whatever. In actual fact, however, spacing effects appear to be relatively independent of stimulus materials. Belleza, Winkler and Andrasik (1975) and Tzeng (1973) could obtain no differential effects of stimulus meaningfulness on the spacing function. Given Martin's evidence for the correlation between the variability of the
encoding process and stimulus meaningfulness, these results are counter
to an encoding variability theory of the spacing effect.

It is difficult to determine which of the assumptions failed. It
is possible to argue that variability is inherent in the encoding process
for all stimuli. To the extent this is true, the failure to obtain
interactions between spacing and optimally chosen stimulus material is
less crucial for the encoding variability explanation. It seems reason-
able that a better test of the theory might involve attempts to gain
control over the encoding process itself. This could be done by manipu-
lating the context in which the stimulus repetitions occur.

One line of such research has taken advantage of the existence of
homographic or polysemous words in natural language (words which have
several distinct meanings such as contract bridge or railroad bridge).
There is substantial evidence that the encoding of different meanings
of such words can be biased by introducing the words into a meaningful
context which is consistent with only one meaning of a word (Light and
Carter-Sobel, 1970). This approach seems to be consistent with Martin's
original definition of encoding variability.

The majority of such research has been carried out in the free
recall paradigm. Predictions from encoding variability are not as
clear-cut as in paired-associates learning because it is difficult to
identify the nature of the retrieval cues or stimuli which generate
performance. Nevertheless, if the manipulation of semantic contexts
establishes multiple codes in memory, it seems reasonable to assume that
the nominal stimulus can be recalled in multiple ways (that is, by
recalling any of the several meanings of a polysemous word). The spacing effect, then, could be explained through a process which allows the probability of encoding multiple meanings to increase with spacing. By presenting TBR words along with cues which bias different meanings of a word on each presentation, it should be possible to "force" the establishment of multiple codes in memory regardless of the spacing between presentations.

Figure 4 is a reproduction of data obtained in an experiment by McCormack and Carboni (1973). The data are somewhat contradictory in this area, but the pattern of recall has been replicated in other experiments (Madigan, 1969, exp. 2; Thios, 1972). McCormack and Carboni presented subjects with homographic nouns embedded in sentences. The sentences were constructed so that either the same meaning or a different meaning was biased across repetitions. Although the authors interpreted their results as being in favor of encoding variability theory, there are some serious questions about an encoding variability interpretation of the data in Figure 4. If encoding variability theory of the sort operationalized by McCormack and Carboni is the source of the spacing effect, and encodings are forced to be the same on each presentation, there would seem to be no reason to expect a spacing effect. However, a substantial spacing effect is apparent in the "same" curve of Figure 4. Secondly, the recall levels at lags five and thirteen pose a considerable problem for basic assumptions about encoding variability and recall. Multiple codes in memory are supposed to enhance performance. However, the means for the "different" condition are considerably below those for
Figure 4. Percent recall for homographic nouns as a function of biasing context (same or different meaning) and interpresentation lag. Data are from McCormack and Carboni (1973).
the "same" condition except at lag zero.

A slightly different version of the encoding variability hypothesis holds more promise for the understanding of spacing effects in free recall. Consider, for a moment, that free recall performance is based on the encoding of a network of interitem associations. The basic question concerns what kind of associative structures would lead to enhanced recall. If an item could be accessed or retrieved from a number of different points in the network, its likelihood of retrieval from memory would be enhanced considerably. Because spaced repetitions occur in a different list context on each repetition, the number of retrieval routes, or associations with other list items, is likely to be greater than that for massed repetitions. In this sense, encoding variability is interpreted in terms of multiple retrieval routes to the same functional code in memory, rather than in terms of multiple codes or representations of the same nominal stimulus.

Glenberg (1976, 1977) has provided some convincing evidence favoring this last version of encoding variability. Originally, he sought to use the encoding variability concept to explain some of the non-monotonic effects of spacing reported in paired-associate experiments. He was able to show that the nonmonotonicity is lawfully related to the length of the retention interval employed. Nonmonotonicity is present only at short retention intervals. The spacing function becomes increasingly monotonic as the retention interval is lengthened. These results have been obtained in free recall (Glenberg, 1977) and in paired-associates learning and recognition memory (Glenberg, 1976, exps.
I and III respectively).

The exposition of encoding variability theory developed previously seems adequately to account for the Glenberg results at the longer retention intervals (the monotonic effects) and for the initial portions of the spacing curves at the shorter retention intervals (the increasing portion of the nonmonotonic effect)--spaced repetitions establish multiple retrieval routes in memory, giving spaced items an advantage over massed items. What is lacking is a reasonable explanation of the declining portion of the spacing function at the shorter retention intervals. In this regard, the notion of contextual change and its influence on the encoding process seem to provide a natural account. Consider the situation in which P1 and P2 are widely spaced in time, but test occurs shortly after P2. By the assumptions of the theory, the two presentations should provide a different set of potential retrieval cues. However, recall will be primarily a function of the proximity of P2 and test in the list--P1 contributes only a small amount to recall relative to P2. As the P1-P2 interval becomes smaller, the encoding at P1 becomes increasingly likely to contribute to recall until it reaches some optimal value. In other words, the optimal combined contributions of P1 and P2 occur at some intermediate value between very short and very long P1-P2 intervals. Thus, a nonmonotonic function is obtained.

Glenberg has applied this same basic explanation to account for the nonmonotonic effects he obtained in several different memory tasks. Such cross paradigm verification is rare in human memory research and
certainly constitutes strong evidence for the theory.

In summary, encoding variability theory has generated both positive and negative evidence. Glenberg's theory nearly explains the retention interval by spacing interactions and no other theory seems capable of handling these data. On the other hand, it is not certain how to reconcile these findings with the negative evidence discussed earlier. The predicted interactions with stimulus material have not been established. Some very natural interpretations of encoding variability theory in terms of the various meanings associated with a word in memory do not accord well with the encoding variability theory of spacing effects. It should be clear that we do not yet fully understand the effects of variable encoding on memory performance. However, encoding variability theory has strong intuitive appeal, and it continues to stimulate research. It therefore remains as a strong potential explanation for spacing phenomena.

**Deficient processing theories**

The second major class of theories dealing with spacing phenomena has been called "deficient processing" theory by Hintzman (1976). The idea common to these theories is very simple. The cause of the spacing effect is the tendency for differential processing at P2 for items occurring at various P1-P2 intervals. Specifically, it is assumed that the second presentation of items occurring at small P1-P2 intervals is not processed as "adequately" or as "thoroughly" as it is for items occurring at large P1-P2 intervals. While there is some empirical
support for this class of theories, a problem lies in specifying a mechanism mediating the differential processing. This point will become clearer in the following discussion.

There is a good evidence that some type of deficient processing at P2 occurs for small values of spacing. For example, consider that the final recall of a repeated item could be based on the recall of information from P1 and/or P2. If it is assumed that the events resulting from each presentation of an item are independent, performance on double presentation items should be consistent with the following equation from probability theory:

\[ P(\text{correct recall}) = P(P1) + P(P2) - P(P1)P(P2) \]

where \( P(P1) \) and \( P(P2) \) represent the probability of recalling information stored on P1 and P2 respectively. Estimating \( P(P1) \) and \( P(P2) \) from the recall percentages for single presentation items, Glanzer (1969) has determined that items occurring at small values of spacing fall consistently below this independence baseline, while items occurring at large values of spacing approach or surpass the baseline.

The major division among deficient processing theories is based on the issue of subject control over processes responsible for the spacing effect. Following Hintzman (1974, 1976), we will consider this to be the issue of voluntary or automatic processes in the spacing effect.

Theories based on automatic or involuntary deficits. Perhaps the most remarkable fact about the spacing effect is its considerable
generality over wide differences in experimental procedure. This
generality has led to speculation that the spacing effect reflects the
operation of some involuntary cognitive mechanism. The additional fact
that spacing effects have been well established in incidental learning
paradigms (Rowe and Rose, 1976; Rose and Rose, 1977; Shaughnessy, 1976)
has also been taken as support for automatic or involuntary deficits.
Several theories have been based on the assumption that two presentations
of the same item occurring close in time are mutually inhibitory in some
fashion.

One such theory, proposed by Peterson (1966) and later by Landauer
(1969), is based on the classical notion of memory trace consolidation
over time. Research on memory consolidation has been generally confined
to clinical and animal research areas (McGaugh and Hertz, 1972). The
attempts to operationalize a consolidation process experimentally have
been somewhat disappointing (Miller and Springer, 1973). Nevertheless,
the concept has some explanatory value in the context of repetition
effects. For our purposes, consolidation refers to a process which
(1) operates autonomously in the interval immediately following stimulus
presentation, (2) does not require the presence of a physical stimulus,
and (3) which transforms a transient memory trace into a permanent,
durable memory. To apply the concept to memory experiments, long term
memory is considered to be proportional to the total amount of consoli-
dation which has taken place. A stimulus repetition which occurs close
in time to the original stimulus presentation is thought to interrupt
any ongoing consolidation processes left over from the first presenta-
tion, and recruit consolidation activity for itself. Thus, when repetitions are massed, the total amount of consolidation is less than when items are repeated after the consolidation process from the initial presentation has been allowed to run its course.

This particular version of consolidation theory can be rejected because it predicts that it is the first presentation of a repeated item which suffers from massed or closely spaced presentations. There is some evidence which suggests that it is the item repetitions which are deficiently or inadequately processed.

The most conclusive evidence comes from a study by Hintzman, Block and Summers (1973). Subjects were presented with a list of repeated items in mixed auditory and visual modalities. Because subjects are capable of remembering the modality of the first or second presentation with some accuracy, the modality information served, in effect, as tags which provided Hintzman et al. with a basis for determining which presentation was the primary determinant of memory performance. Three theoretical models were fit to the data. The best fitting model was one in which the source of the deficient processing was located at the second presentation of a massed or closely-spaced item.

However, consolidation theory can be revised to account for these results. It seems just as plausible to assume that the ongoing consolidation of P1 blocks or prevents P2 from entering into the consolidation process. This version of consolidation theory is compatible with the Hintzman et al. data. However, this should not be considered a strength of consolidation theory. Because consolidation is a poorly operational-
ized concept, it lacks specification of testable hypotheses and is difficult to refute.

The habituation-recovery theory proposed by Hintzman (1974) shares some characteristics with consolidation theory. The habituation mechanism is based on the assumption that stimulus presentation causes some internal representation of a TBR item to become adapted or habituated. During the period of adaptation, the internal representation is insensitive to activation by subsequent repetitions. As time passes, the internal representation recovers from adaptation and again becomes sensitive to repetition. The recovery process is assumed to follow the same time course as the spacing effect.

Because the effectiveness of stimulus repetition obviously depends on the state of recovery from adaptation, items repeated after long P1-P2 intervals are more likely to occur after the internal representation has fully recovered from adaptation. Thus, spaced items are more likely to be recalled than are massed items. This theory has the advantage of predicting unambiguously that it is the item repetition (P2) which suffers from massed presentations.

Hintzman, Summers and Block (1975) attempted to test the habituation-recovery hypothesis by influencing the time course of recovery through overhabituation, i.e., increased stimulus duration. If adaptation accumulates during stimulus presentation, greater amounts of adaptation should be produced with greater stimulus durations. In the series of experiments reported by Hintzman, Summers and Block, however, stimulus duration had no detectable effect on the spacing function for frequency
judgments. Whether this is considered to be evidence against the habituation-recovery hypothesis, of course, depends on what specific characteristics one assumes to be essential to the process of adaptation and recovery. For example, if adaptation is considered to be complete within a brief period of time following stimulus onset (say, within 2.2 seconds, which was the shortest duration studied by Hintzman et al.), there would be no reason to expect the spacing function to change with stimulus duration.

Both the consolidation and the habituation-recovery theories are difficult to refute because there are few ties between experimentally controllable variables and the internal constructs of the theories. The most convincing evidence in their favor is the well documented generality of spacing phenomena. The insensitivity of spacing phenomena to experimental manipulation indicates that the underlying mechanisms may be automatic or involuntary. However, the theoretical work on these ideas will remain speculative until the effects of spacing can be manipulated with theoretically relevant variables.

Voluntary processing strategies. It should be clear that the generality of spacing phenomena poses a considerable challenge for theory development. Different tasks undoubtedly involve different underlying processes--in fact, that is the purpose of studying many different memory tasks. An adequate theory of spacing effects therefore must be compatible with a large variety of process models.

Assuming that all spacing effects have the same general cause,
there are essentially two ways of making a theory of spacing effects compatible with a large range of memory tasks. The first is to locate the source of the effect in a stage which is common to all such tasks. For example, encoding variability theory, consolidation theory and the habituation-recovery hypothesis suggest that there are certain characteristics which are common to the encoding stage of all memory tasks, and that these characteristics can account for the spacing effects found in each task. However, there is a second possibility which appears to hold some promise. This is to attribute the effect to a theoretical level above the mechanisms which perform the basic information processing operations on incoming stimuli—in other words, to a control structure which organizes and guides the operation of encoding, storage and retrieval mechanisms.

Recent developments in cognitive theory have been possible, in part, because of a willingness to speculate about an executive which controls the flow of information within the human information processing system. To be sure, there are certain philosophical problems associated with this step (cf. Skinner, 1953). However, the problems are less serious if the rules and strategies by which the hypothetical executive operates can be specified.

A number of researchers have suggested that spacing effects could arise from the voluntary allocation of attention or processing activity to items at different values of spacing (Bjork, 1970; Greeno, 1970a; Hintzman, 1974; Underwood, 1970; Waugh, 1970). The problem is in specifying a set of rules that would lead to such allocation strategies.
This section reviews the available literature on this question and then attempts to develop these ideas in some detail.

The possibility that subjects may simply fail to attend to massed repetitions was suggested by Underwood (1970), and has become known as the inattention hypothesis (Crowder, 1976, chapter 9). In his original paper, Underwood reported a series of five experiments showing a remarkable consistency in the spacing effect for free recall over a variety of conditions. Because the effect did not seem to be related to traditionally potent variables in verbal learning (list length, presentation rate, etc.), Underwood worried that the effect might be psychologically trivial—that subjects simply may become bored with the task and fail to attend to repeated items. Although some empirical support in favor of an inattention hypothesis could be cited, it will not be discussed in this context because the hypothesis has several undesirable shortcomings. The most serious is that it lacks theoretical precision. A simple inattention to massed repetitions would explain most empirical results, but it does not lead to many interesting questions. In order to be fully adequate, the inattention hypothesis must be supplemented with some theoretical mechanism to explain why subjects do not attend to massed repetitions but continue to process spaced repetitions.

A more interesting explanation is the selective-coding hypothesis outlined by Greeno (1967) to account for a variety of effects in paired-associates learning. The selective-coding hypothesis characterizes learning as a trial and error process with selection from a set or
pool of memory codes or mnemonic devices. According to this position, a paired-associates study trial permits the selection of a coded representation for a stimulus-response pair. Such codes may or may not be adequate to support long term retention. However, the adequacy of the code is unknown until the paired-associate item is repeated on a subsequent trial. Greeno assumed that subjects adopted a rule based on memory retrieval for the replacement of codes. If a coded representation of an item cannot be retrieved from memory, it is replaced with a new code from the pool of potential codes. Because virtually any code will support retention for a short time, repetitions at small P1-P2 intervals are not sufficiently long enough to identify inadequate codes. Some number of such codes will therefore be retained at short P1-P2 intervals when they otherwise would be replaced. This gives an advantage to spaced repetitions. If a code can be retrieved after a long P1-P2 interval, it is likely to support long term retention. Codes which cannot be retrieved are replaced, thus adding to the probability that an adequate code will be found.

The major thrust of Greeno's 1967 paper was the introduction of an elegant Markov model used to predict a number of empirical phenomena in paired-associated learning. However, with regard to spacing effects the model has several difficulties. First, in its fully developed form, the model does not produce unique parameter estimates. Greeno dealt with this problem by making a priori assumptions about the values of certain parameters, thus reducing the number of free parameters required by the model. Secondly, and the most important from our point
of view, the model will not predict the nonmonotonic effects which have been well established in paired-associates learning (cf. Glenberg, 1976). Young (1971) has developed a five state Markov model which will predict nonmonotonic effects and which has a coding interpretation similar to that discussed above. However, the model has a large number of free parameters (seven) which make it difficult to evaluate.

It is interesting to note that Greeno's selective-coding hypothesis can be restated as an encoding or processing strategy in which encoding effort is applied only to repetitions which have been forgotten from earlier presentations. Although the mathematical model developed by Greeno is specific to paired-associates learning, the notion of a trial and error coding process based on forgetting in the spacing interval can easily be extended to other tasks such as free recall, recognition memory and memory judgments.

However, Hintzman (1974; 1976) has argued against this idea. Several studies (Hintzman and Rogers, 1973; Hintzman, Summers and Block, 1975; Hintzman, Summers, Eki and Moore, 1975) have obtained spacing effects in frequency judgment tasks for a variety of pictorial materials that are likely to have recognition probabilities close to 1.0 even for long spacing intervals. Thus, spacing effects can be obtained in frequency-judgment tasks with stimulus materials that are not likely to be forgotten during the spacing interval. Nevertheless, it is reasonable to suppose that items having a high probability of recognition may still vary in terms of memory accessibility. If this is so, then the strategy of allocating encoding effort to the least accessible items
may still be a viable explanation for the spacing effect.

A different but related theory of spacing effects focuses on the role of a short-term or primary memory system. The basic idea was first suggested by Brelsford, Shiffrin and Atkinson (1968) but was later proposed by Glanzer (1969, 1972). Because the earlier work relied on a comprehensive theory of memory developed by Atkinson and Shiffrin (1968) which will be useful as an analytical framework, the account of Brelsford et al. will be emphasized in this section.

The Atkinson and Shiffrin model is based on the theoretical separation of memory into several distinct memory stores. The short term store (STS) is a limited capacity processor which serves as temporary storage while information is transferred to long term memory. According to this model, information can be transferred to permanent memory only during the period it resides in STS. Long term memory (LTS) is assumed to be a static repository of permanent memories. STS and LTS along with a sensory memory store, comprise the model's major structural components.

Brelsford et al. accounted for the spacing effect with two assumptions: (1) when a new list item is presented, it replaces one of the items already residing in STS with some nonzero probability; (2) if a list item already resides in the STS, a second copy cannot be entered into the system. Thus, at short spacing intervals, there is a high probability that P1 would be in STS when P2 occurs, thereby preventing P2 from being processed. From various assumptions about STS, it can be shown that the probability of this event occurring decreases
with increased spacing. On the average, therefore, P2 will receive less processing with short as opposed to long spacing intervals.

An important result of the Atkinson and Shiffrin model was the development of a hypothetical set of control processes which may be presumed to regulate the operation of and interplay between various memory structures (Atkinson and Shiffrin, 1972). A number of such control processes have subsequently been described. Among them are rehearsal, mnemonics, decision rules, coding strategies, rules for allocation of attentional capacity, and so on. These control processes give the model the flexibility to adapt to changing task demands. In order to apply the model to a given task situation, it is necessary to select the appropriate set of control processes which govern the operation of the structural aspects of the model. The selection of control processes is assumed to be influenced by instructions, independent variables, payoffs and other value-cost considerations that fall under the general rubric of demand characteristics. It is assumed that the subject has knowledge concerning his information processing capabilities and that he uses this knowledge in conjunction with information about the experimental task to formulate rules or strategies to guide his task behavior. From this point of view, the major problem for the experimental psychologist is in understanding how information processing capacities are deployed and coordinated in performing various tasks. Thus, all experimental tasks may be viewed as having a problem solving aspect, in that the subject must find a way of adapting control-process capabilities to the demands of the experimental task set before him.
Using this frame of reference, a theory of the spacing effect requires a specification of the rules or strategies that are employed by the subject under conditions of intraserial repetition. For our purposes, the assumption of a specific structural model of memory is unnecessary. Memory experiments are usually designed so that the average person cannot remember all of the TBR items which are presented for memorization. Thus, some method of allocating limited processing capacity to the various items must be devised. When intraserial repetitions are part of the task, an efficient method of study might involve attending only to those items which are not yet well learned—perhaps those which fall below some level of accessibility. Thus, when accessibility of information stored from P1 is high, as it would tend to be for small values of spacing, the subject elects to allocate processing capacity to other TBR items. For repetitions occurring at large values of spacing, when the accessibility of previously stored information is more likely to be low, processing capacity is allocated to the repeated item and this results in enhanced memory performance at test.

There is some reason to believe that the strategy of allocating processing capacity to items which are not well learned is common to a variety of tasks. In most experiments, the experimenter attempts to control the subject's distribution of study or encoding effort by carefully controlling the presentation of his stimulus material. While the subject undoubtedly adopts a processing strategy of some kind in these experiments, there is no means by which the strategy can be directly
observed. However, by allowing the subject to select his own presentation schedule, it may be possible to gain some insight into the control processes selected by the subject for a particular task.

Atkinson (1972) and Ciccone and Brelsford (1974) allowed subjects to self-select items for presentation in a paired-associates task. The most general conclusion that can be reached from these studies is that subjects are actually quite good at optimizing learning when given the freedom to use their own processing strategies. The Atkinson study is particularly interesting because he was able to compare learning performance in the self-selection condition with optimal presentation schedules derived from a simple mathematical model of the learning process. These optimal schedules identify and present those items which are not yet well learned. During the acquisition period, performance under optimal conditions is therefore quite low because the subjects are continually tested on items that are poorly learned. Atkinson's subjects in the self-selection condition displayed exactly this kind of performance; low recall scores during the acquisition period, followed by relatively high recall scores on tests of long term retention. This suggests that the subjects in the self-selection condition tried to adopt a strategy similar to that of the optimization schedules derived from the mathematical model.

The Atkinson study did not include information about lag (spacing) schedules adopted by subjects. This was remedied by Ciccone and Brelsford (1976) who also allowed subjects to self-select items for presentation in one condition of a paired-associate experiment. The
most interesting finding of their study is the probability distribution of lag values generated in the self-selection condition. These values were clustered around an optimum lag of four, with virtually no choices at lag zero or lags greater than seven. This suggests two things. First, an item's accessibility is an important determinant of the effectiveness of a repetition, and, secondly, that subjects possess an ability to evaluate the status of an item in memory. These conclusions fit nicely with a processing strategy interpretation of the spacing effect in which subjects are assumed to allocate processing capacity on the basis of an evaluation of an item's status in memory.

**Summary and comparisons of the theories.** Although a number of theoretical interpretations have been proposed, there are basically two classes of theory that have stimulated the majority of the research on spacing phenomena—encoding variability theory and the various deficient processing theories. It is interesting to note that each class of theory is concerned with somewhat different questions. The predominant theme of research on encoding variability concerns the kinds of structural changes that take place in the memory trace as a function of stimulus repetition. Encoding variability theory can be made to specify rather precisely the kinds of structural changes that may lead to enhanced retrieveability—the encoding of multiple codes or retrieval routes. Spacing effects can be predicted by tying the encoding process to contextual changes which take place as a natural consequence of spacing repetitions.
Deficient processing theories have tended to ask a different kind of question. Such theories have attempted to specify cognitive mechanisms which explain the depressed performance characteristic of spacing effects in many different tasks. Specific accounts of how a memory trace is structurally altered have not been addressed in detail by this class of theories. Ultimately, an adequate theory of spacing effects must incorporate some ideas as to how a memory trace changes with stimulus repetition. However, it may be unnecessary to specify completely a structural theory of the memory trace before approaching some of the questions that are raised by the various deficient processing theories.

Two kinds of cognitive mechanisms have been discussed in the review of deficient processing theories. The first identifies the spacing effect with an involuntary or automatic mechanism which inhibits or prevents repetitions at small values of spacing from being effective. Several such mechanisms have been proposed—consolidation theory and the habituation-recovery hypothesis. Neither has been developed in sufficient detail to permit conclusive tests to be conducted. A second kind of deficient processing theory is based on the notion that spacing effects may be the outcome of common processing strategies or rules under subject control. There is some evidence that subjects may allocate processing capacity to repetitions based on accessibility of the repeated information. It has been argued that strategies based on an accessibility rule would account qualitatively for the monotonic increase in performance with the spacing of repetitions.
Processing strategies or rules may be valid under a large range of assumptions about memory codes and structures. Hence, a theory based on processing strategies has the potential for explaining the effects of spacing repetitions within a large range of memory tasks.

This latter interpretation suggests a flexibility that contrasts with both the habituation-recovery hypothesis and consolidation theory. The concept of automatic or involuntary processing deficits suggests that the returns on additional processing or encoding effort at small values of spacing would be limited at best. By designing tasks which vary the processing demands on the subject it should be possible to determine whether additional processing effort allocated to repetitions at small values of spacing has corresponding effects on memory performance. Once this question is resolved, it may be possible to make some cautious inferences about how the memory trace is structurally altered with stimulus repetition to improve its retrieveability.

Before moving to a presentation of the experiments which were designed to answer some of these questions, it will be useful to present and discuss a selective review of the experimental literature which will serve to narrow the focus of the present research.

DEFICIENT PROCESSING THEORIES:

EXPERIMENTAL EVIDENCE

A processing strategy hypothesis such as developed in the previous section makes several clear-cut predictions regarding the outcome of certain experimental manipulations related to the spacing function. However, some of these predictions are also compatible with the idea
of automatic or involuntary processing deficits at small values of spacing. A discussion of these predictions and a review of the experimental work relating to each is included at this point.

**Covariance of processing activity and spacing.** The processing strategy hypothesis is based on the idea that subjects devote differential amounts of encoding effort to P2 at different values of spacing. The hypothesis therefore predicts that experiments designed to monitor the allocation of processing activity should show depressed processing activity at P2 for small values of spacing. Although failure to find consistent relationships between measures of processing activity and the spacing of repetitions would cause serious difficulty for the processing strategy hypothesis, the empirical confirmation of these relationships would not necessarily eliminate the possibility of automatic or involuntary processing deficits at small values of spacing. A correlational relationship between encoding effort and the spacing of repetitions could result from a strategy adopted by subjects to contend with inhibitory mechanisms which prevent effective encoding effort from being allocated to repetitions at small values of spacing.

There have been several studies which have attempted to monitor the encoding effort of a subject as he proceeds through a list of words with distributed repetitions. These studies provide substantial correlational evidence to indicate that subjects elect to process P2 less thoroughly at small spacing intervals. Rundus (1971) required subjects to rehearse overtly while studying a list of words. He then
analyzed the rehearsal protocols. Over four experiments, he found the number of rehearsals to be highly correlated with the recall of an item. Given that subjects are overtly rehearsing, the number of rehearsals may therefore be a reasonably good indicator of the amount of attention or processing time that is allotted to a given item. In one experiment, Rundus presented subjects with a list of distributed repetitions. Spacing was indexed by interpresentation lag. There was a high correspondence between the number of rehearsals and recall for a given lag condition, both increasing as lag increased. Furthermore, an analysis of the rehearsal protocols indicated that items with short spacing intervals tended to receive fewer rehearsals than items with long spacing intervals at two points, (1) during the P1-P2 interval and (2) in the interval following P2. The number of rehearsals occurring during the P1-P2 interval is limited by the time available for rehearsal within that interval, and therefore may not be reflective of the processing strategy adopted by the subject. However, the small number of rehearsals following P2 for items with short spacing intervals is supportive of the idea that the spacing effect is a result of differential processing allotted to short and long lag items.

Another means of monitoring processing activity developed out of an attempt to reconcile the spacing effect with the total time principle. This principle states that the amount of recall is some function of the total time allotted for study regardless of how that time is distributed (Cooper and Pantle, 1967). Early studies indicated that the total time principle held under conditions of intraserial
repetition (Waugh, 1963, 1967), but the confirmed existence of the spacing effect is a clear violation of this principle. However, if items at different values of spacing do not receive equal amounts of functional processing time, say, due to inattention at P2, the spacing effect may not be a fair test of the principle. The total time principle may still hold if one could separate functional processing time from observable study time in an experiment. In two separate experiments (Shaughnessy, Zimmerman and Underwood, 1972; Zimmerman, 1975), subjects were allowed to pace themselves through a list with distributed repetitions. The inter-response times (time between the onset of one item presentation and the next) were recorded. This technique did not seem to change the typical form of the spacing function in the free recall data for either study. The response time distributions revealed that items with short spacing intervals consistently received less study time at P2. Still, these distributions did not account for all of the variance due to the spacing effect. If the total time principle does hold under the conditions of these experiments, the ratio of the total number of items recalled for a given lag condition to the total amount of time allotted to the study of those items should remain constant across lag conditions. In both experiments, the number of items recalled from the massed presentation (lag zero) condition was less than the total time principle predicted.

Hintzman (1976) has reanalyzed Zimmerman's data on the assumption that the self-pacing strategy did not eliminate all of the "dead" time on each presentation. By incorporating an average dead time of
four seconds, Hintzman was able to derive an estimate of processing time which accurately predicted the recall data. Furthermore, the relationship between estimated processing time and recall was linear with the data for items presented twice and those presented three times falling on the same linear function. Before accepting the obvious conclusion from Hintzman's reanalysis, it should be noted that four seconds is a rather large amount of dead time relative to study time. Average study time for Zimmerman's experiment was 7.00 seconds per presentation.

A different method of monitoring processing activity was employed by Johnston and Uhl (1976). Subjects were required to perform a secondary reaction time task to an auditory stimulus while studying a list with distributed repetitions. It was proposed that subjects who processed less actively at P2 for short P1-P2 intervals should have more processing capacity available for the reaction time task. Thus, the data of interest are the reaction times when the stimulus occurred during P2. The reaction times were fastest for items with small spacing intervals and increased as the spacing interval increased. The Johnston and Uhl (1976) data clearly support the notion that there is less active processing at P2 for short spacing intervals.

A final study relating to the monitoring of processing activity is an experiment reported by Hintzman, Summers, Eki and Moore (1975, exp. III). These experimenters presented subjects with a series of complex pictures. Some were repeated at various interpresentation lags, and the number of eye fixations on each presentation was recorded.
Despite the fact that there was a substantial spacing effect in judged frequency, there was no difference in the number of eye fixations at P2 for any of the lag conditions. If the number of eye fixations can be taken as a measure of processing activity, the results seem to be in conflict with several of the studies discussed above.

It is difficult to determine the differences in procedure which may have caused the different outcomes. The Rundus (1971), Shaughnessy et al. (1972), Zimmerman (1975) and Johnston and Uhl (1976) studies all employed linguistic items as stimuli whereas Hintzman et al. (1975) employed complex pictures that could not be easily recoded into verbal or acoustic units. In addition, Hintzman et al. used a judged frequency measure, while the other studies all used free recall measures. More needs to be known about the kinds of control processes a subject employs to deal with various types of stimulus material and also what types of information are provided by various dependent measures. If there is a direct relationship between processing activity and number of eye fixations, the Hintzman et al. study appear to favor the automatic deficit theories. On the face of it, eye fixations appear to be a reasonable measure of processing activity (Loftus, 1972), but nevertheless this may be a faulty assumption. The bulk of the evidence clearly favors the prediction that encoding or processing effort is allocated differentially to repetitions at different values of spacing.

Performance tradeoffs between repeated and adjacent list items. We now turn to consider a second prediction of the processing strategy
hypothesis which concerns the possible reallocation of processing time during item repetitions at small values of spacing. There is some evidence that subjects do some self-selection of study items even in cases where the experimenter rigidly controls the presentation schedule of the TBR items. Studies in which the subject is asked to rehearse overtly indicate that many different items are rehearsed during the presentation of a single item (Rundus, 1971). If the processing strategy hypothesis is correct, it seems reasonable to expect that some reallocation of processing capacity from massed repetitions to surrounding list items might occur. Reallocation of this sort could be detected by examining recall for surrounding list items.

This prediction was investigated by Potts (1972), who examined recall for items surrounding the second presentation of repeated items at various lag values in paired-associates learning. His experiments revealed no such increments. Elmes, Greener and Wilkinson (1972) explored the same possibility in free recall. They found significantly higher recall for items following massed presentation than for items following spaced presentation in two experiments. This suggests that during P2 of a massed item, subjects may be preparing to encode the following item.

An experiment by Waugh (1970, exp. II) supports the finding of Elmes, Greener and Wilkinson. In this study, subjects were presented with lists comprised of single presentation items and repeated items at either massed or spaced presentations. Spacing was thus manipulated between lists rather than within lists as in the usual design. Her
data indicate that the recall of single presentation items is strongly dependent on whether the repeated items are presented with massed or spaced presentation. This would occur if subjects tended to reallocate processing time to single presentation items in the massed presentation lists.

Hintzman and Stern (1977) replicated the essential details of the Elmes, Greener and Wilkinson studies in three experiments, one employing frequency judgments and two in free recall. There was no statistical evidence in any of the experiments that performance for items surrounding repeated items differed as a function of lag. The conditions under which tradeoffs between repeated and surrounding list items occur or do not occur have clearly not been established.

**Sensitivity of the spacing effect to instructional variables.** A third prediction of the processing strategy hypothesis is related to the fact that, according to the theory, the spacing effect is the result of a subject-controlled process. The form of the effect should therefore be sensitive to manipulations of instructions and other variables such as rehearsals or payoffs. This prediction is particularly crucial to distinguish between the processing strategy hypothesis and the two automatic deficit theories, memory consolidation and habituation-recovery. The latter two theories predict that additional processing effort allocated to repetitions at small P1-P2 intervals should be limited at best. By designing tasks which vary the processing demands on the subject it should be possible to test these assumptions.
Flexible allocation of processing resources resulting in enhanced memory performance would tend to disconform both consolidation theory and the habituation hypothesis. On the other hand, if subjects cannot allocate effective encoding effort to repetitions at small P1-P2 intervals doubt would be cast on the processing strategy hypothesis. Unfortunately, attempts to manipulate the processing strategy of the subject with instructional variables have not produced a coherent pattern of results.

Several attempts to manipulate the amount of processing at P2 for short P1-P2 intervals were reported by Hintzman, Summers, Eki and Moore (1975, exp. I and II). In these experiments, the dependent measure was judged frequency. In the first experiment, subjects were presented with a series of complex pictures in which some pictures were repeated at various interpresentation lags. A tone, paired with a subset of critical pictures, indicated that there would be a high payoff for correct responding to these pictures. The data of interest are the frequency judgments when the tone accompanied P2 for a critical picture. A substantial spacing effect was observed for both the high payoff items and for control items which were not paired with the tone. Although the spacing function for the high payoff items was displaced higher on the ordinate than that for control items, the form of the function for the two kinds of items did not differ statistically (i.e., there was no interaction between spacing and payoff condition).
Although these results appear to be nonsupportive of a processing strategy interpretation, the absence of a significant interaction actually renders the results of their study somewhat ambiguous. The inferential problem for the researcher attempting to understand these results lies in the consequences of accepting the null hypothesis for interaction. Because of the additivity implied by the absence of a significant interaction, it seems appropriate to conclude that the extra processing effort induced by the payoff manipulation was equally effective in all spacing conditions—thus suggesting that the underlying mechanism is not automatic or involuntary in nature. Predictions based on the processing strategy hypothesis, however, depend on how one assumes that the payoff manipulation is affecting the distribution of encoding effort. If the payoff manipulation achieved a redistribution of encoding effort towards equal processing of repetitions at all values of spacing, the experiment should have produced a flattened spacing function in the high payoff condition. If the payoff manipulation added a constant amount of encoding effort to repetitions at all values of spacing, there would be little reason to expect a substantial change in the spacing function. This, of course, was the finding of Hintzman et al.

A second experiment reported by Hintzman et al. required subjects to rehearse a constant number of times during the presentation of each item in a list of words with distributed repetitions. Whatever else it accomplishes, the experimental technique of having subjects rehearse overtly insures that attention will be paid to an item during the
rehearsal period. This makes it difficult for the subject to reallot
processing activity at P2 for short spacing intervals and probably
disrupts the normal rehearsal activity of the subject. The results of
this second study indicated that a typical spacing effect was obtained
for judged frequency with this manipulation. As in the payoff study,
the form of the spacing function was not statistically different from
that of a control condition in which no restriction was placed on the
items which could be rehearsed.

This same manipulation has been used in several other experiments
with different dependent measures. Using a Shepard and Teightsoonian
(1961) recognition task, Ciccone and Brelsford (1974) controlled the
number of times a subject was allowed to rehearse overtly during each
presentation of an item. Contrary to the Hintzman et al. result, this
manipulation produced a flattened spacing function which was statisti-
cally different from a condition in which rehearsal was free to vary.
Glenberg (1977, exp. II) and Wright and Brelsford (in press) had
subjects rehearse in groups of items or allowed them to rehearse only
the item currently being presented. Free recall showed a spacing
effect in both conditions but it was much reduced from the condition
in which their rehearsal processes were restricted to only the current
item.

It is not clear what effect instructional variables have on the
spacing function. Instructions on how to rehearse seem to have some
effect on the spacing function for recall and recognition but not for
frequency judgments. This leads one to question the prevailing view
that the same process underlies the effects of spacing repetitions in all tasks.

**Effects of spacing in the incidental learning paradigm.** The fourth prediction concerning the processing strategy hypothesis arises from a consideration of the conditions under which subjects could be expected to use rules or strategies in dealing with a particular task situation. Incidental learning tasks provide a convenient situation in which the use of strategies and rules can be carefully controlled through instructions. In this paradigm, subjects are not informed as to the true nature of the memory task, but instead are asked to perform an "orienting" task on the TBR items, usually a rating task of some kind. By careful selection of orienting task, the experimenter gains a certain amount of control over the subject's processing of the stimulus material.

Hintzman (1976) has suggested that strategies and rules of the kind discussed in the previous section are not employed in incidental learning tasks. The effects of distributing repetitions in incidental tasks would therefore provide a powerful means of distinguishing between the automatic deficit theories and the processing strategy hypothesis.

There are several experiments which indicate that the spacing effect occurs for some orienting tasks in the incidental paradigm (Shaughnessy, 1976; Rowe and Rose, 1976; Rose and Rowe, 1977). However, the conclusions that may be drawn from these studies depend on the
theoretical basis for memory performance established by the orienting task. For most of the tasks employed in incidental learning paradigms, the subject is asked to make a rating of item characteristics such as meaningfulness, imagery value, pleasantness, word frequency, etc. In order to perform these tasks with consistency, it is probably necessary for the subject to have information about his previous ratings. The subject therefore must continually be storing, retrieving and comparing list information in a way that will make it accessible when he needs it as he proceeds through the list. When the subject is suddenly presented with a request for recall at the conclusion of a list, he has a substantial body of organized information that can be used in his retrieval attempts. In short, the processes employed in performing the orienting tasks may bear considerable similarity to those employed under intentional learning instructions. This basic similarity is supported by the finding that, under certain conditions, incidental task performance may approach that of intentional learning (Hyde and Jenkins, 1969, 1973).

Once the orienting tasks are understood in these terms, it is not difficult to think of ways to accommodate the effects of spacing repetitions in such tasks. If the rating task is the same on each repetition of an item, the easiest course of action for the subject is simply to remember his previous rating. With increased spacing, however, the subject is less likely to be able to remember it. He is thus more likely to repeat the process of comparison with other list items in order to decide on a rating. He will, therefore, encode more
information about the spaced item, thereby making it easier to retrieve when he is asked to recall it. This interpretation is admittedly post hoc, but more precise ideas about the processes involved in incidental task performance are needed before reaching conclusions about the spacing effect.

**Summary of predictions and experimental evidence.** A number of theoretical predictions and many experiments have been discussed and it will be useful to summarize briefly the material contained in the section. At present, the experimental data provide somewhat mixed support for the involvement of possible processing strategies in the spacing effect. For the first prediction, that measures of encoding or processing effort should covary with the spacing of repetitions, the evidence is supportive. The several experiments which have collected measures of encoding effort suggest that different amounts of effort are expanded for repetitions at different values of spacing (Johnston and Uh1, 1976; Rundus, 1971; Shaughnessy, Zimmerman and Underwood, 1972; Zimmerman, 1975). The one exception is the eye movement experiment of Hintzman, Summers, Eki and Moore (1975). Until more is understood about attention, and the relation between attention and encoding effort, it is probably fruitless to try to reconcile this experiment with the others.

The evidence concerning the second prediction, that of tradeoffs in memory performance between repetitions at different values of spacing and nearby list items, is perhaps more ambiguous. In free recall, three experiments support the existence of such tradeoffs (Elmes,
Greener and Wilkinson, 1973, exp. I and II; Waugh, 1970). Hintzman and Stern (1977), however, were unable to replicate these findings in a frequency judgment task and in two studies in free recall. The conditions under which tradeoffs between repeated and surrounding list items need to be established.

The third prediction was that the manipulation of processing or encoding effort should demonstrate flexible allocation of effective processing activity regardless of the spacing of repetitions. The experimental evidence is again conflicting. The manipulation of rehearsal patterns through instructions appears to have rather large effect on the spacing function in recognition (Ciccone and Brelsford, 1974) and recall (Glenberg, 1977; Wright and Brelsford, in press), but not in a frequency judgment task (Hintzman, Summers, Eki and Moore, 1975). This evidence suggests that the frequency judgment task may be somewhat of a special case since the disconfirming evidence was all obtained with frequency judgments as the dependent measure.

The fourth prediction concerns the effects of spacing repetitions in an incidental learning paradigm. There is no empirical disagreement on this question. Spacing effects can be obtained in incidental learning paradigms (Rowe and Rose, 1976; Rose and Rowe, 1977). However, the meaning of this finding is somewhat ambiguous. Hintzman (1976) has suggested that these findings may eliminate all but the theories based on automatic or involuntary mechanisms. However, this seems to be a premature conclusion until the basis of incidental learning task performance is better understood.
This section has reviewed the experimental evidence relating to deficient processing theories. It should be clear that no single theory has received unequivocal support and there are several unanswered questions. The payoff study reported by Hintzman, Summers, Eki and Moore (1975) is particularly important for evaluating the theoretical approaches and predictions developed in the previous two sections. Any theory which incorporates the assumption that two stimulus presentations occurring close in time are mutually inhibitory in some way, such as the habituation-recovery hypothesis or consolidation theory, should predict that the gains due to additional encoding effort allocated to repetitions at small values of spacing should be marginal at best. On the other hand, the processing strategy hypothesis, or other theories based on control process notions, do not place such restrictions on the effectiveness of encoding effort. The simplest method of manipulating encoding effort is by providing some incentive for remembering certain items. However, valid inferences from such a study require some knowledge of how a subject adjusts his allocation strategy in the light of incentives. Without this information, it is difficult to interpret the failure of Hintzman et al. to obtain differences in the spacing function with differential incentives.

The present research consists of three experiments in the free recall paradigm. The first two experiments were designed to provide information on how subjects adjust to the introduction of differential incentives in simple and straightforward ways. The third experiment is an attempt to extend the resulting interpretation to the incidental
learning paradigm.

EXPERIMENT I

The self-pacing technique employed by Shaughnessy, Zimmerman and Underwood (1972) and Zimmerman (1975) provides a straightforward way of acquiring information about allocation or processing strategies employed in experiments on intraserial repetition. The major outcome of these experiments was that subject-controlled study times are consistent with the general form of the spacing effect. Study time tended to increase monotonically over spacing conditions although it did not appear to account for all of the variance in recall due to spacing. If nominal study time is considered to be an approximate measure of internal processing activity, the self-pacing technique provides a simple way of determining the effects of incentive on the processing strategy of the subject. The design of experiment I involved a simple factorial combination of spacing and differential payoffs in a free recall task where subjects were allowed to determine the amount of study time for each TBR item.

Because of the importance of processing tradeoffs between repeated and surrounding list items for the processing strategy hypothesis, a set of critical single presentation items, similar to those employed in the Elmes, Greener and Wilkinson experiments, were incorporated into the list structures for the present experiment. In the following text, these items are referred to as adjacent items. For scoring purposes to determine the amount earned by each subject,
all adjacent items were treated as low payoff items. However, for purposes of data analysis, each adjacent item was classed as either high or low payoff depending on whether it followed a high or low payoff repeated item during list presentation. Recall performance and study time measures on such items should provide information on tradeoffs between repeated and adjacent items.

Several predictions can be made concerning the nature of possible processing tradeoffs in this task. First, because it is assumed that adjacent list items compete for access to a limited capacity working memory, the added processing effort should subtract from that normally allocated to adjacent items in the list. Thus, the effects of payoff on recall for adjacent items should be reversed from that of repeated items; in other words, recall should be lower for high payoff than for low payoff adjacent items. Secondly, recall of adjacent items should decline with increasing spacing as observed by Elmes, Greener and Wilkinson (1973).

Based on the outcome of the Hintzman et al. study, spacing effects can be expected for repeated items in both the low and high payoff groups. Of course, better performance is expected for the high-payoff repeated items. The contribution of the present experiment is to determine whether the additivity of spacing and incentives observed by Hintzman et al. can be accounted for by the allocation strategies adopted by subjects under these conditions. For example, if the effect of incentives is simply to add a constant to the encoding effort allocated to repetitions at different values of spacing, there is no reason
to predict that the effects of spacing will be different for the low and high payoff groups.

**Method**

**Materials, apparatus, and design.** Five presentation sequences of 124 separate item presentations were constructed. There were 30 critical double-presentation items in each sequence, 10 items each at interpresentation lags of zero, four, and 12 items. There were also 40 critical single presentation items. The sequences were constructed so that the second presentation of each double presentation item was immediately followed in the list by a critical single presentation (adjacent) item. Ten additional single presentation items were distributed through the sequence. There were also noncritical, 12-item primacy and recency buffers, each of which included four repeated items and four single presentation items. Eighty-six common nouns (above 20 per million word frequency, Thorndike-and Lorge, 1944) were randomly assigned to each sequence, thus comprising five word lists. Half of the double presentation items (including the buffer items) were designated high-payoff items. All single presentation items were considered low payoff items. The experiment was thus a within-subject factorial combination of payoff and spacing (lag) conditions.

The experiment took place in a small, sound-attenuated cubicle. A television monitor and a keyboard, each controlled by a PDP/8 computer, were placed inside the cubicle. Presentation was visual and the payoff status of each double presentation item was indicated
only at P2. High payoff items were accompanied on the screen by a star which appeared immediately to the left of each high payoff word during the second presentation. Low payoff items were not accompanied by an identifying symbol. Subjects could advance the list one item at a time by pressing the space bar of the keyboard.

Procedure. Each participant in the study was run individually. Following an explanation of the apparatus, somewhat modified free recall instructions were given. Several points should be made concerning these instructions. (1) Each participant was allowed to control his own presentation rate in the experiment. Because pilot subjects had indicated that some participants might take abnormally long amounts of time to study the list, each participant was instructed to average approximately five seconds per word. (2) It was required that every word in the list be studied at least some minimal amount of time. That is, subjects were required to expand some effort to memorize each word in the list. (3) Monetary payoffs were not awarded. Instead, the participants were asked to study as if the high payoff items were worth approximately twice as much as the low payoff items. Ten points were assigned arbitrarily to each high payoff item and five points to each low payoff item. Participants were asked to maximize the number of points they earned rather than the number of words they recalled. (4) Finally, participants were explicitly instructed to trade off study time between items. If they studied one item for eight seconds, they should make up for this spending less than five
seconds on subsequent items.

Following the instructions, participants were allowed to practice on a short list similar to the critical list previously described. During this practice session, the experimenter gave performance feedback which consisted primarily of information on the length of time spent studying each item.

Following the practice list, a single experimental list was presented. The order of lists was predetermined so that each of the five unique lists was used an equal number of times. Study times, as measured by the interresponse times for keyboard presses, were recorded by the computer. Immediately following list presentation, a cue for recall appeared on the television monitor. Five minutes was allowed for written free recall.

Subjects. Forty undergraduates at Rice University participated in the study voluntarily. Two participants were lost because of an error on the part of the experimenter and a third was not included in the data analysis because of failure to conform to the instructional requirements of the study.

Results and Discussion

Figure 5 shows percent recall and mean study time for the repeated items at lags zero, four and twelve. The lag functions display similar trends for both percent recall and study time in the two payoff conditions. Both dependent variables tend to increase over lag in the low payoff condition, but there is no consistent increase over lag for
Figure 5. Percent recall and mean study time for each of the six repeated item conditions in experiment I.
either variable in the high payoff condition.

We will first consider the percent recall data for the repeated items. An analysis of variance showed that the payoff instructions resulted in significantly higher recall for the high payoff than for the low payoff items, $F(1,32) = 10.89, p < .01$. The main effect of interpretation lag was also significant, $F(2,72) = 6.57, p < .01$. However, as can be seen in Figure 5, the lag function departs from its established monotonic form in the high payoff condition. The departure was sufficient to produce a significant interaction between lag and payoff condition, $F(2,72) = 3.61, p < .05$, which supports the conclusion that the form of the spacing or lag function is different for the two payoff conditions. It should be noted that this finding is at variance with the payoff study reported by Hintzman, Summers, Eki and Moore (1975). Because of this discrepancy, the basis for the interaction needs to be explored further and supported with additional experimentation.

According to the processing strategy hypothesis, the redistribution of encoding effort among lag conditions should result in a corresponding effect in recall. Given that different patterns of recall have been observed in the two payoff conditions, there are two questions of interest concerning the allocation of processing activity and its relationship to recall performance. First, we need to determine whether the payoff conditions show a redistribution of processing activity as indicated in the mean study times. Secondly, we must ask how well this redistribution accounts for the differences in recall.
In general, the statistical analysis of the study time data agrees qualitatively with that for percent recall. As expected, subjects allotted more study time to high than to low payoff conditions, $F(1,32) = 15.02, p < .001$. There was no main effect of lag, $F(2,64) < 1.0$. However, the interaction between payoff condition and lag was found to be insignificant, $F(2,64) = 6.12, p < .01$. In the low payoff condition study time increased over lag, but there was a slight decrease in study time over lag in the high payoff condition. The a priori hypothesis was that added incentive would simply add a constant amount of encoding effort to repetitions at different values of lag. Instead, there is a clear indication that the payoff manipulation has served to redistribute the relative amount of encoding effort within the high payoff condition.

Because there are likely to be qualitative differences in the processes going on during different item presentations, recorded study time is probably not an ideal measure of encoding effort. Nevertheless, the relationships between study time and recall will provide some important information on the underlying strategies subjects use in dealing with stimulus repetition.

In an effort to gain some information about this relationship, the number of items recalled per unit study time for each subject was determined for each double presentation condition by dividing the number of items recalled by the total study time for those items. These values, averaged over subjects, are shown in Table 2. These values do not seem to vary greatly except for the lag zero items in the low pay-
<table>
<thead>
<tr>
<th>Payoff Condition</th>
<th>Interpresentation Lag</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zero</td>
</tr>
<tr>
<td>High</td>
<td>5.52</td>
</tr>
<tr>
<td>Low</td>
<td>3.30</td>
</tr>
</tbody>
</table>

Table 2. Mean recall/study time ratios for the six repeated item conditions in experiment I. Ratios were determined by the formula (number recalled/total study time) X 100.
off condition. The lower value for that condition indicates that study time is not as effective as for the other conditions in the experiment. The recall/study time ratios were evaluated using an analysis of variance employing individual comparisons. The low-payoff lag-zero mean was compared to the five other means in Table 2 and was found to be significantly different, $F(1,72) = 15.20$, $p < .001$. The orthogonal comparison of the high-payoff lag-zero mean with the four remaining means was not significant, $F < 1.0$. No other significant differences could be obtained with the four remaining degrees of freedom.

Although the analysis of the repeated items did not confirm the initial hypotheses, the results are compatible with a processing strategy interpretation of the spacing effect. The low payoff condition replicates what essentially has been established using the self-pacing technique in free recall—namely that the distribution of self-paced study time is qualitatively in agreement with the form of the spacing function, but that it does not account for all the variance in recall due to the spacing of repetitions (Shaughnessy, Zimmerman and Underwood, 1972; Zimmerman, 1975). There are several reasons why study time might not account for all of the variance due to spacing. The first is that there may be some mutual inhibition between item presentations when the spacing between repetitions is small—basically the consolidation or habituation explanations. The second is that the self-pacing technique may not eliminate all of the difference between nominal and function study time. Subjects may continue to reallocate study time.
to other list items even when they control the presentation rate. The fact that a substantial decrement in study time effectiveness has been demonstrated only for lag zero items in the low payoff condition favors the second explanation. The enhanced effectiveness of study time for lag zero items in the high payoff condition suggests that effective study time can be allocated to lag-zero items under the proper conditions.

**Adjacent items.** The descriptive statistics for the single presentation items adjacent to each repeated item are shown in Table 3. None of the predictions regarding recall of adjacent items appear to be supported. There are no appreciable differences in recall, study time, or recall per unit study time as a function of spacing or payoff condition. This was supported by an analysis of variance, F's < 1.0. These results, of course, do not agree with those of Elmes, Greener and Wilkinson who found substantial differences in recall for such items. However, it is possible that the introduction of the self-pacing technique substantially changes the nature of the free recall task, although the recall data for the repeated items do not support this interpretation. Because of this possibility, experiment II maintained the design feature of having critical items presented adjacent to each repeated item.

The results of this experiment suggest that flexible allocation of effective encoding effort is possible even at small values of spacing. Because different results have been obtained in previous
| Interpresentation Lag | High Payoff Items | | Low Payoff Items | | Control Items |
|----------------------|------------------|------------------|------------------|------------------|
|                      | Zero  | Four  | Twelve | Zero  | Four  | Twelve | Percent Recall | 26   | 24   | 26   |  29   | 30   | 26   |
|                      | Study Time | 4.70  | 4.84  | 4.53  | 4.70  | 4.62  | 4.58  | Percent Recall | 26   |        |      |        |      |      |
|                      | Recall/Study Time Ratios | 6.23  | 5.69  | 5.91  | 6.83  | 6.93  | 6.19  | Recall/Study Time Ratios | 6.14  |        |      |        |      |      |

Table 3. Percent recall, mean study time, and recall/study time ratios ((number recalled/study time) x 100) for each single presentation condition in experiment I. Low and high payoff items occurred adjacent to a critical repeated item in the respective lag conditions. Control items were distributed through the list structure.
research using other measures of memory performance (e.g., Hintzman et al., 1975), additional confirmation and a closer examination of the sources of the effect are necessary. Experiment II addresses these refinements.

EXPERIMENT II

A useful method of determining which items are selected for extended processing is the overt rehearsal technique introduced into the free recall paradigm by Rundus and Atkinson (1970). The application of the technique is simple and straightforward. Subjects are required to rehearse TBR items overtly during list presentation. The rehearsal protocols obtained in this manner are taped recorded, coded, and analyzed by computer. Plots of rehearsal activity and recall then can be plotted to suggest possible relationships between rehearsal patterns and memory performance. In general, rehearsal activity has been shown to be an excellent predictor of recall probability over a number of different experimental conditions (Hogan, 1975; Hogan and Hogan, 1975; Rundus, 1970; Loftus and Dark, 1977). Although there is some disagreement over the interpretation of the relationship between rehearsal protocols and recall performance (cf. Bartz and Jacoby, 1973), the overt rehearsal technique may provide information about processing strategies that is difficult to obtain through other methods.

For our purposes, overt rehearsal is considered an indirect indicator of encoding or processing time. Under the appropriate experimental conditions, enough of the relationship between rehearsal
and recall should be preserved to make this a reasonable assumption. By observing the rehearsal protocols, it should be possible to determine whether repeated items are selected for extra processing (reflected by the rehearsal activity) and whether this processing draws or subtracts from that normally allocated to surrounding list items.

The second experiment employs most of the design features of the first experiment, using high and low payoff items, spacing values of zero, four, and twelve, and the single presentation adjacent items. The principle changes are that overt rehearsal was required of the subjects, presentation rate was experimenter-controlled and monetary payoffs were used.

Method

The same stimulus materials, lists structures, and apparatus were employed in the second experiment as in the first, except that the stimulus words were rerandomized into the list structures. The same method of identifying high payoff items was used (a "star" at P2 only). The six adjacent item conditions also were retained. The experiment took place in a small experimental cubicle which had a one-way mirror for viewing participant behavior. A tape recorder and television monitor were placed inside the cubicle. A teletype was programmed to advance its carriage on each presentation. The sound of the advancing carriage was clearly audible inside the experimental cubicle and was used to mark the beginning of each stimulus presentation on the recording tape. The experiment was thus designed as a
within-subject factorial combination of payoffs and lag.

Procedure. Subjects were asked to participate for five consecutive days in order to practice and adapt to the overt rehearsal requirement. The first day of the experiment was used as practice for the subject. Following an explanation of the overt rehearsal task and the free recall procedure, a short demonstration list was presented. The presentation of one of the five experimental lists followed. Presentation rate was five seconds per item with a one-second interpresentation interval. The sequence of lists viewed by each subject was determined by a complete Latin square rotation so that each list was used an equal number of times on each day. At the end of each list, a cue for recall appeared on the television monitor. Five minutes was allowed for written free recall.

On the second day, instructions concerning the two payoff conditions were given. Participants received three cents for each low payoff item they recalled and six cents for each high payoff item. After the instructions, a short practice list was presented and subjects were allowed to ask questions about the procedure. The experimental list for that day was then presented. The only procedural change between day one and day two was the introduction of new instructions. Days three, four and five were different from day two only in that short "reminder" instructions were read at the beginning of each session and the practice list was omitted.
Subjects. Thirty undergraduates at Rice University volunteered to participate in the study. They were paid $1.50 per hour in addition to the monetary payoffs.

Results

Percent recall for repeated items averaged over days and subjects is shown in Figure 6. The displayed trends bear considerable similarity to those obtained in experiment I. Recall increases monotonically over spacing in the low payoff condition, but fails to show the typical spacing effect in the high payoff condition. The statistical analysis of percent recall supports these conclusions and confirms the payoff effects found in experiment I.

An analysis of variance indicated that significantly more high payoff items than low payoff items were recalled, $F(1,29) = 20.98$, $p < .001$. This is consistent with the hypothesis that high payoff items are marked or selected for additional processing at P2. The effect of interpretation lag averaged over payoff conditions was significant, $F(2,58) = 6.83$, $p < .01$. However, the interaction between payoff condition and spacing was also significant, $F(2,58) = 5.83$, $p < .01$, indicating that the form of the spacing function differed for the two payoff conditions. The spacing effect was clearly reduced for the high payoff condition and there was no statistical evidence that any of the three means in the high payoff condition arose from a different population, $F(2,58) < 1.0$. These results support experiment I in the conclusion that there are some
Figure 6. Percent recall for the six repeated item conditions in experiment II.
qualitative differences in the processing for the two payoff conditions.

The analysis of variance also showed that recall increased over days in the experiment—percent recall for repeated items averaged over payoffs and lag was 43, 47, 56 and 60 for days two, three, four and five respectively, $F(3,87) = 14.97$, $p < .001$. However, these increases were consistent over all conditions in the experiment and none of the interaction effects involving days were significant, $F's < 1.44$, $p > .10$. Nonspecific transfer effects are common to most kinds of memory tasks, including free recall (Postman, Burns and Hasher, 1970) and are to be expected in any multi-list experiment which provides rest periods between sessions. Because there were no systematic changes in the patterns of recall or in rehearsal, averaging over days should not change the interpretation of Figure 6.

The data in Figure 6 suggest that flexible allocation of processing or encoding effort to repetitions is possible, and is consistent with the hypothesis that both value and accessibility may be used as criteria for selecting an item for extra processing. This rule-based flexibility is counter to any theory which explains spacing effects on the basis of an automatic or involuntary mechanism operating at small values of spacing. However, such conclusions must be qualified by the Hintzman, Summers, Eki and Moore (1975) study which was similar in intent and design but did not obtain the critical spacing by payoff interaction. There are several differences between these studies, but a highly salient difference is the use of different dependent measures, frequency judgments by Hintzman et al. and free recall in
the present study. This suggests some alternative explanations for the payoff effect found in experiments I and II.

**Output order.** The major procedural difference between frequency judgment experiments and free recall is in the kind of response required of the subject. In the frequency judgment paradigm, subjects are presented a test with the stimulus items and asked to estimate the frequency with which each occurred in the experimental list. Thus, it establishes a greater amount of control over the test phase of an experiment than does free recall. Because the free recall task leaves the order of response output uncontrolled, it is possible that the introduction of incentive has its influence on the decoding or retrieval stage rather than on encoding. In other words, the encoding of all lag zero items may be deficient in the sense discussed earlier, but high payoff items may have a retrieval advantage which obscures this deficient encoding. For example, it is well documented that the recall of an item in free recall decreases the probability that other unrecalled items will also be recalled (Smith, 1971; Roediger, 1973, 1974). Variables which affect recall order therefore have the potential of producing spurious effects on the spacing function. For example, if lag zero items in the high payoff condition tended to be recalled before other list items, the spacing function would tend to be flattened much like that observed in experiments I and II.

Although it is difficult to think of reasonable theoretical hypotheses that would lead to such output order effects, we can easily
examine the empirical hypothesis that the payoff effects are an artifact of recall order. The mean and variance of the output position for each of the six repeated item conditions is shown in Table 4. High payoff items tend to be recalled earlier in the output sequence, but there is no differential tendency to recall lag zero before lag four or lag twelve items. If, for some reason, lag zero items tended to be recalled both first and last, this would show up as larger variance for such items. However, the variances in Table 4 do not support this interpretation.

A second possibility is that the flattened recall function for the high payoff condition could be produced by an interaction between recall order and spacing. Suppose the spacing function for the first few items recalled was flat. Because high payoff items tend to be recalled first, the spacing function for high payoff items would tend to be flatter than that for the low payoff items. To provide information on this hypothesis, each subject's recall was divided into thirds. Since there are several choices for the denominator to determine percent recall, the total number recalled for each repeated item condition is plotted for the first third and the last two thirds of recall in Figure 7. The spacing function for the first third of the items recalled is indeed somewhat flatter for the low payoff condition. However, the payoff effect (a flattened spacing function) is quite pronounced in both panels of the figure. Thus, it is unlikely that the payoff effect could be interpreted as an interaction between spacing and recall order. Because none of this evidence suggests
<table>
<thead>
<tr>
<th></th>
<th>Zero</th>
<th>Four</th>
<th>Twelve</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Payoff</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpresentation Lag</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>21.69</td>
<td>19.74</td>
<td>22.14</td>
</tr>
<tr>
<td>Variance</td>
<td>664.61</td>
<td>577.92</td>
<td>716.10</td>
</tr>
<tr>
<td><strong>Low Payoff</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpresentation Lag</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>21.98</td>
<td>22.47</td>
<td>22.99</td>
</tr>
<tr>
<td>Variance</td>
<td>723.28</td>
<td>690.64</td>
<td>705.96</td>
</tr>
</tbody>
</table>

Table 4. Mean and variance of output position for repeated items at each level of payoff and lag in Experiment II.
Figure 7. Total number recalled in each of the six repeated item conditions for the first third of items recalled (bottom) and last two thirds of items recalled (top).
that the interaction between payoff and spacing is an artifact of recall order, the working hypothesis adopted here is that the locus of the effect is in the encoding stage, perhaps as reflected in the patterns of rehearsal activity.

Rehearsal. Table 5 shows the mean rehearsals per item broken into four periods relative to the first and second presentations of a repeated item. Total rehearsal per item is shown in the final column. The use of mean rehearsal as an indicator of processing or encoding effort requires the establishment of some systematic relationship between rehearsal and recall performance. Given that some relationship has been established, the pattern of rehearsal can be used to infer something about the nature of processing or encoding strategies in a given task situation. The relationships between percent recall and mean rehearsal deviate somewhat from what would be expected ideally, primarily because mean rehearsal tends to increase over lag in contrast to the rather flat function obtained for recall. Nevertheless, the correlation between mean rehearsal and recall was .87, accounting for nearly 80% of the variance. This appears to be strong enough to merit a closer examination of the rehearsal data.

The total rehearsal per item in the last column show that high payoff items received more rehearsals than low payoff items. This is consistent with the recall data, suggesting that high payoff items are selected or marked for additional processing and perhaps maintained in a transient state of accessibility while information is being encoded
<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P1-P2</th>
<th>P2</th>
<th>P2-Test</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Payoff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 0</td>
<td>3.20</td>
<td>-----</td>
<td>2.94</td>
<td>4.82</td>
<td>10.96</td>
</tr>
<tr>
<td>Lag 4</td>
<td>3.13</td>
<td>1.50</td>
<td>2.82</td>
<td>4.56</td>
<td>12.01</td>
</tr>
<tr>
<td>Lag 12</td>
<td>3.19</td>
<td>2.06</td>
<td>3.04</td>
<td>4.93</td>
<td>13.22</td>
</tr>
<tr>
<td><strong>Low Payoff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 0</td>
<td>3.41</td>
<td>-----</td>
<td>2.73</td>
<td>2.64</td>
<td>8.78</td>
</tr>
<tr>
<td>Lag 4</td>
<td>3.30</td>
<td>1.25</td>
<td>2.82</td>
<td>3.54</td>
<td>10.91</td>
</tr>
<tr>
<td>Lag 12</td>
<td>3.26</td>
<td>1.79</td>
<td>2.85</td>
<td>3.18</td>
<td>11.08</td>
</tr>
</tbody>
</table>

Table 5. Mean rehearsals during and after P1 and P2 as a function of payoff and interpresentation lag. Totals are shown in the final column.
or assembled into a more stable form. A portion of the rehearsal differences between spacing conditions is attributable to the fact that there is more opportunity for rehearsal in the interval between presentations for longer values of lag. Hintzman (1974) has argued that this cannot be used to explain the spacing effect in free recall because it suggests that the source of deficient processing is P1 rather than P2. It should also be noted, however, that some of the rehearsal differences between spacing conditions is found in the interval between P2 and test. A similar pattern was obtained by Rundus (1971). In addition, differences in rehearsal due to payoffs are substantial in the P2-test interval. The events in this interval, therefore, may be important in understanding the effects of incentive and how it interacts with spacing to produce a flattened spacing function.

Table 6 shows the mean number of rehearsals per critical item at P2 and for ten item presentations following P2. The final column shows the mean number of rehearsals per item which occurred in the remainder of the list presentation. There is a small amount of displaced rehearsal activity in each item presentation following P2 (displaced in the sense that rehearsal of a critical item does not coincide with the item's own presentation). Displaced rehearsal processes have been reported in many overt rehearsal studies using adult subjects in the free recall task (Hogan, 1975; Hogan and Hogan, 1975; Rundus, 1971; Rundus and Atkinson, 1970). The amount of rehearsal during any given item presentation following P2 is small,
<table>
<thead>
<tr>
<th></th>
<th>P2</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
<th>+4</th>
<th>+5</th>
<th>+6</th>
<th>+7</th>
<th>+8</th>
<th>+9</th>
<th>+10</th>
<th>Remainder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag 0</td>
<td>2.73</td>
<td>.56</td>
<td>.33</td>
<td>.25</td>
<td>.13</td>
<td>.07</td>
<td>.04</td>
<td>.05</td>
<td>.04</td>
<td>.03</td>
<td>.06</td>
<td>1.08</td>
</tr>
<tr>
<td>Low Payoff</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 4</td>
<td>2.82</td>
<td>.54</td>
<td>.23</td>
<td>.13</td>
<td>.11</td>
<td>.13</td>
<td>.07</td>
<td>.08</td>
<td>.10</td>
<td>.10</td>
<td>.15</td>
<td>1.90</td>
</tr>
<tr>
<td>Lag 12</td>
<td>2.85</td>
<td>.42</td>
<td>.31</td>
<td>.21</td>
<td>.05</td>
<td>.10</td>
<td>.04</td>
<td>.12</td>
<td>.08</td>
<td>.10</td>
<td>.06</td>
<td>1.67</td>
</tr>
<tr>
<td>Lag 0</td>
<td>2.94</td>
<td>.69</td>
<td>.42</td>
<td>.25</td>
<td>.23</td>
<td>.18</td>
<td>.21</td>
<td>.14</td>
<td>.15</td>
<td>.20</td>
<td>.10</td>
<td>2.25</td>
</tr>
<tr>
<td>High Payoff</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 4</td>
<td>2.82</td>
<td>.54</td>
<td>.34</td>
<td>.28</td>
<td>.13</td>
<td>.18</td>
<td>.16</td>
<td>.24</td>
<td>.22</td>
<td>.20</td>
<td>.16</td>
<td>2.11</td>
</tr>
<tr>
<td>Lag 12</td>
<td>3.04</td>
<td>.54</td>
<td>.31</td>
<td>.32</td>
<td>.21</td>
<td>.25</td>
<td>.20</td>
<td>.18</td>
<td>.22</td>
<td>.17</td>
<td>.15</td>
<td>2.38</td>
</tr>
</tbody>
</table>

Table 6. Mean number of rehearsals for each item presentation in the P2-test interval. The first ten item presentations following P2 are shown. Remaining rehearsals not occurring in the first ten presentations after P2 are shown in the final column.
but over many item presentations the amount of displaced rehearsal is large enough to play an important role in the overall process. It is clear from Table 6 that high payoff items maintain a small advantage throughout this interval, and that lag zero items in the low payoff condition show the least amount of displaced rehearsal. This pattern of rehearsal is consistent with the hypothesis that certain classes of repeated items may be selected or marked for additional processing at P2. A simple two-stage decision rule in which an item was first selected for processing on the basis of its value, and secondly on the basis of its accessibility in memory, would account for such patterns. With this information in mind, it is now possible to return to the question of possible tradeoffs between repeated and adjacent single presentation items in free recall.

**Adjacent items.** Percent recall and mean rehearsal for the adjacent items is shown in Table 7. Neither percent recall nor rehearsal shows any systematic trends as a function of lag or payoffs. The analysis of variance indicated that the only significant effect was that of days, $F(3,87) = 20.03$, $p < .001$. As in the case of the repeated items, percent recall increased over days—26, 33, 42, 43 for days two, three, four and five. None of the other main effects or interactions approached significance, $F's < 1.54$, $p > .10$.

There have now been several attempts to replicate the findings of Elmes, Greener and Wilkinson (1972), the two experiments reported here and the three experiments reported by Hintzman and Stern (1977).
<table>
<thead>
<tr>
<th>Interpresentation Lag</th>
<th>High Payoff Items</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zero</td>
<td>Four</td>
<td>Twelve</td>
</tr>
<tr>
<td>Recall</td>
<td>37</td>
<td>34</td>
<td>33</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>6.44</td>
<td>5.75</td>
<td>6.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interpresentation Lag</th>
<th>Low Payoff Items</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zero</td>
<td>Four</td>
<td>Twelve</td>
</tr>
<tr>
<td>Recall</td>
<td>38</td>
<td>35</td>
<td>37</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>5.82</td>
<td>5.57</td>
<td>6.77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control Items</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>37</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>5.79</td>
</tr>
</tbody>
</table>

Table 7. Percent recall and mean rehearsal for each of the single presentation item conditions in experiment II. High and low payoff items occurred adjacent to a critical repeated item in the indicated condition. Control items were distributed throughout the presentation sequence.
These studies qualify the generality of the Elmes et al. results, perhaps limiting them to a very narrow range of conditions in free recall. However, these negative findings do not eliminate the possibility that tradeoffs between repeated and single presentation items occur, only that the specific tradeoffs reported by Elmes et al. are not general.

The rehearsal data for repeated items in Table 6 suggest a possible reason why the present experiment failed to replicate Elmes et al. Although both high payoff and spaced repetitions are indeed rehearsed more than low payoff items, the rehearsal advantage is rather small for any given item presentation. Given limited resources, the extra rehearsal allocated to high payoff and spaced repetitions may subtract from the amount of processing that can be allocated to other list items. However, such tradeoffs are possibly generalized over a lengthy section of list presentation. This would prevent tradeoffs from being localized at the serial positions nearest P2. In such cases, only between list manipulations like that employed by Waugh (1970) would be sensitive to tradeoffs.

Discussion

Several alternative hypotheses for the payoff effects in experiments I and II have been examined. The rehearsal data in Table 6 suggest that displaced rehearsal in the P2-test interval may be the mediating factor for these effects. We may consider the potential functions that such rehearsal processes could serve in the free recall
task. There appear to be several. First, displaced rehearsal will serve to reactivate the original encoding of an item in memory. It seems reasonable that an item could be maintained in an active state for some time with an occasional rehearsal, even if sufficient information to support long term recall had not yet been encoded into memory. According to the theoretical position developed earlier, displaced rehearsal could be triggered by a variety of criteria aimed at achieving certain goals for the subject. Under most experimental conditions, the subject's implicit goal is to remember as many items as possible. An optimal strategy would therefore be based on selecting only those items for rehearsal which were least accessible in memory. However, if a subject is attempting to maximize points or earnings in situations where items have different values, optimal strategies probably involve giving highly valued items rehearsal priority. A decision rule in which items are selected for extended rehearsal first by their value and secondly on the basis of accessibility would account for the rehearsal data of Table 6 and perhaps for the recall patterns observed in experiments I and II.

A second function of displaced rehearsal is that each rehearsal would serve to introduce a TBR item into a different list context. Modern theoretical interpretations of free recall emphasize the importance of interitem associations (Anderson, 1972; Postman, 1972). If recall performance is considered to be based on associative encoding, displaced rehearsal may function to set up multiple associative encodings between list items which provide multiple access or retrieval
routes in memory. Thus, repeated items which are selected or marked for extra rehearsal following P2 would have a higher probability of recall.

Speculation about the encoding of multiple retrieval routes to information in memory is, of course, the central idea in encoding variability theory. The foregoing account suggests a possible relationship between the notion of processing or encoding strategies and encoding variability theory. From this point of view, the processing strategy hypothesis and encoding variability theory may be complimentary, rather than competing, explanations.

EXPERIMENT III

Experiments I and II were intended to provide some empirical support for the utility of processing strategy concepts in understanding repetition effects in free recall. A major question for this approach concerns the spacing effects which have been obtained in incidental learning tasks. Because subjects are not informed of a forthcoming memory test, the use of rules or strategies in these situations seems to be unlikely. Thus, such spacing effects have been used to support automatic or involuntary theories of the spacing effect, (Hintzman, 1976, p. 79).

Shaughnessy (1976) studied the effects of spacing repetitions in several different incidental learning tasks. These tasks required judgments or ratings of various sorts (imageability, pleasantness of meaning, duration, and word frequency) on common English nouns pre-
ceding an unanticipated recall test. These results appear to be most consistent with an automatic or involuntary theory of spacing effects in which the repetitive activation of the same internal representation close in time is inhibitory in some fashion. However, the interpretation of these findings depends on the theoretical processes believed to underlie incidental learning performance.

There are several alternative hypotheses that can be entertained and which are based on the strategy a subject adopts in contending with an incidental learning task. One account is a simple version of the inattention theory proposed by Underwood (1970). When the spacing between repetitions is small, it seems reasonable that the initial judgment will still be in active memory and can be repeated without additional processing of the item. With larger values of spacing, however, it is more likely that the initial judgment will be forgotten, thus requiring additional processing to reformulate the original judgment. This hypothesis can be tested by comparing tasks in which either different or the same judgments are required on each repetition of an item. The requirement of different judgments on each repetition would prevent the strategy of simply repeating the previous judgment at small values of spacing.

A more interesting hypothesis can be devised from a consideration of the processes involved in generating accurate judgments of verbal concepts. First, we might consider what kinds of information the subject must have available if he is to perform adequately on the rating task. Because rating assignments are relative to the set of
stimuli chosen for an experiment, it may be necessary for the subject to make comparisons between successive list items as he proceeds through the task. That is, on each trial a set of comparison stimuli must be retrieved from the set of list items previously judged before an accurate category assignment can be made.

One might note the similarity between a hypothetical comparison operation in the rating task and the displaced rehearsal discussed in the previous experiment. The comparison process may serve to set up a network of interitem associations much like that assumed to form the basis for intentional recall. With these ideas, spacing effects can be predicted in terms of the contextual changes that are a natural consequence of distributing repetitions. When repetitions are spaced, a new pool of comparison items is introduced between the first and second presentations. The re-evaluation of an item on the basis of a new context may lead to a richer network of associations for recall. At very small values of spacing, the pool of items on which the implicit comparison process is based has changed very little. Thus, comparison will lead to the encoding of only a few new interitem associations.

In order to test this latter hypothesis, it is necessary to devise a task in which judgmental comparison is not between successive list items. The simplest way to accomplish this is to provide explicit comparison or reference "anchor points" on each item presentation. Since comparison will be between list item and anchor points, contextual changes due to the spacing of repetitions will not result in the encoding of additional interitem associations.
To distinguish between the automatic or inhibitory mechanism hypothesis, the version of the inattention hypothesis offered above, and the hypothesis based on interitem comparisons and contextual change, three incidental learning tasks were developed. All three tasks required concreteness judgments similar to those used to develop concreteness scales (Pavio, Madigan and Yuille, 1968). Although this particular rating task has not been studied previously, it was selected because it meets the criteria for a good incidental task in that concreteness judgments must be based on semantic features of the rated items. In the first experimental task, the procedures were similar to those employed by Shaughnessy (1976). Subjects made concreteness judgments on a seven-point unanchored category scale (unanchored rating task). The other two tasks differed from the first in that explicit anchor points or reference items were provided on each item presentation. These tasks included instructions to make all judgments relative only to the reference item presented on each trial. In one of these latter tasks, the reference items changed on each repetition of a TBR item (anchored rating task-different). The other task differed only in that the reference items were the same on each repetition (anchored rating task-same).

The pattern of effects obtained in these three tasks can serve to distinguish the three hypotheses. If spacing effects result from some inhibitory mechanism which operates at small values of lag, the provision of reference items should make no difference in the task and spacing effects are predicted in all three incidental tasks. According
to the simple version of inattention theory, spacing effects are predicted in the two tasks where the judgment is the same on each repetition, but not in the different task. The list comparison hypothesis predicts a spacing effect only in the unanchored rating task since the interitem comparisons encouraged by the provision of anchor or reference items will be irrelevant to the recall process.

Method

Materials and design. Three list structures were constructed, each 66 item presentations in length, with slots for 10 item repetitions each at interpresentation lags of zero and four. There were also 10 critical single presentation items in each list and eight-item primacy and recency buffers. Forty-six common nouns (above 20 per million Thorndike-Lorge frequency) were selected from the Pavio, Madigan and Yuille (1968) norms with concreteness ratings in the range 3.0 to 5.0. Sixteen of these items were randomly selected as buffer items for inclusion in the primacy and recency buffers. The other thirty items were divided in groups of ten each. These groups of items were rotated through conditions (single presentation, repeated-lag zero, and repeated-lag four) in the three list structures to comprise three different lists with different items serving as single presentation, lag zero and lag four items in each list. Four additional items were selected from the Pavio, Madigan and Yuille norms as reference items in two of the incidental tasks. These items were table (mean concreteness judgment of 7.00), knowledge (mean judgment of 1.56), idea (1.42) and product
(5.80). Four combinations of these items were used as anchor points, knowledge-table, knowledge-product, idea-table, and idea-product. A practice list, ten items in length was constructed of items from the same source. For the two tasks with anchor points, the same combinations of the four items were used in the practice list. The basic design of the experiment was thus a three (instructional group) X two two (spacing) factorial with repeated measures on the last factor.

Instructions and procedure. Three sets of instructions were written to correspond to the three incidental tasks described above. Each task was introduced as a procedure for developing concreteness scales for subsequent memory research. In the unanchored rating task, subjects were instructed to make their ratings relative to the total set of items presented for rating. In the two tasks with reference items, the instructions were to make the ratings relative only to the specific target and reference items presented on a given trial. No specific instructions were given on how to deal with the item repetitions, although each subject was made aware that item repetitions would occur in the rating task.

Subjects participated in groups of two with each individual in a separate experimental cubicle where a television monitor was placed. Stimulus presentation was visual with sequences and rates controlled by a PDP/8 computer. Each target item (the items to be rated on the concreteness scale) was presented for five seconds. There was a two second interpresentation interval. On each presentation the target
item was presented in the approximate middle of the display. At the top was a reminder of the scale type, the numbers one through seven spaced across the display, and at the bottom was a trial number for each presentation. For the unanchored rating task, the word "abstract" appeared above the scale value "1" at the top of the screen, and the word "concrete" appeared above the scale value "7". For the two tasks with anchor points, the two reference items appeared in place of the words "abstract" and "concrete". The leftmost word was always the more abstract of the two as determined by the Pavio, Madigan and Yuille norms. The instructions asked subjects to consider the leftmost reference item to have an arbitrary scale value of "1" and the rightmost item to have an arbitrary value of "7". The rating task was thus to assign the target item an appropriate value on the scale of concreteness relative to the two end points. Subjects were also instructed to use the five-second item presentation to formulate their judgment and the two-second interpresentation interval to record their rating assignments on prepared response sheets.

Following the reading of instructions, the practice list was presented. Questions were allowed during this period. Subjects were told that the practice list was merely to allow them to see the display, adjust it if necessary, and to get used to the pace of the task. One of the three experimental lists was presented, with subjects assigned to lists on the basis of the order in which they signed up for the experiment. Following list presentation, a surprise recall test was requested by presenting a recall sheet with printed instructions. The instructions emphasized that only target items would be counted as
correct recalls. Five minutes was allowed for recall. Subjects were then questioned as to whether they expected a recall test and debriefed as to the purpose and nature of the experiment.

**Subjects.** Seventy-two Rice University undergraduates volunteered to participate in the study. Twenty-four subjects were assigned to each of the three incidental tasks, eight subjects in each group to the three different lists. None indicated that they expected the surprise recall test.

**Results and Discussion**

Percent recall for both single and double presentation items is shown in Figure 8. The data were analyzed with a series of planned comparisons. The first hypothesis tested was that performance in the various incidental tasks was based on judgmental comparisons between list items. Such interitem comparisons are more likely to occur in the unanchored rating task. Thus, one would expect higher performance in this group. The comparison of the overall performance in the unanchored rating task versus the average performance in the two anchored rating tasks was significant, $F(1,69) = 10.84$, $p < .005$. The two groups which employed anchor points did not differ significantly between themselves, $F(1,69) < 1.0$. According to the hypothesis, this result would occur if the comparison operation set up a network of interitem associations in which the recall of one item served as a cue for the recall of subsequent items.

The list comparison hypothesis also predicts that a spacing effect
Figure 8. Percent recall for both repeated and single presentation items in the three incidental tasks for experiment III.
should be obtained only in the unanchored rating task. Figure 8 indicates that recall for lag four items is greater than recall for lag zero items only in the unanchored task. Statistical analysis showed that the spacing effect was significant in the unanchored rating task, $F(1,69) = 8.35$, $p < .01$, but not in the two anchored rating tasks, $F's < 1.0$. The overall task by spacing interaction was significant, $F(2,69) = 4.14$, $p < .05$. However, the component of the interaction most relevant to the list comparison hypothesis is the test for differences in the effect of spacing for the unanchored task versus the average effect for the two anchored tasks. This component was significant, $F(1,69) = 7.42$, $p < .01$, and accounted for 90% of the interaction sums of squares.

The important finding, of course, is the failure to obtain spacing effects in the two anchored rating tasks. Since this requires, in effect, accepting the null hypothesis, it will be useful to examine confidence intervals for the differences between lags zero and four in the two anchored rating tasks. Stated as percent recall, the 99% confidence interval calculated for the two groups combined is $-6.3 < \mu < 3.3$. The likelihood that the effect of spacing under these conditions is as great as 4% is therefore quite small. It is not unusual to find spacing effects exceeding 10% in the literature. For example, all effects obtained by Shaughnessy (1976) for judgments of imageability, pleasantness of meaning, and word frequency exceeded 10% for spacing values of zero and four.

It seems reasonable to conclude that spacing effects do not occur
automatically in incidental learning tasks. Instead, such effects may be the outcome of the same kinds of associative encoding processes that form the basis for intentional recall. Judgmental comparisons between list items can be considered as one strategy of contending with the problem of generating accurate performance on unanchored category scales. In this sense, the results of experiment III can be viewed as an outcome of a processing or encoding strategy. However, the pattern of effects are not consistent with a simple inattention hypothesis. The notion of encoding or processing strategies needs to be supplemented with some more basic ideas as to the structural changes which take place in a memory trace to make it more retrievable. In this case, the encoding variability concept appears to provide a natural account. Encoding variability explains the effects of stimulus repetition in terms of the number of associative relationships established between list items. The retrieveability of an item, then, may be enhanced by experimental procedures which encourage an item to be associatively encoded with many other list items (spacing of repetitions), or by subject-controlled processing strategies (the displaced rehearsal of experiment II) which, in effect, accomplish the same end result.

GENERAL DISCUSSION

The predominant assumption in research on the spacing effect has been that a common, underlying cause is responsible for the effects of spacing repetitions found in many different memory tasks. Apart
from the similarity of the spacing functions obtained across different tasks, there are several ways to confirm this assumption. One method is to find independent variables which manipulate the form of the spacing function in similar ways. Glenberg (1976, 1977) has provided evidence for this assumption by showing that the manipulation of the P2-test interval consistently produces nonmonotonic and monotonic effects in a similar fashion in free recall, recognition memory and paired associates learning. Additional evidence is found in studies which restrict rehearsal strategies through instructions. Ciccone and Brelsford (1974) reported that restricting the natural rehearsal strategies both lowered recognition performance and reduced the magnitude of the spacing effect. Essentially the same findings have been established in free recall (Glenberg, 1977; Wright and Brelsford, in press).

On the other hand, the extent to which independent variables produce differential effects on the spacing function in different tasks can be used to infer that the underlying processes are not identical. There is now some evidence to indicate that certain instructional variables produce different effects in the frequency judgment task than have been found in other memory tasks. Hintzman, Summers, Eki and Moore (1975, exp. II) reported that rehearsal restrictions have no effect on the spacing function for frequency judgments. Given the effects of such manipulations in free recall and recognition, this finding indicates that the effects of spacing in the frequency judgment paradigm may require a different interpretation.
To this evidence, it is now possible to add the payoff manipulation reported in experiments I and II, which obtained somewhat different results from a similar manipulation using frequency judgments (Hintzman, Summers, Eki and Moore, 1975, exp. I).

There are, in fact, other kinds of evidence to suggest that the frequency judgment task is somewhat different from free recall and recognition. For example, developmental changes in the ability to estimate frequency are either small or nonexistent (Hasher and Chromiak, 1977). In contrast, there are rather large developmental changes in both recall and recognition (cf. Flavell, 1976). The lack of developmental changes in frequency estimation is interesting within the present framework because of growing evidence in the developmental literature that developmental changes are always obtained when strategies form the basis for task performance (Brown, 1974).

Other evidence for the uniqueness of the frequency judgment task can be cited. Knowledge of what kind of memory test is forthcoming appears to have little effect on the ability to estimate frequency (Howell, 1973), in contrast to the effects such knowledge has on both recall and recognition (Tversky, 1973). The bulk of this evidence clearly suggests that frequency information accrues automatically with event perception.

This robust ability to estimate frequency suggests a possible interpretation of the different payoff effects obtained in the frequency judgment task and in free recall. The analysis of rehearsal processes in experiment II indicates that the majority of the process-
ing advantage of high-payoff items occurs subsequent to P2. Few theories of memory would predict that this increased rehearsal would not lead to enhanced performance. However, there is some reason to believe that such rehearsal processes may not lead to enhanced frequency judgments. A recent study reported by Johnson, Taylor, and Raye (1977) indicates that people have a substantial although imperfect ability to separate the frequency of internally generated events from that of externally generated events. Displaced rehearsal clearly qualifies as an internally generated event. Its effects on frequency judgments would therefore be limited by the extent to which internally generated and externally generated events are confused. If we assume a reasonably competent ability to separate these two sources of information about frequency, displaced rehearsal would probably have little effect on frequency judgments.

The uniqueness of the frequency judgment task is further suggested by the classification of studies presented in Table 8. Each study in the table was discussed to some degree in the introduction as supporting or not supporting one of the four predictions regarding the processing strategy hypothesis. The three experiments subsequently reported are also included in this table. Frequency judgment studies are identified by underlining. As can be seen from the table, the majority of the nonsupporting evidence has come from the frequency judgment paradigm. In addition, the exceptions found in the left column have been provided with interpretations within the framework of the present research. Assuming that these interpretations are acceptable, Table 8
<table>
<thead>
<tr>
<th>Non-supporting Evidence</th>
<th>Supporting Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effects of spacing</strong>&lt;br&gt;in incidental learning tasks.</td>
<td>Rose and Rowe (1976), Rowe and Rose (1977), Shaughnessy (1976)</td>
</tr>
</tbody>
</table>

Table 8. Classification of studies relating to the four predictions discussed in the introduction. Studies in the frequency judgment paradigm are underlined.
strongly indicates that the spacing effects obtained in the frequency judgment task should be given a different interpretation. However, such interpretations are beyond the scope of the present paper.

Thus far, the theoretical explanation of spacing and payoff effects have relied heavily on the concept of rehearsal and other covert processes which bear a considerable similarity to rehearsal. One question that may be asked, then, concerns whether the interpretation of spacing effects offered here is limited to those conditions where rehearsal can be employed. For example, Hintzman (1974) has taken the position that explanations based on differential rehearsal cannot form the basis for a general theory of the spacing effect for the simple reason that the effect occurs in paradigms where the opportunity for rehearsal is carefully controlled or with stimulus materials that are difficult to rehearse. Rehearsal, however, is merely one form of subject-controlled processing. It seems reasonable that the strategy of allocating processing capacity on the basis of accessibility in memory will be manifested within the framework of the processing techniques available to a subject during a particular task. The free recall task lends itself to analysis in terms of rehearsal and interitem associative encoding. It is only natural that explanations of intraserial repetitions in free recall be stated in these terms. The processing strategy hypothesis is not invalidated simply because subjects adopt different processing techniques in other paradigms. Of course, this does not obviate the need to investigate the moment in which processing strategies might fit within the framework of process
models designed to explain other kinds of memory performance.

With a few exceptions, previous theories have tended to be stated at a level which is independent of any existing task specific processes. Eventually such theories must be related to the specific process models which are used to account for memory performance. The value of present research may perhaps be to show how the notion of processing or encoding strategies is compatible with some basic ideas about performance in the free recall task.
REFERENCES


Glanzer, M. Distance between related words in free recall: Trace of the STS. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 105-111.


Howell, W. C. Storage of events and event frequencies: A comparison


Potts, G. R. Distance from a massed double presentation or blank trial as a factor in paired associates learning. *Journal of Verbal Behavior, 1972, 11*, 375-386.


Tversky, B. Encoding processes in recall and recognition. *Cognitive


