2 HUMAN MEMORY: A PROPOSED SYSTEM AND ITS CONTROL PROCESSES[1]

R. C. Atkinson and R. M. Shiffrin

STANFORD UNIVERSITY
STANFORD, CALIFORNIA

[1] This research was supported by the National Aeronautics and Space Administration, Grant No. NGR-05-020-036. The authors are indebted to W. K. Estes and G. H. Bower who provided many valuable suggestions and comments at various stages of the work. Special credit is due J. W. Brelsford who was instrumental in carrying out the research discussed in Section IV and whose overall contributions are too numerous to report in detail. We should also like to thank those co-workers who carried out a number of the experiments discussed in the latter half of the paper; rather than list them here, each will be acknowledged at the appropriate place.

I. Introduction

This paper is divided into two major portions; the first outlines a general theoretical framework in which to view human memory, and the second describes the results of a number of experiments designed to test specific models that can be derived from the overall theory.

The general theoretical framework, set forth in Sections II and III, categorizes the memory system along two major dimensions. One categorization distinguishes permanent, structural features of the system from control processes that can be readily modified or reprogrammed at the will of the subject. Because we feel that this distinction helps clarify a number of results, we will take time to elaborate it at the outset. The permanent features of memory, which will be referred to as the memory structure, include both the physical system and the built-in processes that are unvarying and fixed from one situation to another. Control processes, on the other hand, are selected, constructed, and used at the option of the subject and may vary dramatically from one task to another even though superficially the tasks may appear very similar. The use of a particular control process in a given situation will depend upon such factors as the nature of the instructions, the meaningfulness of the material, and the individual subject's history.

A computer analogy might help illustrate the distinction between memory structure and control processes. If the memory system is viewed as a computer under the direction of a programmer at a remote console, then both the computer hardware and those programs built into the system that cannot be modified by the programmer are analogous to our structural features; those programs and instruction sequences which the programmer can write at his console and which determine the operation of the computer, are analogous to our control processes. In the sense that the computer's method of processing a given batch of data depends on the operating program, so the way a stimulus input is processed depends on the particular control processes the subject brings into play. The structural components include the basic memory stores; examples of control processes are coding procedures, rehearsal operations, and search strategies.

Our second categorization divides memory into three structural components: the sensory register, the short-term store, and the long-term store. Incoming sensory information first enters the sensory register, where it resides for a very brief period of time, then decays and is lost. The short-term store is the subject's working memory; it receives selected inputs from the sensory register and also from long-term store. Information in the short-term store decays completely and is lost within a period of about 30 seconds, but a control process called rehearsal can

maintain a limited amount of information in this store as long as the subject desires. The long-term store is a fairly permanent repository for information, information which is transferred from the short-term store. Note that "transfer" is not meant to imply that information is removed from one store and placed in the next; we use transfer to mean the copying of selected information from one store into the next without removing this information from the original store.

In presenting our theoretical framework we will consider first the structural features of the system (Section II) and then some of the more generally used control processes (Section III). In both of these sections the discussion is organized first around the sensory register, then the short-term store, and finally the long-term store. Thus, the outline of Sections II and III can be represented as follows:

	Sensory register	Short-term store	Long-term store
Structure	Sec. II,A	Sec. II,B	Sec. II,C
Control processes	Sec. III,A	Sec. III,B	Sec. III,C

These first sections of the paper do not present a finished theory; instead they set forth a general framework within which specific models can be formulated. We attempt to demonstrate that a large number of results may be handled parsimoniously within this framework, even without coming to final decisions at many of the choice points that occur. At some of the choice points several hypotheses will be presented, and the evidence that is available to help make the choice will be reviewed. The primary goal of Sections II and III is to justify our theoretical framework and to demonstrate that it is a useful way of viewing a wide variety of memory phenomena.

The remaining sections of the paper present a number of precise models that satisfy the conditions imposed by our general theoretical framework. These sections also present data from a series of experiments designed to evaluate the models. Section IV is concerned with an analysis of short-term memory; the model used to analyze the data emphasizes a control process based in the short-term store which we designate a rehearsal buffer. Section V presents several experiments that shed some light upon processes in the long-term store, especially subject-controlled search processes. Some of the experiments in Sections IV and V have been reported by us and our co-workers in previous publications, but the earlier treatments were primarily mathematical whereas the present emphasis is upon discussion and overall synthesis.

If the reader is willing to accept our overall framework on a provisional

basis and wishes to proceed at once to the specific models and experiments, then he may begin with Section IV and as a prerequisite need only read that portion of Section III,B concerned with the rehearsal buffer.

II. Structural Features of the Memory System

This section of the paper will describe the permanent, structural features of the memory system. The basic structural division is into the three components diagrammed in Fig. 1: the sensory register, the short-term store, and the long-term store.

When a stimulus is presented there is an immediate registration of that stimulus within the appropriate sensory dimensions. The form of this registration is fairly well understood in the case of the visual system (Sperling, 1960); in fact, the particular features of visual registration (including a several hundred millisecond decay of an initially accurate visual image) allow us positively to identify this system as a distinct component of memory. It is obvious that incoming information in other sense modalities also receives an initial registration, but it is not clear whether these other registrations have an appreciable decay period or any other features which would enable us to refer to them as components of memory.

The second basic component of our system is the short-term store. This store may be regarded as the subject's "working memory." Information entering the short-term store is assumed to decay and disappear completely, but the time required for the information to be lost is considerably longer than for the sensory register. The character of the information in the short-term store does not depend necessarily upon the form of the sensory input. For example, a word presented visually may be encoded from the visual sensory register into an auditory short-term store. Since the auditory short-term system will play a major role in subsequent discussions, we shall use the abbreviation a-v-l to stand for auditory-verbal-linguistic store. The triple term is used because, as we shall see, it is not easy to separate these three functions.

The exact rate of decay of information in the short-term store is difficult to estimate because it is greatly influenced by subject-controlled processes. In the a-v-l mode, for example, the subject can invoke rehearsal mechanisms that maintain the information in STS and thereby complicate the problem of measuring the structural characteristics of the decay process. However, the available evidence suggests that information represented in the a-v-l mode decays and is lost within a period of about 15–30 seconds. Storage of information in other modalities

is less well understood and, for reasons to be discussed later, it is difficult to assign values to their decay rates.

The last major component of our system is the long-term store. This store differs from the preceding ones in that information stored here does not decay and become lost in the same manner. All information eventually is completely lost from the sensory register and the short-term store,

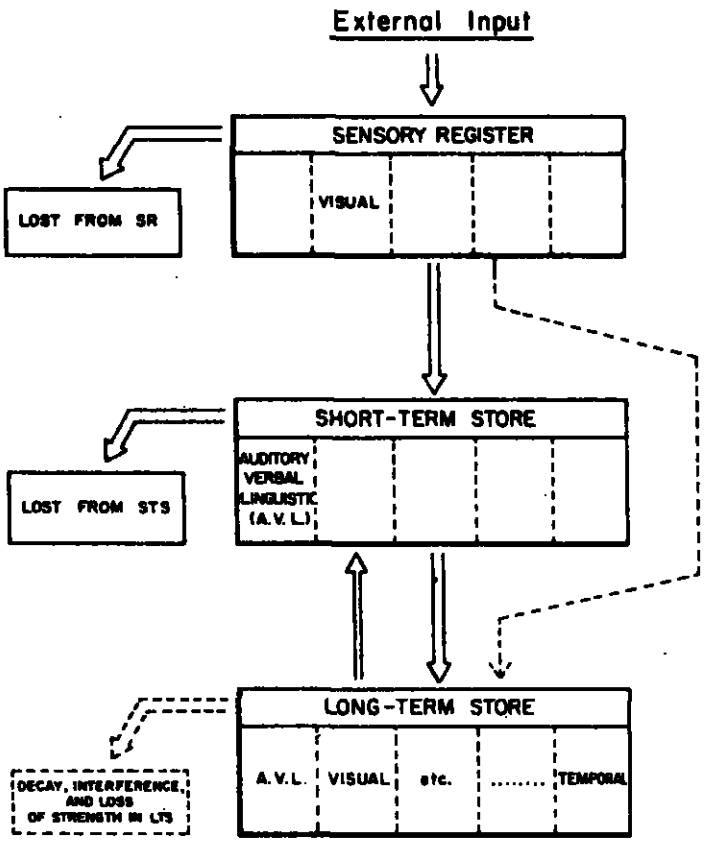


FIG. 1. Structure of the memory system.

whereas information in the long-term store is relatively permanent (although it may be modified or rendered temporarily irretrievable as the result of other incoming information). Most experiments in the literature dealing with long-term store have been concerned with storage in the a-v-l mode, but it is clear that there is long-term memory in each of the other sensory modalities, as demonstrated by an ability to recognize stimuli presented to these senses. There may even be information

in the long-term store which is not classifiable into any of the sensory modalities, the prime example being temporal memory.

The flow of information among the three systems is to a large degree under the control of the subject. Note that by information flow and transfer between stores we refer to the same process: the copying of selected information from one store into the next. This copying takes place without the transferred information being removed from its original store. The information remains in the store from which it is transferred and decays according to the decay characteristics of that store. In considering information flow in the system, we start with its initial input into the sensory register. The next step is a subject-controlled scan of the information in the register; as a result of this scan and an associated search of long-term store, selected information is introduced into short-term store. We assume that transfer to the long-term store takes place throughout the period that information resides in the short-term store, although the amount and form of the transferred information is markedly influenced by control processes. The possibility that there may be direct transfer to the long-term store from the sensory register is represented by the dashed line in Fig. 1; we do not know whether such transfer occurs. Finally, there is transfer from the long-term store to the short-term store, mostly under the control of the subject; such transfer occurs, for example, in problem solving, hypothesis testing, and "thinking" in general.

This brief encapsulation of the system raises more questions than it answers. Not yet mentioned are such features as the cause of the decay in each memory store and the form of the transfer functions between the stores. In an attempt to specify these aspects of the system, we now turn to a more detailed outline, including a review of some relevant literature.

A. Sensory Register

The prime example of a sensory register is the short-term visual image investigated by Sperling (1960, 1963), Averbach and Coriell (1961), Estes and Taylor (1964, 1966), and others. As reported by Sperling (1967), if an array of letters is presented tachistoscopically and the subject is instructed to write out as many letters as possible, usually about six letters are reported. Further, a 30-second delay between presentation and report does not cause a decrement in performance. This fact (plus the facts that confusions tend to be based on auditory rather than visual similarities, and that subjects report rehearsing and subvocalizing the letters) indicates that the process being examined is in the a-v-l short-term store; i.e., subjects scan the visual image and transfer a number of letters to the a-v-l short-term store for rehearsal and output.

In order to study the registered visual image itself, partial-report procedures (Averbach & Coriell, 1961; Averbach & Sperling, 1961; Sperling, 1960, 1963) and forced-choice detection procedures (Estes, 1965; Estes & Taylor, 1964, 1966; Estes & Wessel, 1966) have been employed. The partial-report method typically involves presenting a display (usually a 3 × 4 matrix of letters and numbers) tachistoscopically for a very brief period. After the presentation the subject is given a signal that tells him which row to report. If the signal is given almost immediately after stimulus offset, the requested information is reported with good precision, otherwise considerable loss occurs. Thus we infer that a highly accurate visual image lasts for a short period of time and then decays. It has also been established that succeeding visual stimulation can modify or possibly even erase prior stimulation. By using a number of different methods, the decay period of the image has been estimated to take several hundred milliseconds, or a little more, depending on experimental conditions; that is, information cannot be recovered from this store after a period of several hundred milliseconds.

Using the detection method, in which the subject must report which of two critical letters was presented in a display, Estes and Taylor (1964, 1966) and Estes and Wessel (1966) have examined some models for the scanning process. Although no completely satisfactory models have yet been proposed, it seems reasonably certain that the letters are scanned serially (which letters are scanned seems to be a momentary decision of the subject), and a figure of about 10 msec to scan one letter seems generally satisfactory.

Thus it appears fairly well established that a visual stimulus leaves a more or less photographic trace which decays during a period of several hundred milliseconds and is subject to masking and replacement by succeeding stimulation. Not known at present is the form of the decay, that is, whether letters in a display decay together or individually, probabilistically or temporally, all-or-none, or continuously. The reader may ask whether these results are specific to extremely brief visual presentations; although presentations of long duration complicate analysis (because of eye movements and physical scanning of the stimulus), there is no reason to believe that the basic fact of a highly veridical image quickly decaying after stimulus offset does not hold also for longer visual presentations. It is interesting that the stimulation seems to be transferred from the visual image to the a-v-l short-term store, rather than to a visual short-term store. The fact that a written report was requested may provide the explanation, or it may be that the visual short-term store lacks rehearsal capacity.

There is not much one can say about registers in sensory modalities other than the visual. A fair amount of work has been carried out on the

auditory system without isolating a registration mechanism comparable to the visual one. On the other hand, the widely differing structures of the different sensory systems makes it questionable whether we should expect similar systems for registration.

Before leaving the sensory register, it is worth adding a few comments about the transfer to higher order systems. In the case of the transfer from the visual image to the a-v-l short-term store, it seems likely that a selective scan is made at the discretion of the subject.[2] As each element in the register is scanned, a matching program of some sort is carried out against information in long-term store and the verbal "name" of the element is recovered from long-term memory and fed into the short-term store. Other information might also be recovered in the long-term search; for example, if the scanned element was a pineapple, the word, its associates, the taste, smell, and feel of a pineapple might all be recovered and transferred to various short-term stores. This communication between the sensory register and long-term store does not, however, permit us to infer that information is transferred directly to long-term store from the register. Another interesting theoretical question is whether the search into long-term store is necessary to transfer information from the sensory register to the short-term store within a modality. We see no a priori theoretical reason to exclude nonmediated transfer. (For example, why should a scan or match be necessary to transfer a spoken word to the a-v-l short-term store?) For lack of evidence, we leave these matters unspecified.

B. Short-Term Store

The first point to be examined in this section is the validity of the division of memory into short- and long-term stores. Workers of a traditional bent have argued against dichotomizing memory (e.g., Melton, 1963; Postman, 1964). However, we feel there is much evidence indicating the parsimony and usefulness of such a division. The argument is often given that one memory is somehow "simpler" than two; but quite the opposite is usually the case. A good example may be found in a comparison of the model for free recall presented in this paper and the model proposed by Postman and Phillips (1965). Any single-process system making a fair attempt to explain the mass of data currently available must, of necessity, be sufficiently complex that the term *single process* becomes a misnomer. We do not wish, however, to engage in the controversy here. We ask the reader to accept our model provisionally until its power to deal with data becomes clear. Still, some justification

[2] Sperling (1960) has presented evidence relating the type of scan used to the subject's performance level.

of our decision would seem indicated at this point. For this reason, we turn to what is perhaps the single most convincing demonstration of a dichotomy in the memory system: the effects of hippocampal lesions reported by Milner (1959, 1966, 1968). In her words:

"Bilateral surgical lesions in the hippocampal region, on the mesial aspect of the temporal lobes, produce a remarkably severe and persistent memory disorder in human patients, the pattern of breakdown providing valuable clues to the cerebral organization of memory. Patients with these lesions show no loss of pre-operatively acquired skills, and intelligence as measured by formal tests is unimpaired, but, with the possible exception of acquiring motor skill, they seem largely incapable of adding new information to the long-term store. This is true whether acquisition is measured by free recall, recognition, or learning with savings. Nevertheless, the immediate registration of new input (as measured, for example, by digit span and dichotic listening tests) appears to take place normally and material which can be encompassed by verbal rehearsal is held for many minutes without further loss than that entailed in the initial verbalization. Interruption of rehearsal, regardless of the nature of the distracting task, produces immediate forgetting of what went before, and some quite simple material which cannot be categorized in verbal terms decays in 30 seconds or so, even without an interpolated distraction. Material already in long-term store is unaffected by the lesion, except for a certain amount of retrograde amnesia for preoperative events" (Milner, 1966).

Apparently, a short-term store remains to the patients, but the lesions have produced a breakdown either in the ability to store new information in long-term store or to retrieve new information from it. These patients appear to be incapable of retaining new material on a long-term basis.[3]

As with most clinical research, however, there are several problems that should be considered. First, the patients were in a general sense abnormal to begin with; second, once the memory defect had been discovered, the operations were discontinued, leaving only a few subjects for observation; third, the results of the lesions seem to be somewhat variable, depending for one thing upon the size of the lesion, the larger lesions giving rise to the full syndrome. Thus there are only a few patients who exhibit the deficit described above in full detail. As startling as these patients are, there might be a temptation to discount them as anomalies but for the following additional findings. Patients who had

[3] A related defect, called Korsakoff's syndrome, has been known for many years. Patients suffering from this abnormal condition are unable to retain new events for longer than a few seconds or minutes (e.g., they cannot recall the meal they have just eaten or recognize the face of the doctor who treated them a few minutes earlier), but their memory for events and people prior to their illness remains largely unimpaired and they can perform adequately on tests of immediate memory span. Recent evidence suggests that Korsakoff's syndrome is related to damage of brain tissue, frequently as the result of chronic alcoholism, in the hippocampal region and the mammillary body (Barbizet, 1963).

known damage to the hippocampal area in one hemisphere were tested for memory deficit after an intracarotid injection of sodium amytal temporarily inactivated the other hemisphere. Controls were patients without known damage, and patients who received injections inactivating their damaged side. A number of memory tests were used as a criterion for memory deficit; the easiest consisted of presenting four pictures, distracting the patient, and then presenting nine pictures containing the original four. If the patient cannot identify the critical four pictures then evidence of memory deficit is assumed. The results showed that in almost all cases memory deficit occurs only after bilateral damage; if side A is damaged and side B inactivated, memory deficit appears, but if the inactivated side is the damaged side, no deficit occurs. These results suggest that the patients described above by Milner were not anomalous cases and their memory deficits therefore give strong support to the hypothesis of distinct short- and long-term memory stores.

1. *Mechanisms Involved in Short-Term Store*

We now turn to a discussion of some of the mechanisms involved in the short-term store. The purpose of this section is not to review the extensive literature on short-term memory, but rather to describe a few experiments which have been important in providing a basis for our model. The first study in this category is that of Peterson and Peterson (1959). In their experiment subjects attempted to recall a single trigram of three consonants after intervals of 3, 6, 9, 12, 15, and 18 seconds. The trigram, presented auditorily, was followed immediately by a number, and the subject was instructed to count backward by three's from that number until he received a cue to recall the trigram. The probability of a correct answer was nearly perfect at 3 seconds, then dropped off rapidly and seemed to reach an asymptote of about .08 at 15–18 seconds. Under the assumption that the arithmetic task played the role of preventing rehearsal and had no direct interfering effect, it may be concluded that a consonant trigram decays from short-term store within a period of about 15 seconds. In terms of the model, the following events are assumed to occur in this situation: the consonant trigram enters the visual register and is at once transferred to the a-v-l short-term store where an attempt is made to code or otherwise "memorize" the item. Such attempts terminate when attention is given to the task of counting backward. In this initial period a trace of some sort is built up in long-term store and it is this long-term trace which accounts for the .08 probability correct at long intervals. Although discussion of the long-term system will come later, one point should be noted in this context; namely, that the long-term trace should be more powerful the more

repetitions of the trigram before arithmetic, or the longer the time before arithmetic. These effects were found by Hellyer (1962); that is, the model predicts the probability correct curve will reach an asymptote that reflects long-term strength, and in the aforementioned experiment, the more repetitions before arithmetic, the higher the asymptote.

It should be noted that these findings tie in nicely with the results from a similar experiment that Milner (1968) carried out on her patients. Stimuli that could not be easily coded verbally were used; for example, clicks, light flashes, and nonsense figures. Five values were assigned to each stimulus; a test consisted of presenting a particular value of one stimulus, followed by a distracting task, followed by another value of the stimulus. The subject was required to state whether the two stimuli were the same or different. The patient with the most complete memory deficit was performing at a chance level after 60 seconds, whether or not a distracting task was given. In terms of the model, the reduction to chance level is due to the lack of a long-term store. That the reduction occurred even without a distracting task indicates that the patient could not readily verbalize the stimuli, and that rehearsal in modes other than the verbal one was either not possible or of no value. From this view, the better asymptotic performance demonstrated by normal subjects on the same tasks (with or without distraction) would be attributed to a long-term trace. At the moment, however, the conclusion that rehearsal is lacking in nonverbal modes can only be considered a highly tentative hypothesis.

We next ask whether or not there are short-term stores other than in the a-v-l mode, and if so, whether they have a comparable structure. A natural approach to this problem would use stimuli in different sense modalities and compare the decay curves found with or without a distracting task. If there was reason to believe that the subjects were not verbally encoding the stimuli, and if a relatively fast decay curve was found, then there would be evidence for a short-term memory in that modality. Furthermore, any difference between the control group and the group with a distracting task should indicate the existence of a rehearsal mechanism. Posner (1966) has undertaken several experiments of this sort. In one experiment the subject saw the position of a circle on a 180-mm line and later had to reproduce it; in another the subject moved a lever in a covered box a certain distance with only kinesthetic feedback and later tried to reproduce it. In both cases, testing was performed at 0, 5, 10, and 20 seconds; the interval was filled with either rest, or one of three intervening tasks of varying difficulty. These tasks, in order of increasing difficulty, consisted of reading numbers, adding numbers, and classifying numbers into categories. For the kinesthetic task there was a decline in performance over 30 seconds,

but with no obvious differences among the different intervening conditions. This could be taken as evidence for a short-term kinesthetic memory without a rehearsal capability. For the visual task, on the other hand, there was a decline in performance over the 30 seconds only for the two most difficult intervening tasks; performance was essentially constant over time for the other conditions. One possibility, difficult to rule out, is that the subjects' performance was based on a verbal encoding of the visual stimulus. Posner tends to doubt this possibility for reasons that include the accuracy of the performance. Another possibility is that there is a short-term visual memory with a rehearsal component; this hypothesis seems somewhat at variance with the results from Milner's patient who performed at chance level in the experiment cited above. Inasmuch as the data reported by Posner (1966) seem to be rather variable, it would probably be best to hold off a decision on the question of rehearsal capability until further evidence is in.

2. *Characteristics of the a-v-l Short-Term Store*

We restrict ourselves in the remainder of this section to a discussion of the characteristics of the a-v-l short-term store. Work by Conrad (1964) is particularly interesting in this regard. He showed that confusions among visually presented letters in a short-term memory task are correlated with the confusions that subjects make when the same letters are read aloud in a noise background; that is, the letters most confused are those sounding alike. This might suggest an auditory short-term store, essentially the auditory portion of what has been called to this point an a-v-l store. In fact, it is very difficult to separate the verbal and linguistic aspects from the auditory ones. Hintzman (1965, 1967) has argued that the confusions are based upon similar kinesthetic feedback patterns during subvocal rehearsal. When subjects were given white noise on certain trials, several could be heard rehearsing the items aloud, suggesting subvocal rehearsal as the usual process. In addition, Hintzman found that confusions were based upon both the voicing qualities of the letters and the place of articulation. The place-of-articulation errors indicate confusion in kinesthetic feedback, rather than in hearing. Nevertheless, the errors found cannot be definitely assigned to a verbal rather than an auditory cause until the range of auditory confusions is examined more thoroughly. This discussion should make it clear that it is difficult to distinguish between the verbal, auditory, and linguistic aspects of short-term memory; for the purposes of this paper, then, we group the three together into one short-term memory, which we have called the a-v-l short-term store. This store will henceforth be labeled STS. (Restricting the term STS to the a-v-l mode

does not imply that there are not other short-term memories with similar properties.)

The notation system should be made clear at this point. As just noted, STS refers to the auditory-verbal-linguistic short-term store. LTS will refer to the comparable memory in long-term store. It is important not to confuse our theoretical constructs STS and LTS (or the more general terms short-term store and long-term store) with the terms short-term memory (STM) and long-term memory (LTM) used in much of the psychological literature. These latter terms have come to take on an operational definition in the literature; STM refers to the memory examined in experiments with short durations or single trials, and LTM to the memory examined in long-duration experiments, typically list learning, or multiple-list learning experiments. According to our general theory, both STS and LTS are active in both STM and LTM experiments. It is important to keep these terms clear lest confusion results. For example, the Keppel and Underwood (1962) finding that performance in the Peterson situation is better on the first trials of a session has been appropriately interpreted as evidence for proactive interference in short-term memory (STM). The model we propose, however, attributes the effect to changes in the long-term store over the session, hence placing the cause in LTS and not STS.

At this point a finished model would set forth the structural characteristics of STS. Unfortunately, despite a large and growing body of experiments concerned with short-term memory, our knowledge about its structure is very limited. Control processes and structural features are so complexly interrelated that it is difficult to isolate those aspects of the data that are due solely to the structure of the memory system. Consequently, this paper presumes only a minimal structure for STS; we assume a trace in STS with auditory or verbal components which decays fairly rapidly in the absence of rehearsal, perhaps within 30 seconds. A few of the more promising possibilities concerning the precise nature of the trace will be considered next. Because most workers in this area make no particular distinction between traces in the two systems, the comments to follow are relevant to the memory trace in the long-term as well as the short-term store.

Bower (1967a) has made a significant exploration of the nature of the trace. In his paper, he has demonstrated the usefulness of models based on the assumption that the memory trace consists of a number of pieces of information (possibly redundant, correlated, or in error, as the case may be), and that the information ensemble may be construed as a multicomponent vector. While Bower makes a strong case for such a viewpoint, the details are too lengthy to review here. A somewhat different approach has been proposed by Wickelgren and Norman (1966)

who view the trace as a unidimensional strength measure varying over time. They demonstrate that such a model fits the results of certain types of recognition-memory experiments if the appropriate decay and retrieval assumptions are made. A third approach is based upon a phenomenon reported by Murdock (1966), which has been given a theoretical analysis by Bernbach (1967). Using methods derived from the theory of signal detectability, Bernbach found that there was an all-or-none aspect to the confidence ratings that subjects gave regarding the correctness of their response. The confidence ratings indicated that an answer was either "correct" or "in error" as far as the subject could tell; if intermediate trace strengths existed, the subject was not able to distinguish between them. The locus of this all-or-none feature, however, may lie in the retrieval process rather than in the trace; that is, even if trace strengths vary, the result of a retrieval attempt might always be one of two distinct outcomes: a success or a failure. Thus, one cannot rule out models that assume varying trace strengths. Our preference is to consider the trace as a multicomponent array of information (which we shall often represent in experimental models by a unidimensional strength measure), and reserve judgment on the locus of the all-or-none aspect revealed by an analysis of confidence ratings.

There are two experimental procedures which might be expected to shed some light on the decay characteristics of STS and both depend upon controlling rehearsal; one is similar to the Peterson paradigm in which rehearsal is controlled by an intervening activity and the other involves a very rapid presentation of items followed by an immediate test. An example of the former procedure is Posner's (1966) experiment in which the difficulty of the intervening activity was varied. He found that as the difficulty of an intervening task increased, accuracy of recall decreased.

Although this result might be regarded as evidence that decay from STS is affected by the kind of intervening activity, an alternative hypothesis would ascribe the result to a reduction in rehearsal with more difficult intervening tasks. It would be desirable to measure STS decay when rehearsal is completely eliminated, but it has proved difficult to establish how much rehearsal takes place during various intervening tasks.

Similar problems arise when attempts are made to control rehearsal by increasing presentation rates. Even at the fastest conceivable presentation rates subjects can rehearse during presentation if they attend to only a portion of the incoming items. In general, experiments manipulating presentation rate have not proved of value in determining decay characteristics for STS, primarily because of the control processes the subject brings into play. Thus Waugh and Norman (1965) found no

difference between 1-second and 4-second rates in their probe digit experiment; Conrad and Hille (1958) found improvement with faster rates; and Buschke and Lim (1967) found increases in the amount of primacy in their missing-span serial position curves as input rate increased from one item per second to four items per second. Complex results of this sort make it difficult to determine the structural decay characteristics of STS. Eventually, models that include the control processes involved in these situations should help clarify the STS structure.

3. *Transfer from STS to LTS*

The amount and form of information transferred from STS to LTS is primarily a function of control processes. We will assume, however, that transfer itself is an unvarying feature of the system; throughout the period that information resides in the short-term store, transfer takes place to long-term store. Support for such an assumption is given by studies on incidental learning which indicate that learning takes place even when the subject is not trying to store material in the long-term store. Better examples may be the experiments reported by Hebb (1961) and Melton (1963). In these experiments subjects had to repeat sequences of digits. If a particular sequence was presented every several trials, it was gradually learned. It may be assumed that subjects in this situation attempt to perform solely by rehearsal of the sequence within STS; nevertheless, transfer to LTS clearly takes place. This Hebb-Melton procedure is currently being used to explore transfer characteristics in some detail. R. L. Cohen and Johansson (1967), for example, have found that an overt response to the repeated sequence was necessary for improvement in performance to occur in this situation; thus information transfer is accentuated by overt responses and appears to be quite weak if no response is demanded.

The form of the STS-LTS transfer may be probabilistic, continuous, or some combination; neither the literature nor our own data provide a firm basis for making a decision. Often the form of the information to be remembered and the type of test used may dictate a particular transfer process, as for example in Bower's (1961) research on an all-or-none paired-associate learning model, but the issue is nevertheless far from settled. In fact, the changes in the transfer process induced by the subject effectively alter the transfer function form experiment to experiment, making a search for a universal, unchanging process unproductive.

C. Long-Term Store

Because it is easiest to test for recall in the a-v-l mode, this part of long-term store has been the most extensively studied. It is clear, how-

ever, that long-term memory exists in each of the sensory modalities; this is shown by subjects' recognition capability for smells, taste, and so on. Other long-term information may be stored which is not necessarily related to any of the sensory modalities. Yntema and Trask (1963), for example, have proposed that temporal memory is stored in the form of "time-tags." Once again, however, lack of data forces us to restrict our attention primarily to the a-v-l mode, which we have designated LTS.

First a number of possible formulations of the LTS trace will be considered. The simplest hypothesis is to assume that the trace is all-or-none; if a trace is placed in memory, then a correct retrieval and response will occur. Second-guessing experiments provide evidence concerning an hypothesis of this sort.

Binford and Gettys (1965) presented the subject with a number of alternatives, one of which was the correct answer. If his first response is incorrect, he picks again from the remaining alternatives. The results indicate that second guesses are correct well above the chance level to be expected if the subject were guessing randomly from the remaining alternatives. This result rules out the simple trace model described above because an all-or-none trace would predict second guesses to be at the chance level. Actually, the above model was a model of both the form of the trace and the type of retrieval. We can expand the retrieval hypothesis and still leave open the possibility of an all-or-none trace. For example, in searching for a correct all-or-none trace in LTS, the subject might find a similar but different trace and mistakenly terminate the search and generate an answer; upon being told that the answer is wrong the subject renews the search and may find the correct trace the next time. Given this hypothesis, it would be instructive to know whether the results differ if the subject must rank the response alternatives without being given feedback after each choice. In this case all the alternatives would be ranked on the basis of the same search of LTS; if the response ranked second was still above chance, then it would become difficult to defend an all-or-none trace.

A second source of information about the nature of the trace comes from the tip-of-the-tongue phenomenon examined by Hart (1965), R. Brown and McNeill (1966), and Freedman and Landauer (1966). This phenomenon refers to a person's ability to predict accurately that he will be able to recognize a correct answer even though he cannot recall it at the moment. He feels as if the correct answer were on the "tip of the tongue." Experiments have shown that if subjects who cannot recall an answer are asked to estimate whether they will be able to choose the correct answer from a set of alternatives, they often show good accuracy in predicting their success in recognition. One explanation might be that the subject recalls some information, but not enough to generate an

answer and feels that this partial information is likely to be sufficient to choose among a set of alternatives. Indeed, Brown and McNeill found that the initial sound of the word to be retrieved was often correctly recalled in cases where a correct identification was later made. On the other hand, the subject often is absolutely certain upon seeing the correct response that it is indeed correct. This might indicate that some new, relevant information has become available after recognition. In any case, a simple trace model can probably not handle these results. A class of models for the trace which can explain the tip-of-the-tongue phenomenon are the multiple-copy models suggested by Atkinson and Shiffrin (1965). In these schemes there are many traces or *copies of* information laid in long-term store, each of which may be either partial or complete. In a particular search of LTS perhaps only a small number or just one of these copies is retrieved, none complete enough to generate the correct answer; upon recognition, however, access is gained to the other copies, presumably through some associative process. Some of these other copies contain enough information to make the subject certain of his choice. These multiple-copy memory models are described more fully in Atkinson and Shiffrin (1965).

The decay and/or interference characteristics of LTS have been studied more intensively over the past 50 years than any other aspect of memory. Partly for this reason a considerable body of theory has been advanced known as interference theory.[4] We tend to regard this theory as descriptive rather than explanatory; this statement is not meant to detract from the value of the theory as a whole, but to indicate that a search for mechanisms at a deeper level might prove to be of value. Thus, for example, if the interfering effect of a previously learned list upon recall of a second list increases over time until the second list is retested, it is not enough to accept "proactive interference increasing over time" as an explanation of the effect; rather one should look for the underlying search, storage, and retrieval mechanisms responsible.

We are going to use a very restricted definition of interference in the rest of this paper; interference will be considered a structural feature of memory not under the control of the subject. It will refer to such possibilities as disruption and loss of information. On the other hand, there are search mechanisms which generate effects like those of structural interference, but which are control processes. Interference theory, of course, includes both types of possibilities, but we prefer to break down interference effects into those which are structurally based, and those under the control of the subject. Therefore the term *interference* is used henceforth to designate a structural feature of the long-term system.

[4] For an overview of interference theory see Postman (1961).

It is important to realize that often it is possible to explain a given phenomenon with either interference or search notions. Although both factors will usually be present, the experimental situation sometimes indicates which is more important. For example, as we shall see in Section V, the decrease in the percentage of words recalled in a free verbal-recall experiment with increases in list length could be due either to interference between items or to a search of decreasing effectiveness as the number of items increase. The typical free recall situation, however, forces the subject to engage in a search of memory at test and indicates to us that the search process is the major factor. Finally, note that the interference effect itself may take many forms and arise in a number of ways. Information within a trace may be destroyed, replaced, or lessened in value by subsequent information. Alternatively, information may never be destroyed but may become irretrievable, temporarily or permanently.

In this section an attempt has been made to establish a reasonable basis for at least three systems—the sensory register, the short-term store, and the long-term store; to indicate the transfer characteristics between the various stores; and to consider possible decay and interference functions within each store.

III. Control Processes in Memory

The term *control process* refers to those processes that are not permanent features of memory, but are instead transient phenomena under the control of the subject; their appearance depends on such factors as instructional set, the experimental task, and the past history of the subject. A simple example of a control process can be demonstrated in a paired-associate learning task involving a list of stimuli each paired with either an A or B response (Bower, 1961). The subject may try to learn each stimulus-response pair as a separate, integral unit or he may adopt the more efficient strategy of answering B to any item not remembered and attempting to remember only the stimuli paired with the A response. This latter scheme will yield a radically different pattern of performance than the former; it exemplifies one rather limited control process. The various rehearsal strategies, on the other hand, are examples of control processes with almost universal applicability.

Since subject-controlled memory processes include any schemes, coding techniques, or mnemonics used by the subject in his effort to remember, their variety is virtually unlimited and classification becomes difficult. Such classification as is possible arises because these processes, while under the voluntary control of the subject, are nevertheless dependent upon the permanent memory structures described in the

previous section. This section therefore will follow the format of Section II, organizing the control processes into those primarily associated with the sensory register, STS, and LTS. Apart from this, the presentation will be somewhat fragmentary, drawing upon examples from many disparate experiments in an attempt to emphasize the variety, pervasiveness, and importance of the subject-controlled processes.

A. Control Processes in the Sensory Register

Because a large amount of information enters the sensory register and then decays very quickly, the primary function of control processes at this level is the selection of particular portions of this information for transfer to the short-term store. The first decision the subject must make concerns which sensory register to attend to. Thus, in experiments with simultaneous inputs from several sensory channels, the subject can readily report information from a given sense modality if so instructed in advance, but his accuracy is greatly reduced if instructions are delayed until after presentation. A related attention process is the transfer to STS of a selected portion of a large information display within a sensory modality. An example to keep in mind here is the scanning process in the visual registration system. Letters in a tachistoscopically presented display may be scanned at a rate of about 10 msec a letter, the form of the scan being under the control of the subject. Sperling (1960) found the following result. When the signal identifying which row to report from a matrix of letters was delayed for an interval of time following stimulus offset, the subjects developed two observing strategies. One strategy consisted of obeying the experimenter's instructions to pay equal attention to all rows; this strategy resulted in evenly distributed errors and quite poor performance at long delays. The other strategy consisted of anticipating which row would be tested and attending to only that row; in this case the error variance is increased but performance is better at longer delay intervals than for the other strategy. The subjects were aware of and reported using these strategies. For example, one experienced subject reported switching from the first to the second strategy in an effort to maximize performance when the delay between presentation and report rose above .15 seconds. The graph of his probability of a correct response plotted against delay interval, while generally decreasing with delay, showed a dip at about .15 seconds, indicating that he did not switch strategies soon enough for optimal performance.

The decisions as to which sensory register to attend to, and where and what to scan within the system, are not the only choices that must be made at this level. There are a number of strategies available to the subject for matching information in the register against the long-term

store and thereby identifying the input. In an experiment by Estes and Taylor (1966) for example, the subject had to decide whether an F or B was embedded in a matrix display of letters. One strategy would have the subject scan the letters in order, generating the "name" of each letter and checking to see whether it is a B or an F. If the scan ends before all letters are processed, and no B or F has been found, the subject would presumably guess according to some bias. Another strategy might have the subject do a features match on each letter against B and then F, moving on as soon as a difference is found; in this strategy it would not be necessary to scan all features of each letter (i.e., it would not be necessary to generate the name of each letter). A third strategy might have the subject compare with only one of the crucial letters, guessing the other if a match is not found by the time the scan terminates.

B. Control Processes in Short-Term Store

1. *Storage, Search, and Retrieval Strategies*

Search processes in STS, while not as elaborate as those in LTS because of the smaller amount of information in STS through which the search must take place, are nevertheless important. Since information in STS in excess of the rehearsal capability is decaying at a rapid rate, a search for a particular datum must be performed quickly and efficiently. One indirect method of examining the search process consists of comparing the results of recognition and recall experiments in which STS plays the major role. Presumably there is a search component in the recall situation that is absent in the recognition situation. It is difficult to come to strong conclusions on this basis, but recognition studies such as Wickelgren and Norman (1966) have usually given rise to less complicated models than comparable recall experiments, indicating that the search component in STS might be playing a large role.

One result indicating that the STS search occurs along ordered dimensions is based upon binaural stimulus presentation (Broadbent, 1954, 1956, 1958). A pair of items is presented, one to each ear simultaneously. Three such pairs are given, one every half second. Subjects perform best if asked to report the items first from one ear and then the other, rather than, say, in pairs. While Broadbent interprets these results in terms of a postulated time needed to switch attention from one ear to the other (a control process in itself), other interpretations are possible. In particular, part of the information stored with each item might include which ear was used for input. This information might then provide a simple dimension along which to search STS and report during recall. Another related possibility would have the subject group the

items along this dimension during presentation. In any case we would expect similar results if another dimension other than "sides" (which ear) were provided. Yntema and Trask (1963) used three word-number pairs presented sequentially, one every half second; one member of a pair was presented to one ear and the other member to the other ear. There were three conditions: the first in which three words were presented consecutively on one side (and therefore the three numbers on the other), the second in which two words and one number were presented consecutively on one side, the third in which a number separated the two words on one side. Three test conditions were used: the subject was asked to report words, the numbers (types); or to report one ear followed by the other (sides); or the simultaneous pairs in order (pairs). The results are easy to describe. In terms of probability correct, presentation condition one was best, condition two next, and condition three worst. For the test conditions, "types" yielded the highest probability of correct response, followed by "sides" and then "pairs." "Sides" being better than "pairs" was one of the results found by Broadbent, but "types" being even better than "sides" suggests that the organization along available dimensions, with the concomitant increase of efficiency in the search process, is the dominant factor in the situation.

One difficulty in studying the search process in STS is the fact that the subject will perform perfectly if the number of items presented is within his rehearsal span. Sternberg (1966) has overcome this difficulty by examining the latency of responses within the rehearsal span. His typical experiment consists of presenting from one to six digits to the subject at the rate of 1.2 seconds each. Following a 2-second delay, a single digit is presented and the subjects must respond "yes" or "no" depending on whether or not the test digit was a member of the set just presented. Following this response the subject is required to recall the complete set in order. Since the subjects were 98.7% correct on the recognition test and 98.6% correct on the recall test, it may be assumed that the task was within their rehearsal span. Interesting results were found in the latencies of the recognition responses: there was a linear increase in latency as the set size increased from one to six digits. The fact that there was no difference in latencies for "yes" versus "no" responses indicates that the search process in this situation is exhaustive and does not terminate the moment a match is found. Sternberg concludes that the subject engages in an exhaustive serial comparison process which evaluates elements at the rate of 25 to 30 per second. The high processing rate makes it seem likely that the rehearsal the subjects report is not an integral part of the scanning process, but instead maintains the image in STS so that it may be scanned at the time of the test. This conclusion depends upon accepting as a reasonable rehearsal rate

for digits the values reported by Landauer (1962) which were never higher than six per second.

Buschke's (1963) missing-span method provides additional insight into search and retrieval processes in STS. The missing-span procedure consists of presenting in a random order all but one of a previously specified set of digits; the subject is then asked to report the missing digit. This technique eliminates the output interference associated with the usual digit-span studies in which the entire presented set must be reported. Buschke found that subjects had superior performance on a missing-span task as compared with an identical digit-span task in which all of the presented items were to be reported in any order. A natural hypothesis would explain the difference in performance as being caused by output interference; that is, the multiple recalls in the digit-span procedure produce interference not seen in the single test procedure of the missing span. An alternative explanation would hold that different storage and search strategies were being employed in the two situations. Madsen and Drucker (1966) examined this question by comparing test instructions given just prior to or immediately following each presentation sequence; the instructions specify whether the subject is to report the set of presented digits or simply to report the missing digit. Output interference would imply that the difference between missing-span and digit-span would hold up in both cases. The results showed that the missing-span procedure with prior instructions was superior to both missing-span and digit-span with instructions following presentation; the latter two conditions produced equal results and were superior to digit-span with prior instructions. It seems clear, then, that two storage and search strategies are being used: a missing-span type, and a digit-span type. Prior instructions (specifying the form of the subject's report) lead the subject to use one or the other of these strategies, but instructions following presentation are associated with a mixture of the two strategies. It appeared in this case that the strategies differed in terms of the type of storage during presentation; the digit-span group with prior instructions tended to report their digits in their presentation order, while the digit-span group with instructions after presentation more often reported the digits in their numerical order. This indicates that the missing-span strategy involved checking off the numbers as they were presented against a fixed, numerically ordered list, while the digit-span strategy involved rehearsing the items in their presented order. It is interesting to note that if the subjects had been aware of the superiority of the missing-span strategy, they could have used it in the digit-span task also, since the two types of tests called for the same information.

It should be noted that retrieval from STS depends upon a number of factors, some under the control of the subject and some depending upon

the decay characteristics of STS. If the decay is partial in some sense, so that the trace contains only part of the information necessary for direct output, then the problem arises of how the partial information should be used to generate a response. In this case, it would be expected that the subject would then engage in a search of LTS in an effort to match or recognize the partial information. On the other hand, even though traces may decay in a partial manner, the rehearsal capability can hold a select set of items in a state of immediate recall availability and thereby impart to these items what is essentially an all-or-none status. It is to this rehearsal process that we now turn.

2. *Rehearsal Processes*

Rehearsal is one of the most important factors in experiments on human memory. This is particularly true in the laboratory because the concentrated, often meaningless, memory tasks used increase the relative efficacy of rehearsal as compared with the longer term coding and associative processes. Rehearsal may be less pervasive in everyday memory, but nevertheless has many uses, as Broadbent (1958) and others have pointed out. Such examples as remembering a telephone number or table-tennis score serve to illustrate the primary purpose of rehearsal, the lengthening of the time period information stays in the short-term store. A second purpose of rehearsal is illustrated by the fact that even if one wishes to remember a telephone number permanently, one will often rehearse the number several times. This rehearsal serves the purpose of increasing the strength built up in a long-term store, both by increasing the length of stay in STS (during which time a trace is built up in LTS) and by giving coding and other storage processes time to operate. Indeed, almost any kind of operation on an array of information (such as coding) can be viewed as a form of rehearsal, but this paper reserves the term only for the duration-lengthening repetition process.

In terms of STS structure, we can imagine that each rehearsal regenerates the STS trace and thereby prolongs the decay. This does not imply that the entire information ensemble available in STS immediately after presentation is regenerated and maintained at each rehearsal. Only that information selected by the subject, often a small proportion of the initial ensemble, is maintained. If the word "cow" is presented, for example, the sound of the word cow will enter STS; in addition, associates of cow, like milk, may be retrieved from LTS and also entered in STS; furthermore, an image of a cow may be entered into a short-term visual store. In succeeding rehearsals, however, the subject may rehearse only the word "cow" and the initial associates will decay and be lost. The process may be similar to the loss of meaningfulness that occurs when a word is repeated over and over (Lambert & Jakobovitz, 1960).

An interesting question concerns the maximum number of items that can be maintained via rehearsal. This number will depend upon the rate of STS decay and the form of the trace regenerated in STS by rehearsal. With almost any reasonable assumptions about either of these processes, however, an ordered rehearsal will allow the greatest number of items to be maintained. To give a simple example, suppose that individual items take 1.1 seconds to decay and may be restarted if rehearsal begins before decay is complete. Suppose further that each rehearsal takes .25 seconds. It is then clear that five items may be maintained indefinitely if they are rehearsed in a fixed order over and over. On the other hand, a rehearsal scheme in which items are chosen for rehearsal on a random basis will quickly result in one or more items decaying and becoming lost. It would be expected, therefore, that in situations where subjects are relying primarily upon their rehearsal capability in STS, rehearsal will take place in an ordered fashion. One such situation, from which we can derive an estimate of rehearsal capability, is the digit-span task. A series of numbers is read to the subject who is then required to recall them, usually in the forward or backward order. Because the subject has a long-term store which sometimes can be used to supplement the short-term rehearsal memory, the length of a series which can be correctly recalled may exceed the rehearsal capacity. A lower limit on this capacity can be found by identifying the series length at which a subject never errs; this series length is usually in the range of five to eight numbers.[5]

The above estimates of rehearsal capability are obtained in a discrete-trial situation where the requirement is to remember every item of a small input. A very similar rehearsal strategy can be employed, however, in situations such as free recall where a much greater number of items is input than rehearsal can possibly encompass. One strategy in this case would be to replace one of the items currently being rehearsed by each new item input. In this case every item would receive at least some rehearsal. Because of input and reorganization factors, which undoubtedly consume some time, the rehearsal capacity would probably be reduced. It should be clear that under this scheme a constant number of items will be undergoing rehearsal at any one moment. As an analogy, one might think of a bin always containing exactly n items; each new item enters the bin and knocks out an item already there. This process has been called in earlier reports a "rehearsal buffer," or simply a "buffer," and we will use this terminology here (Atkinson & Shiffrin, 1965).

[5] Wickelgren (1965) has examined rehearsal in the digit-span task in greater detail and found that rehearsal capacity is a function of the groupings engaged in by the subject; in particular, rehearsal in distinct groups of three was superior to rehearsal in four's and five's.

In our view, the maintenance and use of the buffer is a process entirely under the control of the subject. Presumably a buffer is set up and used in an attempt to maximize performance in certain situations. In setting up a maximal-sized buffer, however, the subject is devoting all his effort to rehearsal and not engaging in other processes such as coding and hypothesis testing. In situations, therefore, where coding, long-term search, hypothesis testing, and other mechanisms appreciably improve performance, it is likely that a trade-off will occur in which the buffer size will be reduced and rehearsal may even become somewhat random while coding and other strategies increase.

At this point we want to discuss various buffer operations in greater detail. Figure 2 illustrates a fixed-size buffer and its relation to the rest

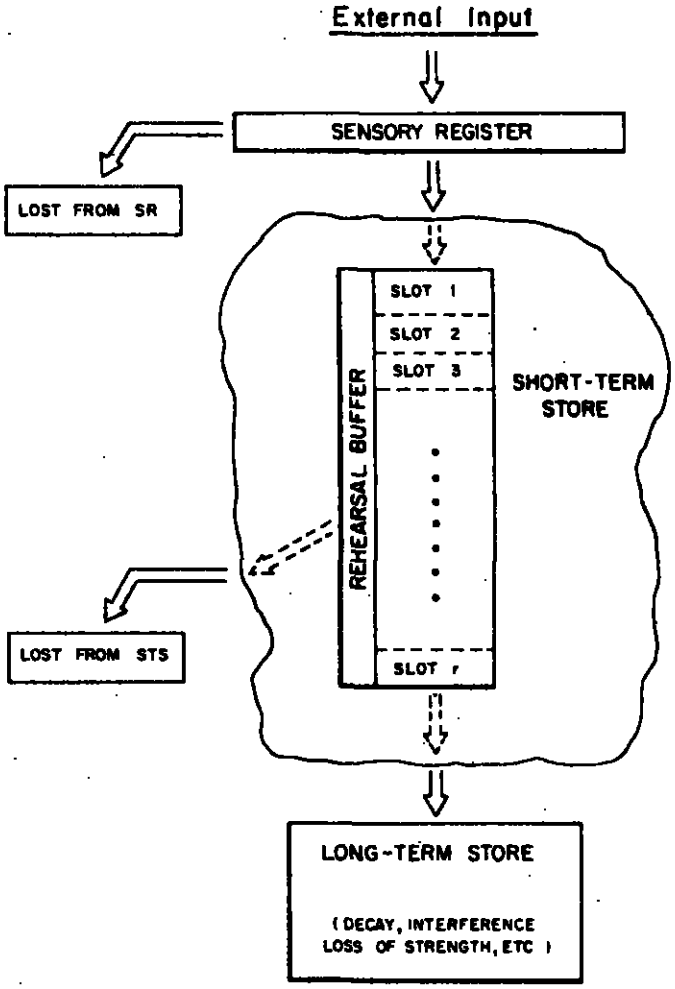


FIG. 2. The rehearsal buffer and its relation to the memory system.

of the memory system. The content of the buffer is constructed from items that have entered STS, items which have been input from the sensory register or from LTS. The arrow going toward LTS indicates that some long-term trace is being built up during an item's stay in the buffer. The other arrow from the buffer indicates that the input of a new item into the buffer causes an item currently in the buffer to be bumped out; this item then decays from STS and is lost (except for any trace which has accumulated in LTS during its stay). An item dropped from the buffer is likely to decay more quickly in STS than a newly presented item which has just entered STS. There are several reasons for this. For one thing, the item is probably already in some state of partial decay when dropped; in addition, the information making up an item in the buffer is likely to be only a partial copy of the ensemble present immediately following stimulus input.

There are two additional processes not shown in Fig. 2 that the subject can use on appropriate occasions. First, the subject may decide not to enter every item into the buffer; the reasons are manifold. For example, the items may be presented at a very fast rate so that input and reorganization time encroach too far upon rehearsal time. Another possibility is that some combinations of items are particularly easy to rehearse, making the subject loath to break up the combination. In fact, the work involved in introducing a new item into the buffer and deleting an old one may alone give the subject incentive to keep the buffer unchanged. Judging from these remarks, the choice of which items to enter into the buffer is based on momentary characteristics of the current string of input items and may appear at times to be essentially random.

The second process not diagrammed in Fig. 2 is the choice of which item to eliminate from the buffer when a new item is entered. There are several possibilities. The choice could be random; it could be based upon the state of decay of the current items; it could depend upon the ease of rehearsing the various items; most important, it could be based upon the length of time the various items have resided in the buffer. It is not unreasonable that the subject knows which items he has been rehearsing the longest, as he might if rehearsal takes place in a fixed order. It is for this reason that the slots or positions of the buffer have been numbered consecutively in Fig. 2; that is, to indicate that the subject might have some notion of the relative recency of the various items in the buffer.

The experimental justification for these various buffer mechanisms will be presented in Section IV. It should be emphasized that the subject will use a fixed-size buffer of the sort described here only in select situations, primarily those in which he feels that trading off rehearsal time for coding and other longer term control processes would not be fruitful. To the extent that long-term storage operations prove to be successful

as compared with rehearsal, the structure of the rehearsal mechanism will tend to become impoverished. One other point concerning the buffer should be noted. While this paper consistently considers a fixed-size short-term buffer as a rehearsal strategy of the subject, it is possible to apply a fixed-size model of a similar kind to the structure of the short-term system as a whole, that is, to consider a short-term buffer as a permanent feature of memory. Waugh and Norman (1965), for example, have done this in their paper on primary memory. The data on the structure of STS is currently so nebulous that such an hypothesis can be neither firmly supported nor rejected.

3. Coding Processes and Transfer between Short- and Long-Term Store

It should be evident that there is a close relationship between the short- and long-term store. In general, information entering STS comes directly from LTS and only indirectly from the sensory register. For example, a visually presented word cannot be entered into STS as an auditory-verbal unit until a long-term search and match has identified the verbal representation of the visual image. For words, letters, and highly familar stimuli, this long-term search and match process may be executed very quickly, but one can imagine unfamiliar stimuli, such as, say, a nonsense scribble, where considerable search might be necessary before a suitable verbal representation is found to enter into STS. In such cases, the subject might enter the visual image directly into his short-term visual memory and not attempt a verbal coding operation.

Transfer from STS to LTS may be considered a permanent feature of memory; any information in STS is transferred to LTS to some degree throughout its stay in the short-term store. The important aspect of this transfer, however, is the wide variance in the amount and form of the transferred information that may be induced by control processes. When the subject is concentrating upon rehearsal, the information transferred would be in a relatively weak state and easily subject to interference. On the other hand, the subject may divert his effort from rehearsal to various coding operations which will increase the strength of the stored information. In answer to the question of what is a coding process, we can most generally state that a coding process is a select alteration and/or addition to the information in the short-term store as the result of a search of the long-term store. This change may take a number of forms, often using strong preexisting associations already in long-term store. A number of these coding possibilities will be considered later.

Experiments may be roughly classified in terms of the control operations the subject will be led to use. Concept formation problems or tasks where there is a clear solution will lead the subject to strategy selection and hypothesis-testing procedures (Restle, 1964). Experiments which

do not involve problem solving, where there are a large number of easily coded items, and where there is a long period between presentation and test, will prompt the subject to expend his efforts on long-term coding operations. Finally, experiments in which memory is required, but long-term memory is not efficacious, will lead the subject to adopt rehearsal strategies that maintain the information the limited period needed for the task. Several examples of the latter experiment will be examined in this paper; they are characterized by the fact that the responses assigned to particular stimuli are continually changing, so that coding of a specific stimulus-response pair will prove harmful to succeeding pairs using the same stimulus. There are experiments, of course, for which it will not be possible to decide on a priori grounds which control processes are being used. In these cases the usual identification procedures must be used, including model fits and careful questioning of the subjects.

There are other short-term processes that do not fit easily into the above classification. They include grouping, organizing, and chunking strategies. One form that organizing may take is the selection of a subset of presented items for special attention, coding and/or rehearsal. This selection process is clearly illustrated in a series of studies on magnitude of reward by Harley (1965a, 1965b). Items in a paired-associate list were given two monetary incentives, one high and one low. In one experiment the subjects learned two paired-associate lists, one consisting of all high incentive items, the other consisting of all low incentive items; there were no differences in the learning rates for these lists. In a second experiment, subjects learned a list which included both high and low incentive items; in this case learning was faster for the high than the low incentive items. However, the overall rate of learning for the mixed list was about the same as for the two previous lists. It seems clear that when the high and low incentive items are mixed, the subject selectively attends to, codes, and rehearses those items with the higher payoffs. A second kind of organizing that occurs is the grouping of items into small sets, often with the object of memorizing the set as a whole, rather than as individual items. Typically in this case the grouped items will have some common factor. A good example may be found in the series of studies by Battig (1966) and his colleagues. He found a tendency to group items according to difficulty and according to degree of prior learning; this tendency was found even in paired-associate tasks where an extensive effort had been made to eliminate any basis for such grouping. A third type of information organization is found in the "chunking" process suggested by Miller (1956). In his view there is some optimal size that a set of information should have in order to best facilitate remembering. The incoming information is therefore organized into chunks of the desired magnitude.

C. Control Processes in Long-Term Store

Control processes to be considered in this section fall roughly into two categories: those concerned with transfer between short-term and long-term store and those concerned with search for and retrieval of information from LTS.

1. *Storage in Long-Term Store*

It was stated earlier that some information is transferred to LTS throughout an item's stay in STS, but that its amount and form is determined by control processes. This proposition will now be examined in greater detail. First of all, it would be helpful to consider a few simple examples where long-term storage is differentially affected by the coding strategy adopted. One example is found in a study on mediators performed by Montague, Adams, and Kiess (1966). Pairs of nonsense syllables were presented to the subject who had to write down any natural language mediator (word, phrase, or sentence associated with a pair) which occurred to him. At test 24 hours later the subject attempted to give the response member of each pair and the natural language mediator (NLM) that had been used in acquisition. Proportion correct for items on which the NLM was retained was 70%, while the proportion correct was negligible for items where the NLM was forgotten or significantly changed. Taken in conjunction with earlier studies showing that a group using NLMs was superior to a group learning by rote (Runquist & Farley, 1964), this result indicates a strong dependence of recall upon natural language mediators. A somewhat different encoding technique has been examined by Clark and Bower (personal communication). Subjects were required to learn several lists of paired-associate items, in which each item was a pair of familiar words. Two groups of subjects were given identical instructions, except for an extra section read to the experimental group explaining that the best method of learning the pairs was to form an elaborate visual image containing the objects designated by the two words. This experimental group was then given a few examples of the technique. There was a marked difference in performance between the groups on both immediate and delayed tests, the experimental group outperforming the control group by better than 40% in terms of probability correct. In fact, postexperimental questioning of the subjects revealed that the occasional high performers in the control group were often using the experimental technique even in the absence of instructions to do so. This technique of associating through the use of visual images is a very old one; it has been described in considerable detail, for example, by Cicero in *De Oratore* when he discusses memory as one of the five parts of rhetoric, and is clearly very effective.

We now consider the question of how these encoding techniques improve performance. The answer depends to a degree upon the fine structure of long-term store, and therefore cannot be stated precisely. Nevertheless, a number of possibilities should be mentioned. First, the encoding may make use of strong preexisting associations, eliminating the necessity of making new ones. Thus in mediating a word pair in a paired-associate task, word *A* might elicit word *A'* which in turn elicits the response. This merely moves the question back a level: how does the subject know which associates are the correct ones? It may be that the appropriate associations are identified by temporal position; that is, the subject may search through the associations looking for one which has been elicited recently. Alternatively, information could be stored with the appropriate association identifying it as having been used in the current paired-associates task. Second, the encoding might greatly decrease the effective area of memory which must be searched at the time of test. A response word not encoded must be in the set of all English words, or perhaps in the set of all words presented "recently," while a code may allow a smaller search through the associates of one or two items. One could use further search-limiting techniques such as restricting the mediator to the same first letter as the stimulus. A third possibility, related to the second, is that encoding might give some order to an otherwise random search. Fourth, encoding might greatly increase the amount of information stored. Finally, and perhaps most important, the encoding might protect a fledgling association from interference by succeeding items. Thus if one encodes a particular pair through an image of, say, a specific room in one's home, it is unlikely that future inputs will have any relation to that image; hence they will not interfere with it. In most cases coding probably works well for all of the above reasons.

There is another possible set of effects of the coding process which should be mentioned here. As background, we need to consider the results of several recent experiments which examine the effect of spacing between study and test in paired-associate learning (Bjork, 1966; Young, 1966). The result of primary interest to us is the decrease in probability correct as the number of other paired-associate items presented between study and test increases. This decrease seems to reach asymptote only after a fairly large number (e.g., 20) of intervening items. There are several possible explanations for this "short-term" effect. Although the effect probably occurs over too great an interval to consider direct decay from STS as an explanation, any of several rehearsal strategies could give rise to an appropriate-looking curve. Since a paired-associate task usually requires coding, a fixed-size rehearsal buffer may not be a reasonable hypothesis, unless the buffer size is fairly small; on the other hand, a variable rehearsal set with semirandomly spaced

rehearsals may be both reasonable and accurate. If, on the other hand, one decides that almost no continuing rehearsal occurs in this task, what other hypotheses are available? One could appeal to retroactive interference but this does little more than name the phenomenon. Greeno (1967) has proposed a coding model which can explain the effect. In his view, the subject may select one of several possible codes at the time of study. In particular, he might select a "permanent" code, which will not be disturbed by any other items or codes in the experiment; if this occurs, the item is said to be learned. On the other hand, a "transitory" code might be selected, one which is disturbed or eliminated as succeeding items are presented. This transitory code will last for a probabilistically determined number of trials before becoming useless or lost. The important point to note here is the fact that a decreasing "*short-term*" effect can occur as a result of solely long-term operations. In experiments emphasizing long-term coding, therefore, the decision concerning which decay process, or combination of decay processes, is operative will not be easy to make in an a priori manner; rather the decision would have to be based upon such a posteriori grounds as goodness-of-fit results for a particular model and introspective reports from the subject.

2. *Long-Term Search Processes*

One of the most fascinating features of memory is the long-term search process. We have all, at one time or another, been asked for information which we once knew, but which is momentarily unavailable, and we are aware of the ensuing period (often lasting for hours) during which memory was searched, occasionally resulting in the correct answer. Nevertheless, there has been a marked lack of experimental work dealing with this rather common phenomenon. For this reason, our discussion of search processes will be primarily theoretical, but the absence of a large experimental literature should not lead us to underestimate the importance of the search mechanism.

The primary component of the search process is locating the sought-for trace (or one of the traces) in long-term store. This process is seen in operation via several examples. The occasionally very long latencies prior to a correct response for well-known information indicates a non-perfect search. A subject reporting that he will think "of it the moment he thinks about something else" indicates a prior fixation on an unsuccessful search procedure. Similarly, the tip-of-the-tongue phenomenon mentioned earlier indicates a failure to find an otherwise very strong trace. We have also observed the following while quizzing a graduate

student on the names of state capitals. The student gave up trying to remember the capital of the state of Washington after pondering for a long time. Later this student quickly identified the capital of Oregon as Salem and then said at once that the capital of Washington was Olympia. When asked how he suddenly remembered, he replied that he had learned the two capitals together. Presumably this information would have been available during the first search if the student had known where to look: namely in conjunction with the capital of Oregon. Such descriptive examples are numerous and serve to indicate that a search can sometimes fail to uncover a very strong trace. One of the decisions the subject must make is when to terminate an unsuccessful search. An important determiner of the length of search is the amount of order imposed during the search; if one is asked to name all the states and does so strictly geographically, one is likely to do better than someone who spews out names in a haphazard fashion. The person naming states in a haphazard fashion will presently encounter in his search for new names those which he has already given; if this occurs repeatedly, the search will be terminated as being unfruitful. The problem of terminating the search is especially acute in the case of recalling a set of items without a good natural ordering. Such a case is found in free-verbal-recall experiments in which a list of words is presented to the subject who must then recall as many as possible. The subject presumably searches along some sort of temporal dimension, a dimension which lets the subject know when he finds a word whether or not it was on the list presented most recently. The temporal ordering is by no means perfect, however, and the search must therefore be carried out with a degree of randomness. This procedure may lead to missing an item which has a fairly strong trace. It has been found in free-verbal-recall experiments, for example, that repeated recall tests on a given list sometimes result in the inclusion on the second test of items left out on the first test. In our own experiments we have even observed intrusions from an earlier list that had not been recalled during the test of that list.

It would be illustrative at this point to consider an experiment carried out by Norma Graham at Stanford University. Subjects were asked to name the capitals of the states. If a correct answer was not given within 5 seconds following presentation of the state name, the subjects were then given a hint and allowed 30 seconds more to search their memory. The hint consisted of either 1, 2, 4, 12, or 24 consecutive letters of the alphabet, one of which was the first letter in the name of the state capital. The probability correct dropped steadily as the hint size increased from 1 to 24 letters. The average response latencies for correct answers, however, showed a different effect; the 1-letter hint was associated with the fastest response time, the 2-letter hint was slower, the 4-letter hint

was slower yet, but the 12- and 24-letter hints were faster than the 4-letter hint. One simple hypothesis that can explain why latencies were slower after the 4-letter hint than after the 12- and 24-letter hints depends upon differing search processes. Suppose the subject in the absence of a hint engages in "normal" search, or N search. When given the first letter, however, we will assume the subject switches to a first letter search, or L search, consisting of a deeper exploration of memory based upon the first letter. This L search might consist of forming possible sounds beginning with the appropriate letter, and matching them against possible city names. When the size of the hint increases, the subject must apply the L search to each of the letters in turn, obviously a time-consuming procedure. In fact, for 12- or 24-letter hints the probability is high that the subject would use up the entire 30-second search period without carrying out an L search on the correct first letter. Clearly a stage is reached, in terms of hint size, where the subject will switch from an L search to N search in order to maximize performance. In the present experiment it seems clear that the switch in strategy occurred between the 4- and 12-letter hints.

In the above experiment there were two search-stopping events, one subject-controlled and the other determined by the 30-second time limit. It is instructive to consider some of the possible subject-controlled stopping rules. One possibility is simply an internal time limit, beyond which the subject decides further search is useless. Related to this would be an event-counter stopping rule that would halt the subject when a fixed number of prespecified events had occurred. The events could be total number of distinct "searches," total number of incorrect traces found, and so on. A third possibility is dependent on a consecutive-events counter. For example, search could be stopped whenever x consecutive searches recovered traces that had been found in previous searches.

It was noted earlier that searches may vary in their apparent orderliness. Since long-term memory is extremely large, any truly random search would invariably be doomed to failure. The search must always be made along some dimension, or on the basis of some available cues. Nevertheless, searches do vary in their degree of order; a letter-by-letter search is highly structured, whereas a free associative search that proceeds from point to point in a seemingly arbitrary manner will be considerably less restrained, even to the point where the same ground may be covered many times. One other possible feature of the search process is not as desirable as the ones previously mentioned. The search itself might prove destructive to the sought-after trace. That is, just as new information transferred to the long-term store might interfere with previous material stored there, the generation of traces during the search might prove to have a similar interfering effect.

A somewhat different perspective on search procedures is obtained by considering the types of experimental tests that typically are used. Sometimes the very nature of the task presumes a specific search procedure. An example is found in the free-verbal-recall task in which the subject must identify a subset of a larger well-learned group of words. A search of smaller scope is made in a paired-associate task; when the set of possible responses is large, the search for the answer is similar to that made in free recall, with a search component and a recognition component to identify the recovered trace as the appropriate one. When the set of responses in a paired-associate task is quite small, the task becomes one of recognition alone: the subject can generate each possible response in order and perform a recognition test on each. The recognition test presumably probes the trace for information identifying it as being from the correct list and being associated with the correct stimulus.

It was said that the primary component of the search process is locating the desired memory trace in LTS. The secondary component is the recovery of the trace once found. It has been more or less assumed for simplicity in the above discussions that the trace is all-or-none. This may not be the case, and the result of a search might be the recovery of a partial trace. Retrieval would then depend either upon correctly guessing the missing information or performing a further search to match the partial trace with known responses. It is possible, therefore, to divide the recovery processes into a search component and retrieval component, both of which must be successfully concluded in order to output the correct response. The two components undoubtedly are correlated in the sense that stronger, more complete traces will both be easier to find and easier to retrieve, having been found.

One final problem of some importance should be mentioned at this time. The effects of trace interference may be quite difficult to separate from those of search failure. Trace interference here refers either to loss of information in the trace due to succeeding inputs or to confusions caused by competition among multiple traces at the moment of test. Search failure refers to an inability to find the trace at all. Thus a decrease in the probability of a correct response as the number of items intervening between study and test increases could be due to trace interference generated by those items. It could also be due to an increased likelihood of failing to find the trace because of the increasing number of items that have to be searched in memory. One way these processes might be separated experimentally would be in a comparison of recognition and recall measures, assuming that a failure to find the trace is less likely in the case of recognition than in the case of recall. At the present, research along these lines has not given us a definitive answer to this question.

IV. Experiments Concerned with Short-Term Processes

Sections II and III of this paper have outlined a theoretical framework for human memory. As we have seen, the framework is extremely general, and there are many alternative choices that can be made in formulating models for particular experimental situations. The many choice points make it impossible for us to examine each process experimentally. Instead we shall devote our attention to a number of processes universally agreed to occur in experiments on memory, namely rehearsal and search processes. In Section V the LTS search processes will be examined in detail; in the present section the major emphasis will be on STS mechanisms, particularly the control process designated as the rehearsal buffer. The sensory registration system is not an important factor in these models; the experiments are designed so that all items enter the sensory register and then are transferred to STS. The long-term store will be presented in the models of this section but only in the simplest possible manner. We now turn to a series of experiments designed to establish in some detail the workings of the buffer mechanism.

A. A Continuous Paired-Associate Memory Task (Experiment 1)

This study is the prototype for a series of experiments reported in this section designed specifically to study buffer processes. The buffer is a fixed-size rehearsal scheme in STS; conditions which prompt the subject to make use of a buffer include difficulty in using long-term store, a large number of short study-test intervals, and a presentation rate slow enough that cognitive manipulations in STS are not excessively rushed. The task that was developed to establish these conditions is described below.[6]

The subject was required to keep track of constantly changing responses associated with a fixed set of stimuli.[7] The stimuli were 2-digit numbers chosen from the set 00–99; the responses were letters of the alphabet. At the start of a particular subject-session a set of s stimuli was chosen randomly from the numbers 00 to 99; these stimuli were not changed over the course of that day's session. To begin the session each stimulus was paired with a letter chosen randomly from the alphabet. Following this initial period, a continuous sequence of trials made up the rest of the session, each trial consisting of a test phase followed by a

[6] The reader may consult Atkinson, Brelsford, and Shiffrin (1967) for details of the experimental procedure and theoretical analyses that are not covered in the present discussion. Also presented there is an account of the mathematics of the model.

[7] The task is similar to those used by Yntema and Mueser (1960, 1962), Brelsford *et al.* (1966), and Katz (1966).

study phase. During the test phase, one of the s stimuli was randomly selected and presented alone for test. The subject was required to respond with the most recent response paired with that stimulus. No feedback was given to the subject. Following his response the study portion of the trial began. During the study portion the stimulus just presented for test was paired with a new response selected randomly from the alphabet; the only restriction was that the previous response (the correct response during the immediately preceding test phase) was not used during the study phase of the same trial. The subject was instructed to forget the previous pairing and try to remember the new pairing currently being presented for study. Following the study period, a stimulus was again selected randomly from the set of s stimuli and the test portion of the next trial began.

The result of this procedure is as follows: a particular stimulus-response pair is presented for study, followed by a randomly determined number of trials involving other stimuli, and then tested. Having been tested, the pair is broken up and the stimulus is paired with a different response; in other words, no stimulus-response pair is presented for study twice in succession. It is easy to imagine the effects of this procedure on the subject's long-term memory processes. If any particular pair is strongly stored in long-term memory, it will interfere with subsequent pairings involving that same stimulus. In addition, the nature of the stimuli and responses used makes coding a difficult task. For these reasons, the subject soon learns that the usual long-term storage operations, such as coding, are not particularly useful; in fact, the subject is forced to rely heavily on his short-term store and his rehearsal capacity. The experimental procedure also was designed so that it would be possible to carry out extensive parametric analyses on data from individual subjects. This was accomplished by running each subject for 12 or more days and collecting the data on a system under the control of a time-sharing computer, a procedure which made the precise sequence of events during each session available for analysis.

1. *Method*

The subjects were nine students from Stanford University who received $2 per experimental session. This experiment, and most of the others reported in this paper, was conducted in the Computer-Based Learning Laboratory at Stanford University. The control functions were performed by computer programs run on a modified PDP-1 computer manufactured by the Digital Equipment Corp., and under control of a time-sharing system. The subject was seated at a cathode-ray-tube display terminal; there were six terminals, each located in a separate 7 × 8 foot sound-shielded room. Stimuli were displayed on the face of

the cathode ray tube (CRT); responses were made on an electric typewriter keyboard located immediately below the lower edge of the CRT.

For each session the subject was assigned to one of the three experimental conditions. The three conditions were defined in terms of s, the size of the set of stimuli to be remembered, which took on the values 4, 6, or 8. An attempt was made to assign subjects to each condition once in consecutive three-session blocks. Every session began with a series of study trials: one study trial for each stimulus to be used in the session. On a study trial the word "study" appeared on the upper face of the CRT. Beneath the word "study" one of the stimuli (a 2-digit number) appeared along with a randomly selected letter from the alphabet. Subjects were instructed to try to remember the stimulus-response pairs. Each of these initial study trials lasted for 3 seconds with a 3-second intertrial interval. As soon as there had been an initial study trial for each stimulus to be used in the session, the session proper began.

Each subsequent trial involved a fixed series of events. (1) The word "test" appeared on the upper face of the CRT. Beneath the word "test" a randomly selected member of the stimulus set appeared. Subjects were instructed that when the word "test" and a stimulus appeared on the CRT, they were to respond with the last response that had been associated with that simulus, guessing if necessary. This test portion of a trial lasted for 3 seconds. (2) The CRT was blacked out for 2 seconds. (3) The word "study" appeared on the upper face of the CRT for 3 seconds. Below the word "study" a stimulus-response pair appeared. The stimulus was the same one used in the preceding test portion of the trial. The response was randomly selected from the letters of the alphabet, with the stipulation that it be different from the immediately preceding response assigned to that stimulus. (4) There was a 3-second intertrial interval before the next trial. Thus a complete trial (test plus study) took 11 seconds. A subject was run for 220 such trials during each experimental session.

2. *Theoretical Analysis*

In order that the reader may visualize the sequence of events which occurs in this situation, a sample sequence of 18 trials is illustrated in Fig. 3. Within the boxes are the displays seen on the CRT screen. In this session the stimulus set includes the four stimuli 20, 31, 42, and 53 (i.e., $s = 4$). On trial n, item 31-Q is presented for study. On trial $n+1$, 42 is tested and 42-B presented for study. Then on trial $n+2$, 31 is tested; the correct answer is Q as is seen by referring to trial n. After the subject answers he is given 31-S to study. He is instructed to forget the previous pair, 31 Q, and remember only the new pair, 31-S. The response letter S was selected randomly from the alphabet, with the restriction that the

previous response, Q, could not be used. A previously used response may through chance, however, be chosen again later in the session; for example, on trial $n+7$, 31-Q is again presented for study. It is also possible that two or more stimuli might be paired with the same response concurrently; as an example, on trial $n+15$, 20 is paired with C and on trial $n+16$, 42 also is paired with C. The stimulus presented on each trial is chosen randomly; for this reason the number of trials intervening

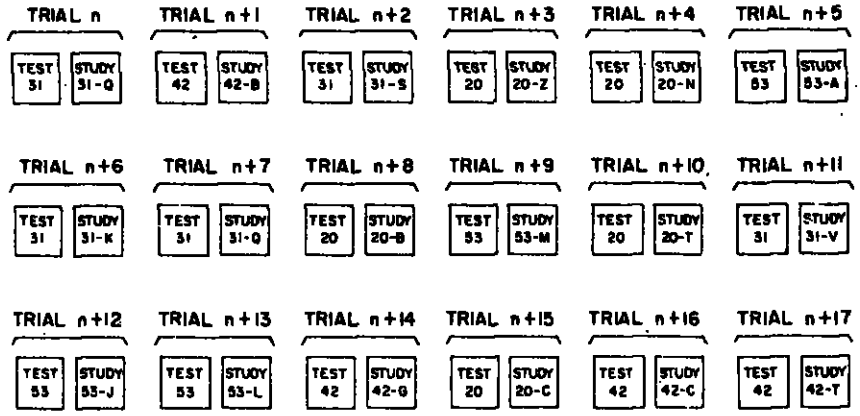


FIG. 3. A sample sequence of trials for Experiment 1.

between study and test is a random variable distributed geometrically. In the analysis of the results, a very important variable is the number of trials intervening between study and test on a particular stimulus-response pair; this variable is called the *lag*. Thus 20 is tested on trial $n+4$ at a lag of 0 because it was studied on trial $n+3$. On the other hand, 42 is tested on trial $n+14$ at a lag of 12, because it was last studied on trial $n+1$.

Consider now the processes the subject will tend to adopt in this situation. The obvious difficulties involved in the use of LTS force the subject to rely heavily upon rehearsal mechanisms in STS for optimal performance.[8] A strategy making effective use of STS is an ordered rehearsal scheme of fixed size called the buffer in Section III,B. The fixed-size requirement may not be necessary for maximal utilization of

[8] The usual examples given for the usefulness of a distinct short-term store do not stress the positive benefits of a memory decaying quickly and completely. Without such a memory, many minor tasks such as adding a long column of numbers might become far more difficult. The current experiment, in which associative bonds are frequently broken and re-formed, is an example of a class of operations for which a short-term store is almost essential.

STS, but is indicated by the following considerations. Keeping the size of the rehearsal set constant gives the subject a great deal of control over the situation; each rehearsal cycle will take about the same amount of time, and it is easier to reorganize the buffer when a new item is introduced. Furthermore, an attempt to stretch the rehearsal capacity to its limit may result in confusion which causes the entire rehearsal set to be disrupted; the confusion results from the variable time that must be allowed for operations such as responding at the keyboard and processing the new incoming items. The hypothesis of an ordered fixed-size buffer is given support by the subjects' reports and the authors' observations while acting as subjects. The reader is not asked, however, to take our word on this matter; the analysis of the results will provide the strongest support for the hypothesis.

It must be decided next just what is being rehearsed. The obvious candidate, and the one reported by subjects, is the stimulus-response pair to be remembered. That is, the unit of rehearsal is the two-digit stimulus number plus the associated response letter. Under certain conditions, however, the subject may adopt a more optimal strategy in which only the responses are rehearsed. This strategy will clearly be more effective because many more items may be encompassed with the same rehearsal effort. The strategy depends upon ordering the stimuli (usually in numerical order in the present case) and rehearsing the responses in an order corresponding to the stimulus order; in this way the subject may keep track of which response goes with which stimulus. For a number of reasons, the scheme is most effective when the size of the stimulus set is small; for a large set the subject may have difficulty ordering the stimuli, and difficulty reorganizing the rehearsal as each new item is presented. When the number of stimulus-response pairs to be remembered is large, the subject may alter this scheme in order to make it feasible. The alteration might consist of rehearsing only the responses associated with a portion of the ordered stimuli. In a previous experiment (Brelsford *et al.*, 1966) with a similar design, several subjects reported using such a strategy when the stimulus set size was four, and an examination of their results showed better performance than the other subjects. Subject reports lead us to believe that this strategy is used infrequently in the present experiment; consequently, our model assumes that the unit of rehearsal is the stimulus-response pair, henceforth called an "item."

Figure 2 outlines the structure of the model to be applied to the data. Despite the emphasis on rehearsal, a small amount of long-term storage occurs during the period that an item resides in the buffer. The information stored in LTS is comparatively weak and decays rapidly as succeeding items are presented. In accord with the argument that the long-term

process is uncomplicated, we assume here that information stored in LTS increases linearly with the time an item resides in the buffer. Once an item leaves the buffer, the LTS trace is assumed to decrease as each succeeding item is presented for study.

Every item is assumed to enter first the sensory register and then STS. At that point the subject must decide whether or not to place the new item in the rehearsal buffer. There are a number of reasons why every incoming item may not be placed in the buffer. For one thing, the effort involved in reorganizing the buffer on every trial may not always appear worthwhile, especially when the gains from doing so are not immediately evident; for another, the buffer at some particular time may consist of a combination of items especially easy to rehearse and the subject may not wish to destroy the combination. In order to be more specific about which items enter the buffer and which do not, two kinds of items must be distinguished. An O item is an incoming stimulus-response pair whose stimulus is currently in the buffer. Thus if 52-L is currently in the buffer, 52 is tested, and 52-G is presented for study, then 52-G is said to be an O item. Whenever an O item is presented it is automatically entered into the buffer; this entry, of course, involves replacing the old response by the appropriate new response. Indeed, if an O item did not enter the buffer, the subject would be forced to rehearse the now incorrect previous response, or to leave a useless blank spot in the buffer; for these reasons, the assumption that O items are always entered into the buffer seems reasonable. The other kind of item that may be presented is an N item. An N item is a stimulus-response pair whose stimulus currently is not in the buffer. Whenever an N item is entered into the buffer, one item currently in the buffer must be removed to make room the new item (i.e., the buffer is assumed to be of fixed size, r, meaning that the number of items being rehearsed at any one time is constant). The assumption is made that an N item enters into the buffer with probability α; whenever an N item is entered, one of the items currently in the buffer is randomly selected and removed to make room for it.

The model used to describe the present experiment is now almost complete. A factor still not specified is the response rule. At the moment of test any item which is in the buffer is responded to correctly. If the stimulus tested is not in the buffer, a search is carried out in LTS with the hope of finding the trace. The probability of retrieving the correct response from LTS depends upon the current trace strength, which in turn, depends on the amount of information transferred to LTS. Specifically we assume that information is transferred to LTS at a constant rate θ during the entire period an item resides in the buffer; θ is the transfer rate per trial. Thus, if an item remains in the rehearsal

buffer for exactly j trials, then that item accumulated an amount of information equal to $j\theta$. We also assume that each trial following the trial on which an item is knocked out of the buffer causes the information stored in LTS for that item to decrease by a constant proportion τ. Thus, if an item were knocked out of the buffer at trial j, and i trials intervened between the original study and test on that item, then the amount of information in LTS at the time of the test would be $j\theta\tau^{i-j}$. We now want to specify the probability of a correct retrieval of an item from LTS. If the amount of information in LTS at the moment of test is zero, then the probability of a correct retrieval should be at the guessing level. As the amount of information increases, the probability of a correct retrieval should increase toward unity. We define ρ_{ij} as the probability of a correct response from LTS for an item that was tested at lag i, and resided in the buffer for exactly j trials. Considering the above specifications on the retrieval process,

$$\rho_{ij} = 1 - (1 - g)\exp[-j\theta(\tau^{i-j})]$$

where g is the guessing probability, which is 1/26 since there were 26 response alternatives.[9]

The basic dependent variable in the present experiment is the probability of a correct response at the time of a test, given lag i. In order to derive this probability we need to know the length of time that an item resides in the memory buffer. Therefore, define $\beta_j =$ probability that an item resides in the buffer for exactly j trials, given that it is tested at a lag greater than j. The probability of a correct response to an item tested at lag i can now be written in terms of the β_j's. Let "C_i" represent the occurrence of a correct response to an item tested at lag i. Then

$$\Pr(C_i) = \left[1 - \sum_{k=0}^{i} \beta_k\right] + \left[\sum_{k=0}^{i} \beta_k \rho_{ik}\right].$$

The first bracketed term is the probability that the item is in the buffer at the time of the test. The second bracket contains a sum of probabilities, each term representing the probability of a correct retrieval

[9] Lest the use of an exponential function seem entirely arbitrary, it should be noted that this function bears a close relation to the familiar linear model of learning theory. If we ignore for the moment the decay feature, then

$$\rho_{ij} = 1 - (1 - g)\exp(-j\theta).$$

It is easily seen that this is the linear model expression for the probability of a correct response after j reinforcements with parameter e^{θ}. Thus, the retrieval function ρ_{ij} can be viewed as a linear model with time in the buffer as the independent variable. To be sure, the decay process complicates matters, but the reason for choosing the exponential function becomes somewhat less arbitrary. A decay process is needed so that the probability of a correct retrieval from LTS will approach chance as the lag tends toward infinity.

from LTS of an item which remained in the buffer for exactly k trials and was then lost.[10] There are four parameters in the model: r, the buffer size which must be an integer; α, the probability of entering an N item into the buffer; θ, the transfer rate of information to LTS; and τ, the decay rate of information from LTS after an item has left the buffer.

One final process must be considered before the model is complete. This process is the recovery of information from STS which is not in the buffer. It will be assumed that the decay of an item which has entered and then left the buffer is very rapid, so rapid that an item which has left the buffer cannot be recovered from STS on the succeeding test.[11] The only time in which a recovery is made from STS, apart from the buffer, occurs if an item is tested immediately following its study (i.e., at a lag of 0). In this case there is virtually no time between study and test and it is assumed therefore that the recovery probability is one, regardless of whether the item was entered into the buffer or not. In other words, the probability correct is one when the lag is zero.

3. *Data Analysis*

Figure 4 presents the probability of a correct response as a function of lag for each of the three stimulus set sizes examined. It can be seen that the smaller the stimulus set size, the better the overall performance. It is important to note that the theory predicts such a difference on the following basis: the larger the size of the stimulus set, the more often an N item will be presented; and the more often N items will be presented, the more often items in the buffer will be knocked out. Recall that only N items can knock items from the buffer; O items merely replace themselves.

It can be seen that performance is almost perfect for lag 0 in all three conditions. This was expected because lag 0 means that the item was tested immediately following its study, and was therefore available in STS. The curves drop sharply at first and slowly thereafter, but have not yet reached the chance level at lag 17, the largest lag plotted. The chance level should be 1/26 since there were 26 response alternatives.

The four parameters of the model were estimated by fitting the model to the lag curves in Fig. 4 using a minimum chi-square as a best fit

[10] One factor which the model as outlined ignores is the probability of recovering from LTS an old, incorrect trace. In the interest of simplicity this process has not been introduced into the model, although it could be appended with no major changes.

[11] Clearly this assumption depends on the time intervals involved. In the present experiment the trials were quite slow; in experiments where a faster presentation rate is used, the model probably would need to be modified slightly to allow a nonzero probability of recovery of an item from STS on the test following its removal from the buffer.

criterion.[12] The solid lines in Fig. 5 give the best fit of the model, which occurred when the parameter values were: $r = 2$, $\alpha = .39$, $\theta = .40$, and $\tau = .93$. It can be seen that the observed data and the predictions from the model are in close agreement. It should be emphasized that the three curves are fit simultaneously using the same parameter values, and the differences between the curves depend only on the value of s (the stimulus

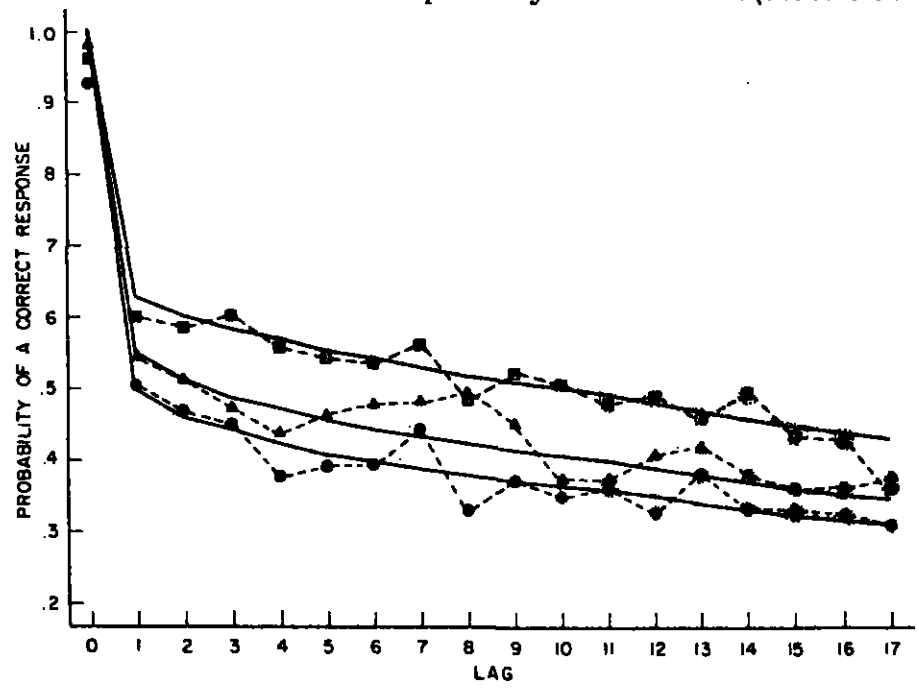


FIG. 4. Observed and theoretical probabilities of a correct response as a function of lag (Experiment 1). --■-- $s = 4$; --▲-- $s = 6$; --●-- $s = 8$;—theory.

set size) which, of course, is determined by the experimenter. The predicted probabilities of a correct response weighted and summed over all lag positions are .562, .469, and .426 for s equal to 4, 6, and 8, respectively; the observed values are .548, .472, and .421.

The estimated value of r might seem surprising at first glance; two items appear to be a rather small buffer capacity. But there are a number of considerations that render this estimate reasonable. It seems clear that the capacity estimated in a task where the subject is constantly interrupted for tests must be lower than the capacity estimated, for example, in a typical digit-span task. This is so because part of the attention time that would be otherwise allotted to rehearsal must be used to search memory in order to respond to the continuous sequence

[12] See Atkinson, Brelsford, and Shiffrin (1967) for details of the estimation procedure and a statistical evaluation of the goodness-of-fit.

of tests. Considering that two items in this situation consist of four numbers and two letters, an estimate of r equal to two is not particularly surprising. The estimated value of α indicates that only 39% of the N items actually enter the buffer (remember that O items always enter the buffer). This low value may indicate that a good deal of mental effort is involved in keeping an item in the buffer via rehearsal, leading to a reluctance to discard an item from the buffer that has not yet been tested. A similar reluctance to discard items would be found if certain

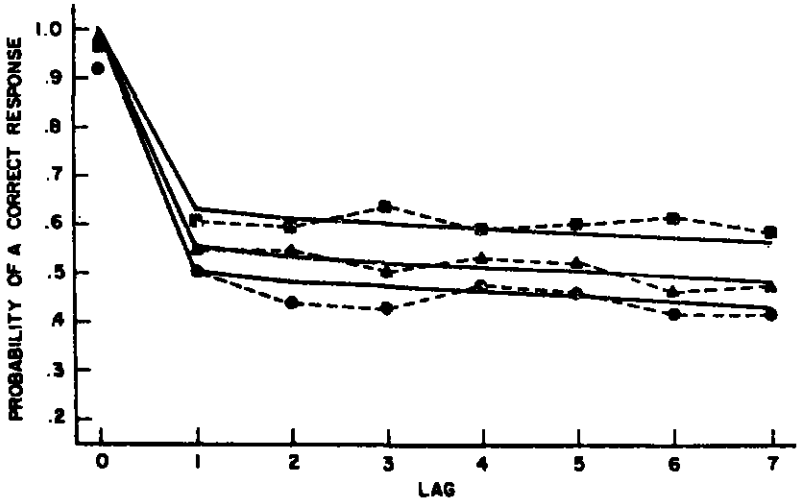


FIG. 5. Observed and theoretical probabilities of a correct response as a function of lag when every intervening item uses the same stimulus (Experiment 1). --■-- $s = 4$; --▲-- $s = 6$; --●-- $s = 8$;—theory.

combinations of items were particularly easy to rehearse. Finally, note that the theory predicts that, if there were no long-term storage, the subject's overall probability of a correct response would be independent of α. Thus it might be expected that α would be higher the greater the effectiveness of long-term storage. In accord with this reasoning, the low value of α found would result from the weak long-term storage associated with the present situation.

In addition to the lag curves in Fig. 4, there are a number of other predictions that can be examined. One aspect of the theory maintains that O items always enter the buffer and replace themselves, while N items enter the buffer with probability α and knock an item out of the buffer whenever they do so. The effects of different stimulus-set sizes displayed in Fig. 5 are due to this assumption. The assumption, however, may be examined in other ways; if it is true, then an item's probability of being correct will be affected by the specific items that intervene

between its initial study and its later test. If every intervening trial uses the same stimulus, then the probability of knocking the item of interest from the buffer is minimized. This is so because once any intervening item enters the buffer, every succeeding intervening item is an O item (since it uses the same stimulus), and hence also enters the buffer. Indeed, if α were one, then every intervening item after the first would be an O

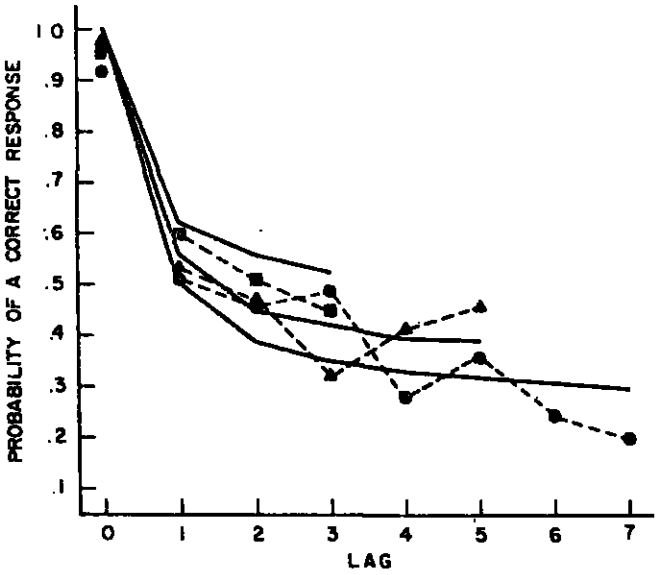


FIG. 6. Observed and theoretical probabilities of a correct response as a function of lag when every intervening item uses a different stimulus (Experiment 1). --■-- $s = 4$; --▲-- $s = 6$; --●-- $s = 8$;— theory.

item, and hence only the first intervening item would have a chance of knocking the item of interest from the buffer; if $\alpha = 1$ and there were no long-term decay, then the lag curve for this condition would be flat from lag 1 onward. In this case, however, α is not equal to one and there is long-term decay; hence the lag curve will decrease somewhat when the intervening items all have the same stimulus, but not to the extent found in Fig. 4. This lag curve, called the "all-same" curve, is shown in Fig. 5; it plots the probability of a correct response as a function of lag, when all the intervening trials between study and test involve the same stimulus. The parameters previously estimated were used to generate predictions for these curves and they are displayed as solid lines. It seems clear that the predictions are highly accurate.

A converse result, called the "all-different" lag curve, is shown in Fig. 6. In this condition, every intervening item has a different stimulus,

and therefore the probability of knocking the item of interest from the buffer is maximized. The lag curves for this condition, therefore, should drop faster than the unconditional lag curves of Fig. 4. Predictions were again generated using the previous parameter values and are represented by the solid lines in Fig. 6. Relatively few observations were available in this condition; considering the instability of the data the predictions seem reasonable.

The procedure used in this experiment is an excellent example of what has been traditionally called a negative transfer paradigm. The problems inherent in such a paradigm were mentioned earlier as contributing to the subjects' heavy reliance upon the short-term store. To the extent that there is any use of LTS, however, we would expect intrusion errors from previously correct responses. The model could be extended in several obvious ways to predict the occurrence of such intrusions. For example, the subject could, upon failing to recover the most recent trace from LTS, continue his search and find the remains of the previous, now incorrect, trace. In order to examine intrusion errors, the proportion of errors which were the correct response for the previous presentation of the stimulus in question were calculated for each lag and each condition. The proportions were quite stable over lags with mean values of .065, .068, and .073 for the 4, 6, and 8 stimulus conditions, respectively. If the previously correct response to an item is generated randomly for any given error, these values should not differ significantly from $1/25 = .04$. In both the $s = 4$ and $s = 6$ conditions seven of the nine subjects had mean values above chance; in the $s = 8$ condition eight of the nine subjects were above chance. Intrusion errors may therefore be considered a reliable phenomenon in this situation; on the other hand, the relatively low frequency with which they occur indicates a rather weak and quickly decaying long-term trace.

A second error category of interest includes those responses that are members of the current set of responses to be remembered but are not the correct responses. This set, of course, includes the set of responses in the buffer at any one time; if the subject tends to give as a guess a response currently in the buffer (and therefore highly available), then the probability of giving as an error a response in the current to-be-remembered set will be higher than chance. Since responses may be assigned to more than one stimulus simultaneously, the number of responses in the to-be-remembered set is bound by, but may be less than, the size of the stimulus set, s. Thus, on the basis of chance the error probabilities would be bounded below.12, .20, and .28 for $s = 4$, 6, and 8, respectively. The actual values found were .23, .28, and .35, respectively. This finding suggests that when the subject cannot retrieve the response from his buffer or LTS and is forced to guess, he has a somewhat greater

than chance likelihood of giving a response currently in the rehearsal set but assigned to another stimulus. It is not surprising that a subject will give as a guess one of the responses in his buffer since they are immediately available.

Other analyses have been performed on the data of this experiment, but the results will not be presented until a second experiment has been described. Before considering the second experiment, however, a few words should be said about individual differences. One of the reasons for running a single subject for many sessions was the expectation that the model could be applied to each subject's data separately. Such analyses have been made and are reported elsewhere (Atkinson, Brelsford & Shiffrin, 1967). The results are too complex to go into here, but they establish that individual subjects by and large conform to the predictions of the model quite well. Since our aim in this paper is to present a nontechnical discussion of the model, to simplify matters we will make most of our analyses on group data.

B. The "All-Different" Stimulus Procedure (Experiment 2)

In the preceding experiment, the number of stimuli used in a given experimental session and the size of the to-be-remembered set were identical. These two factors, however, can be made independent. Specifically, a set of all-different stimuli could be used while keeping the size of the to-be-remembered set constant. The name, all-different, for this experiment results from the use of all-different stimuli, i.e., once a given stimulus-response pair is presented for test, that stimulus is not used again. In other respects the experiment is identical to Experiment 1.

One reason for carrying out an experiment of this type is to gain some information about the replacement hypothesis for O items. In Experiment 1 we assumed that a new item with a stimulus the same as an item currently in the buffer automatically replaced that item in the buffer; that is, the response switched from old to new. In the all-different experiment subjects are instructed, as in Experiment 1, to forget each item once it has been tested. If an item currently in the buffer is tested (say, 52-G) and a new item is then presented for study (say, 65-Q), we might ask whether the tested item will be automatically replaced by the new item (whether 65-Q will replace 52-G in the buffer). This replacement strategy is clearly optimal for it does no good to retain an item in the buffer that already has been tested. Nevertheless, if the reorganization of the buffer is difficult and time consuming, then the replacement of a tested item currently in the buffer might not be carried out. One simple assumption along these lines would postulate that every item has an independent probability α of entering the buffer.

The all-different experiment was identical to Experiment 1 in all

respects except the following. In Experiment 1 the s stimuli were the same throughout an experimental session, with only the associated responses being changed on each trial, whereas in the all-different experiment 100 stimuli were available for use in each session. In fact, every stimulus was effectively new since the stimulus for each study trial was selected randomly from the set of all 100 stimuli under the restriction that no stimulus could be used if it had been tested or studied in the previous 50 trials. There were still three experimental conditions with s equal to 4, 6, or 8 denoting the number of items that the subject was required to try to remember at any point in time. Thus a session began with either 4, 6, or 8 study trials on different randomly selected stimuli, each of which was paired with a randomly selected response (from the 26 letters). On each trial a stimulus in the current to-be-remembered set was presented for test. After the subject made his response he was instructed to forget the item he had just been tested on, since he would not be tested on it again. Following the test a new stimulus was selected (one that had not appeared for at least 50 trials) and randomly paired with a response for the subject to study. Thus the number of items to be remembered at any one time stays constant throughout the session. However, the procedure is quite different from Experiment 1 where the study stimulus was always the one just tested.

Denote an item presented for study on a trial as an O item (old item) if the item just tested was in the buffer. Denote an item presented for study as an N item (new item) if the item just tested was not in the buffer. This terminology conforms precisely to that used to describe Experiment 1. If an O item is presented there will be at least one spot in the buffer occupied by a useless item (the one just tested). If an N item is presented, the buffer will be filled with information of the same value as that before the test. If we assume that an N item has probability α of entering the buffer, and that an O item will always enter the buffer and knock out the item just made useless, then the model for Experiment 1 will apply here with no change whatsoever. In this case we again expect that the lag curves for $s = 4$, 6, and 8 would be separated. In fact, given the same parameter values, exactly the same curves would be predicted for the all-different experiment as for Experiment 1.

As noted earlier, however, there is some doubt that the assumptions regarding N items and O items will still hold for the all-different experiment. In Experiment 1 the stimulus just tested was re-paired with a new response, virtually forcing the subject to replace the old response with a new one if the item was in the buffer. Put another way, if an item is in the buffer when tested, only a minor change need be made in the buffer to enter the succeeding study item: a single response is replaced by another. In the all-different experiment, however, a greater change

needs to be made in order to enter an O item; both a stimulus and a response member have to be replaced. Thus an alternative hypothesis might maintain that every entering item (whether an N item or an O item) has the same probability α of entering the buffer, and will knock out any item currently in the buffer with equal likelihood. In this case we predict no differences among the lag curves for the $s = 4$, 6, and 8 conditions.

1. *Results*

The observed lag curves for Experiment 2 are displayed in Fig. 7. It should be emphasized that, except for the procedural changes described

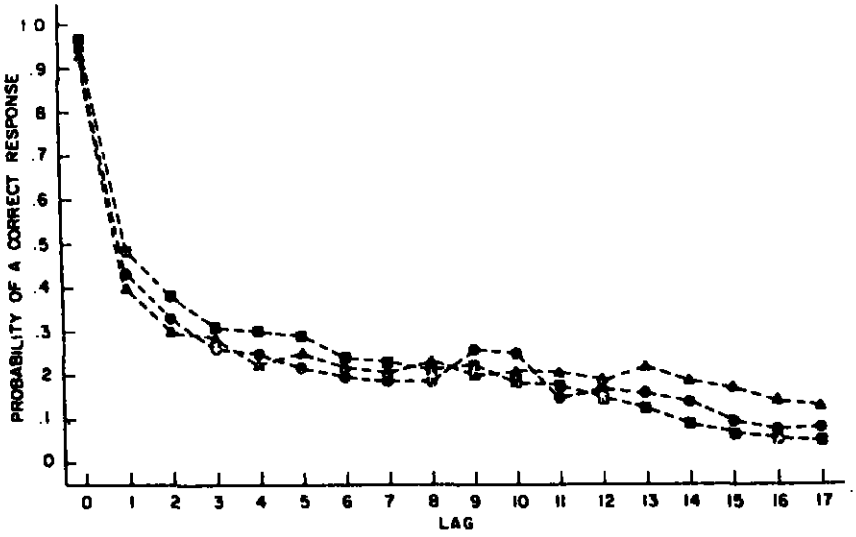


FIG. 7. Observed and theoretical probabilities of a correct response as a function of lag (Experiment 2). --■-- $s = 4$; --▲-- $s = 6$; --●-- $s = 8$.

above and the fact that a new sample of subjects was used, the experimental conditions and operations were identical in Experiments 1 and 2. The important point about this data is that the lag curves for the three conditions appear to overlap.[13] For this reason we lump the three curves to form the single lag curve displayed in Fig. 8.

Because the three curves overlap, it is apparent that the theory used in Experiment 1 needs modification. The hypothesis suggested above

[13] To determine whether the three curves in Fig. 7 differ reliably, the proportions correct for each subject and condition were calculated and then ranked. An analysis of variance for correlated means did not yield significant effects ($F = 2.67$, $df = 2/16$, $p > .05$).

will be used: every item enters the buffer with probability α. If an item enters the buffer it knocks out an item already there on a random basis. This model implies that useless items are being rehearsed on occasion, and subjects reported doing just that despite instructions to forget each item once tested.

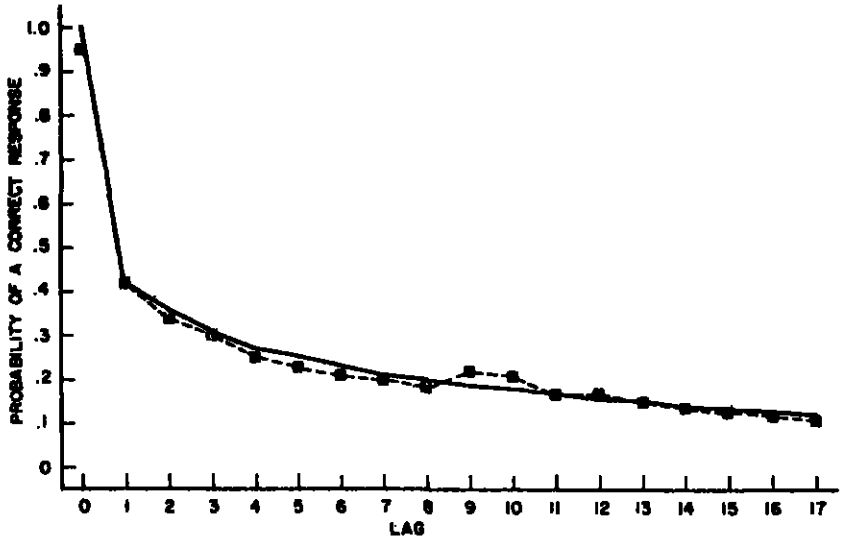


FIG. 8. Observed and theoretical probabilities of a correct response as a function of lag. Data from the $s = 4$, 6, and 8 conditions have been pooled (Experiment 2). --■-- Data;—theory.

The curve in Fig. 8 was fit using a minimum χ^2 procedure; the parameter estimates were $r = 2$, $\alpha = .52$, $\theta = .17$, and $\tau = .90$. It can be seen that the fit is excellent. Except for r, the parameters differ somewhat from those found in Experiment 1, primarily in a slower transfer rate, θ. In Experiment 1 the estimate of θ was .40. This reduction in long-term storage is not too surprising since the subjects were on occasion rehearsing useless information. It could have been argued in advance of the data that the change away from a strong "negative-transfer" paradigm in Experiment 2 would lead to increased use of LTS; that this did not occur is indicated not only by the low θ value, but also by the low probability of a correct response at long lags. One outcome of this result is the possibility that the all-different procedure would give superior long-term memory in situations where subjects could be induced to attempt coding or other long-term storage strategies. It seems apparent that LTS was comparatively useless in the present situation.

2. *Some Statistics Comparing Experiments 1 and 2*

In terms of the model, the only difference between Experiments 1 and 2 lies in the replacement assumption governing the buffer. In Experiment 1, an item in the buffer when tested is automatically replaced by the immediately succeeding study item; if the tested item is not in the buffer, the succeeding study item enters the buffer with probability α, randomly displacing an item already there. In Experiment 2, every study item, independent of the contents of the buffer, enters the buffer with probability α, randomly displacing an item already there. While these assumptions are given credence by the predictions of the various lag curves of Figs. 4 and 8, there are other statistics that can be examined to evaluate their adequacy. These statistics depend upon the fact that items vary in their probability of entering the buffer. Since items which enter the buffer will have a higher probability correct than items which do not, it is relatively easy to check the veracity of the replacement assumptions in the two experiments.

In Experiment 1, the probability that an item will be in the buffer at test is higher the greater the number of consecutive preceding trials that involve the same stimulus. Thus if the study of 42-B is preceded, for example, by six consecutive trials using stimulus 42, there is a very high probability that 42-B will enter the buffer. This occurs because there is a high probability that the stimulus 42 already will be in the buffer when 42-B is presented, and if so, then 42-B will automatically enter the buffer. In any series of consecutive trials all with the same stimulus, once any item in the series enters the buffer, every succeeding item will enter the buffer. Hence, the longer the series of items with the same stimulus, the higher the probability that that stimulus will be in the buffer. Figure 9 graphs the probability of a correct response to the last stimulus-response pair studied in a series of consecutive trials involving the same stimulus; the probability correct is lumped over all possible lags at which that stimulus-response pair is subsequently tested. This probability is graphed as a function of the length of the consecutive run of trials with the same stimulus and is the line labeled Experiment 1. These curves are combined over the three experimental conditions (i.e., $s = 4, 6, 8$). We see that the probability of a correct response to the last item studied in a series of trials all involving the same stimulus increases as the length of that series increases, as predicted by the theory.

In Experiment 2 stimuli are not repeated, so the above statistic cannot be examined. A comparable statistic exists, however, if we consider a sequence of items all of which are tested at zero lag (i.e., tested immediately after presentation). One could hypothesize that the effect displayed in Fig. 9 for Experiment 1 was due to a consecutive sequence of zero-lag tests, or due to factors related to the sequence of

correct answers (at zero-lag an item is always correct). These same arguments would apply, however, to the sequence of zero-lag items in Experiment 2. In Fig. 9, the line labeled Experiment 2 represents a probability measure comparable to the one displayed for Experiment 1. Specifically, it is the probability of a correct response on the eventual test of the last S-R pair studied in a consecutive sequence of trials all involving S-R pairs tested at lag zero, as a function of the length of the

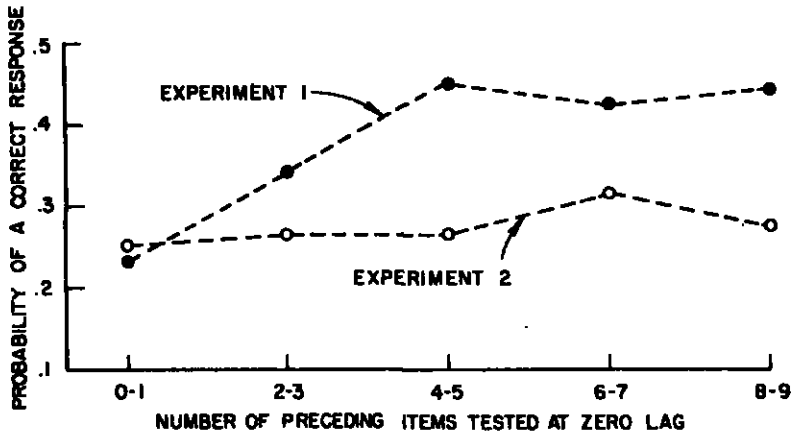


FIG. 9. Probability of a correct response as a function of the number of consecutive preceding items tested at zero lag (Experiment 1 and Experiment 2).

sequence. The model for Experiment 2 with its scheme for entering items in the buffer predicts that this curve should be flat; the data seem to bear out this prediction.

The close correspondence between the predicted and observed results in Experiments 1 and 2 provides strong support for the theory. The assumptions justified most strongly appear to be the fixed-size rehearsal buffer containing number-letter pairs as units, and the replacement assumptions governing O and N items. It is difficult to imagine a consistent system without these assumptions that would give rise to similar effects. Some of the predictions supported by the data are not at all intuitive. For example, the phenomenon displayed in Fig. 9 seems to be contrary to predictions based upon considerations of negative transfer. Negative transfer would seem to predict that a sequence of items having the same stimulus but different responses would lead to large amounts of interference and hence reduce the probability correct of the last item in the sequence; however, just the opposite effect was found. Furthermore, the lack of an effect in Experiment 2 seems to rule out explanations based on successive correct responses or successive zero-lag tests. Intuition notwithstanding, this effect was predicted by the model.

C. A Continuous Paired-Associate Memory Task with Multiple Reinforcements (Experiment 3)

In contrast to a typical short-term memory task, the subjects' strategy in paired-associate learning shifts from a reliance on rehearsal processes to a heavy emphasis on coding schemes and related processes that facilitate long-term storage. There are many factors, however, that contribute to such a shift, and the fact that items are reinforced more than once in a paired-associate learning task is only one of these. In the present experiment, all factors are kept the same as in Experiment 1, except for the number of reinforcements. It is not surprising, then, that subjects use essentially the same rehearsal strategy found in Experiment 1. It is therefore of considerable interest to examine the effects associated with repeated reinforcements of the same item.

In Experiment 3 only one stimulus set size, $s = 8$, was used. Each session began with eight study trials on which the eight stimuli were each randomly paired with a response. The stimuli and responses were two-digit numbers and letters, respectively. After the initial study trials, the session involved a series of consecutive trials each consisting of a test phase followed by a study phase. On each trial a stimulus was randomly selected for testing and the same stimulus was then presented for study on the latter portion of the trial. Whereas in Experiment 1, during the study phase of a trial, the stimulus was always re-paired with a new response, in the present experiment the stimulus was sometimes left paired with the old response. To be precise, when a particular S-R pair was presented for study the first time, a decision was made as to how many reinforcements (study periods) it would be given; it was given either 1, 2, 3, or 4 reinforcements with probabilities .30, .20, .40, and .10 respectively. When a particular S-R pair had received its assigned number of reinforcements, its stimulus was then re-paired with a new response on the next study trial, and this new item was assigned a number of reinforcements using the probability distribution specified above. In order to clarify the procedure, a sample sequence from trials n to $n+19$ is shown in Fig. 10. On trial $n+2$ stimulus 22 is given a new response, L, and assigned three reinforcements, the first occurring on trial $n+2$. The second reinforcement occurs on trial $n+3$ after a lag of zero. After a lag of 6, the third reinforcement is presented on trial $n+10$. After a lag of 8, stimulus 22 is re-paired with a new response on trial $n+19$. Stimulus 33 is sampled for test on trial $n+6$ and during the study phase is assigned the new response, B, which is to receive two reinforcements, the second on trial $n+9$. Stimulus 44 is tested on trial $n+4$, assigned the new response X which is to receive only one reinforcement; thus when 44 is presented again on trial $n+16$ it is assigned another response which by chance also is to receive only one reinforce-

ment, for on the next trial 44 is studied with response Q. The subject is instructed, as in Experiments 1 and 2, to respond on the test phase of each trial with the letter that was *last* studied with the stimulus being tested.

The same display devices, control equipment, and timing relations used in Experiment 1 were used in this study. There were 20 subjects, each run for 10 or more sessions; a session consisted of 220 trials. Details of the experimental procedure, and a more extensive account of the data

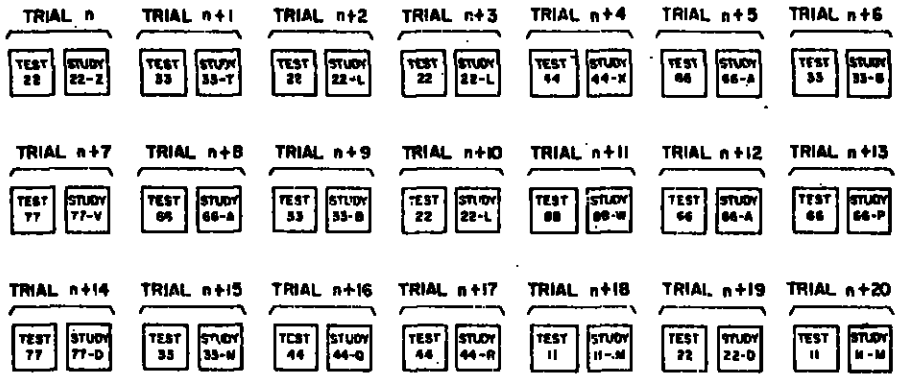


FIG. 10. A sample sequence of trials for Experiment 3.

analysis, including a fit of the model to response protocols of individual subjects, can be found in Brelsford, Shiffrin, and Atkinson (1968).

The model for Experiment 1 may be used without change in the present situation. There is some question, however, whether it is reasonable to do so. The assumptions concerning LTS storage and decay may be applied to items which are given multiple reinforcements: information is transferred to LTS at a rate θ whenever the item resides in the buffer, and decays from LTS by the proportion τ on each trial that the item is not present in the buffer. The assumption regarding O items also may be applied: since the stimulus already is in the buffer, the new response replaces the old one, thereby entering the item in the buffer (if, as is the case in this experiment, the old response is given yet another study, then nothing changes in the buffer). N items, however, are not so easily dealt with. N items, remember, are items whose stimuli are not currently represented in the buffer. In Experiment 1, the stimulus of every N item also was being paired with a new response. In the current experiment this is not always the case; some N items, although not in the buffer, will be receiving their second, third, or fourth reinforcement when presented for study. That is, some N items in this experiment will

already have a substantial amount of information stored on them in LTS. It seems reasonable that subjects may not rehearse an item which has just been retrieved correctly from LTS. The assumption regarding N items is therefore modified for purposes of the present experiment as follows. If a stimulus is tested and is not in the buffer, then a search of LTS is made. If the response is correctly retrieved from LTS, and if that stimulus-response pair is repeated for study, then that item will not be entered into the buffer (since the subject "knows" it already). If a new item is presented for study (i.e., the response to that stimulus is changed), or if the correct response is not retrieved from LTS (even though the subject may have made the correct response by guessing), then the study item enters the buffer with probability α. This slight adjustment of the replacement assumption allows for the fact that some items presented for study may already be known and will not enter the rehearsal buffer. This version of the model is the one used later to generate predictions for the data.

1. *Results*

Figure 11 presents the probability of a correct response as a function of lag for items tested after their first, second, and third reinforcements. The number of observations is weighted not only toward the short lags,

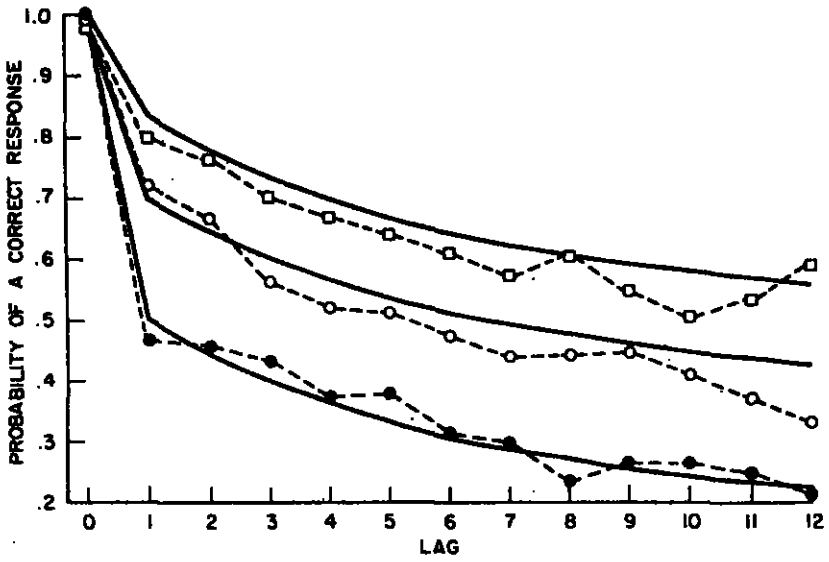
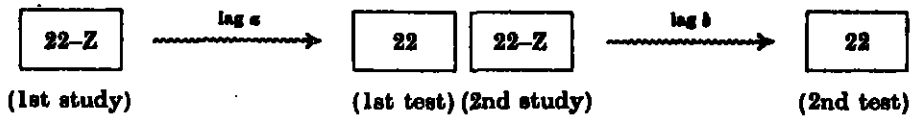


FIG. 11. Observed and theoretical probabilities of a correct response as a function of lag for items tested following their first, second, or third reinforcement (Experiment 3). --□-- Three reinforcements; --○-- two reinforcements; --●-- one reinforcement;— theory.

but also toward the smaller numbers of reinforcements. This occurs because the one-reinforcement lag curve contains not only the data from the items given just one reinforcement, but also the data from the first reinforcement of items given two, three, and four reinforcements. Similarly, the lag curve following two reinforcements contains the data from the second reinforcement of items given two, three, and four reinforcements, and the three-reinforcement curve contains data from the third reinforcement of items given three and four reinforcements. The lag curves in Fig. 11 are comparable to those presented elsewhere in this paper. What is graphed is the probability of a correct response to an item that received its jth reinforcement, and was then tested after a lag of n trials. The graph presents data for n ranging from 0 to 15 and for j equal to 1, 2, and 3. Inspecting the figure, we see that an item which received its first reinforcement and was then tested at a lag of 8 trials gave a correct response about 23% of the time; an item that received its second reinforcement and was then tested at lag 8 had about 44% correct responses; and an item that received its third reinforcement and was then tested at lag 8 had about 61% correct.

The curves in Fig. 11 exhibit a consistent pattern. The probability correct decreases regularly with lag, starting at a higher value on lag 1 the greater the number of prior reinforcements. Although these curves are quite regular, there are a number of dependencies masked by them. For example, the probability of a correct response to an item that received its second reinforcement and was then tested at some later trial will depend on the number of trials that intervened between the first and second reinforcements. To clarify this point consider the following diagram



Item 22-Z is given its first reinforcement, tested at lag a and given a second reinforcement, and then given a second test at lag b. For a fixed lag b, the probability of a correct response on the second test will depend on lag a. In terms of the model it is easy to see why this is so. The probability correct for an item on the second test will depend upon the amount of information about it in LTS. If lag a is extremely short, then there will have been very little time for LTS strength to build up. Conversely, a very long lag a will result in any LTS strength decaying and disappearing. Hence the probability of a correct response on the second test will be maximal at some intermediate value of lag a; namely, at a

lag which will give time for LTS strength to build up, but not so much time that excessive decay will occur. For this reason a plot of probability correct on the second test as a function of the lag between the first and second reinforcement should exhibit an inverted U-shape. Figure 12 is such a plot. The probability correct on the second test is graphed as a function of lag a. Four curves are shown for different values of lag b. The

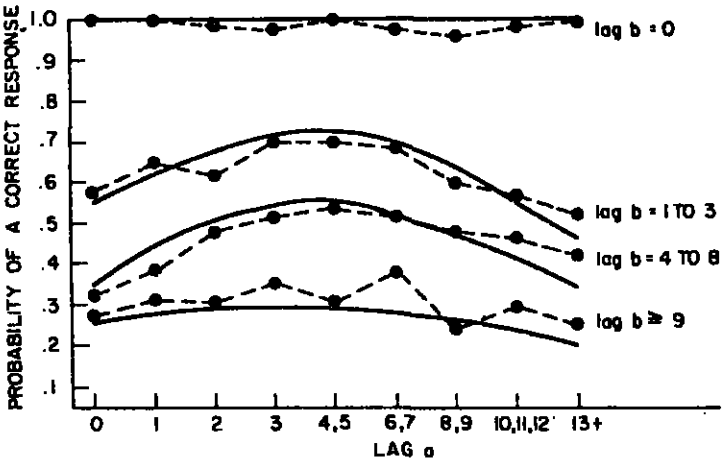


FIG. 12. Observed and theoretical probabilities of a correct response as a function of lag a (the spacing between the first and second reinforcement) (Experiment 3).

four curves have not been lumped over all values of lag b because we wish to indicate how the U-shaped effect changes with changes in lag b. Clearly, when lag b is zero, the probability correct is one and there is no U-shaped effect. Conversely, when lag b is very large, the probability correct will tend toward chance regardless of lag a, and again the U-shaped effect will disappear. The functions shown in Fig. 12 give support to the assumption that information is being transferred to LTS during the entire period an item resides in the buffer. If information is transferred, for example, only when an item first enters the buffer, then our model will not predict the rise in the functions of Fig. 12 for lag a going from zero to about five. The rise is due to the additional information transferred to LTS as lag a increases.

2. *Theoretical Analysis*

A brief review of the model is in order. O items (whose stimulus is currently in the buffer) always enter the buffer. N items (whose stimulus is not currently in the buffer) enter the buffer with probability α if they

are also new items (i.e., receiving their first reinforcement). However, N items do not enter the buffer if they are repeat items and were correctly retreived from LTS on the immediately preceding test; if they are repeat items and a retrieval was not made, then they enter the buffer with probability α. An O item entering the buffer occupies the position of the item already there with the same stimulus; an entering N item randomly replaces one of the items currently in the buffer. During the period an item resides in the buffer, information is transferred to LTS at a rate θ per trial. This information decays by a proportion τ on each trial after an item has left the buffer.[14] The subject is always correct at a lag of zero, or if the item is currently in the buffer. If the item is not in the buffer a search of LTS is made, and the correct response is retrieved with a probability that is an exponential function of the amount of information currently in LTS (i.e., the same function specified for Experiments 1 and 2). If the subject fails to retrieve from LTS, then he guesses. There are four parameters for this model: r, the buffer size; α, the buffer entry probability; θ, the transfer rate of information to LTS; and τ, the parameter characterizing the LTS decay rate once an item has left the buffer.

Estimates of r, α, θ, and τ were made using the data presented in Figs. 11 and 12. We shall not go into the estimation procedures here, for they are fairly complex; in essence they involve a modified minimum χ^2 procedure where the theoretical values are based on Monte Carlo runs. The parameter estimates that gave the best fit to the data displayed in Figs. 11 and 12 were as follows: $r = 3$; $\alpha = .65$; $\theta = 1.24$; and $\tau = .82$. Once these estimates had been obtained they were then used to generate a large-scale Monte Carlo run of 12,500 trials. The Monte Carlo procedure involved generating pseudo-data following precisely the rules specified by the model and consulting a random number generator whenever an event occurred in the model that was probabilistically determined. Thus the pseudo-data from a Monte Carlo run is an example of how real data would look if the model was correct, and the parameters had the values used in the Monte Carlo computation. In all subsequent discussions of Experiment 3, the predicted values are based on the output of the Monte Carlo run. The run was very long so that in all cases the theoretical curves are quite smooth, and we doubt if they reflect fluctuations due to sampling error. A detailed account of the estimation and prediction procedures for this experiment is given in Brelsford, Shiffrin, and Atkinson (1968).

The predictions from the theory are shown as the smooth curves in

[14] In this experiment an item receiving x reinforcements may enter the buffer as many as x times. When the item is in the buffer the θ-process is activated, and when not in the buffer the τ-process takes over.

Figs. 11 and 12. It should be evident that the predicted values are quite close to the observed ones. Note also that the seven curves in the two figures are fit simultaneously with the same four parameter values; the fact that the spacing of the curves is accurately predicted is particularly interesting.

We now examine a number of statistics that were not used in making parameter estimates. First consider the all-same and all-different curves shown in Fig. 13; these are the same functions displayed in Figs. 5 and 6

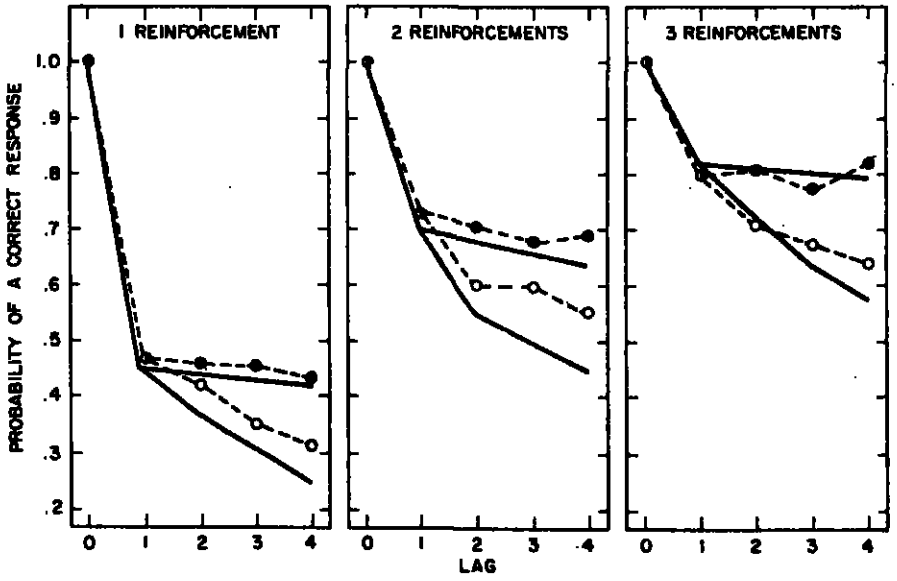


FIG. 13. Observed and theoretical probabilities of a correct response as a function of lag for the "all-same" and "all-different" conditions (Experiment 3). --●-- all-same; --○-- all-different;—theory.

for Experiment 1. For the all-same curve, we compute the probability of a correct response as a function of the lag, when all the intervening items between study and test involve the same stimulus. There are three such curves, depending on whether the study was the first, second, or third reinforcement of the particular S-R pair. The model predicts that once the intervening stimulus enters the buffer, there will be no further chance of any other item being knocked out of the buffer. Hence these curves should drop at a much slower rate than the unconditional lag curves in Fig. 11. The all-different curve plots the probability of a correct response as a function of lag, when the intervening items between study and test all involve different stimuli. Again there are three curves depending on whether the study was the first, second, or third reinforcement of the S-R pair. The all-different sequence maximizes the expected

number of intervening N items and therefore the curve should have a much faster drop than the unconditional lag curves in Fig. 11. The predictions are shown in the figure as solid lines. The correspondence between predicted and observed values is reasonably good. It is particularly impressive when it is noted that the parameter values used in making the predictions were estimated from the previous data.

We next examine the data displayed in Fig. 14. Consider a sequence of consecutive trials all involving the same stimulus, but where the

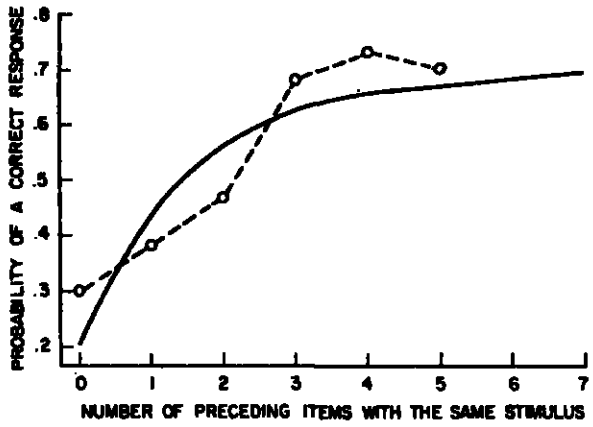
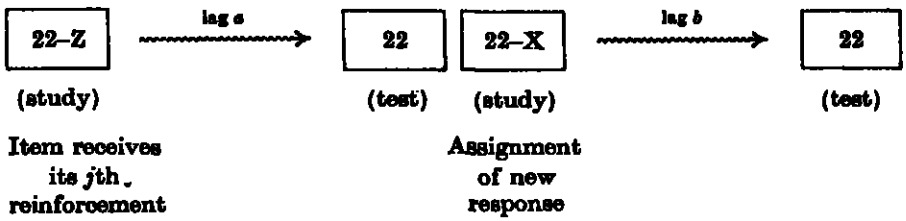


FIG. 14. Observed and theoretical probabilities of a correct response as a function of the number of consecutive preceding items using the same stimulus (Experiment 3).

response paired with the stimulus on the study phase of the last trial in the sequence is different from the response on the immediately preceding trial. Then, the theory predicts that the longer this sequence of consecutive trials, the higher will be the probability of a correct response when the last item studied in the sequence is eventually tested. This is so because the probability of the last item entering the buffer increases as the length of the sequence increases: once any item in the sequence enters the buffer, every succeeding one will. The data is shown in Fig. 14. What is graphed is the length of the sequence of trials all involving the same stimulus versus the probability of a correct response when the last item studied in the sequence is eventually tested. In this graph we have lumped over all lags at which the eventual test of the last item is made. The predictions generated from the previously estimated parameter values are shown as the smooth line. The predicted values, though not perfect, are surprisingly close to the observed proportions correct. It is worth reemphasizing that considerations of negative transfer make this result somewhat unexpected (see p. 140).

We next examine another prediction of the theory that ran counter to our initial intuitions. To make matters clear, consider the following diagram:



Item 22-Z is studied for the jth time and then tested at lag a; on this trial 22 is paired with a new response X, and tested next at lag b. According to the theory, the shorter lag a, the better performance should be when the item is tested after lag b. This prediction is based on the fact that the more recently a stimulus had appeared, the more likely that it was still in the buffer when the next item using it was presented for study; if the stimulus was in the buffer, then the item using it would automatically enter the buffer. In the present analysis, we examine this effect for three conditions: the preceding item using the stimulus in question could have just received its first, second, or third reinforcement. Figure 15 presents the appropriate data. In terms of the above diagram, what is plotted is the value of lag a on the abscissa versus the probability of a correct response lumped over all values of lag b on the ordinate; there is a separate curve for $j = 1$, 2, and 3.

The predicted curves are based upon the previous parameter estimates. The predictions and observations coincide fairly well, but the effect is not as dramatic as one might hope.[15] One problem is that the predicted decrease is not very large. Considerably stronger effects may be expected if each curve is separated into two components: one where the preceding item was correct at test and the other where the preceding item was not correct. In theory the decrease predicted in Fig. 15 is due to a lessened probability of the relevant stimulus being in the buffer as lag a increases. Since an item in the buffer is always responded to correctly, conditionalizing upon correct responses or errors (the center test in the above diagram) should magnify the effect. To be precise, the decrease will be accentuated for the curve conditional upon correct responses, whereas no decrease at all is predicted for the curve conditional upon errors. If an error is made, the relevant stimulus cannot be in the buffer and hence the new item enters the buffer with probability α

[15] A curve comparable to the one displayed in Fig. 15 for the one-reinforcement condition was obtained from the data of Experiment 1. This curve showed a similar but more pronounced drop and was well predicted by the model.

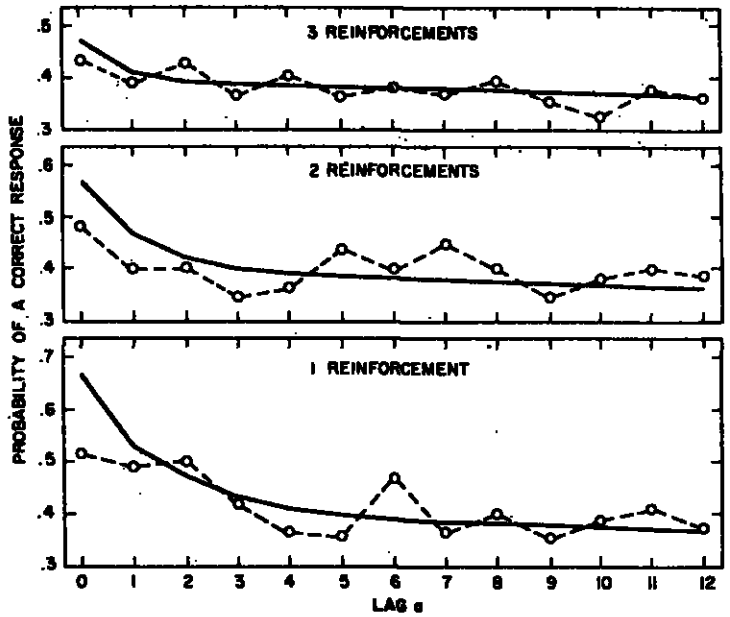


FIG. 15. Observed and theoretical probabilities of a correct response as a function of lag a (the lag of the item preceding the item tested, but using the same stimulus) (Experiment 3). ---o--- Data;—theory.

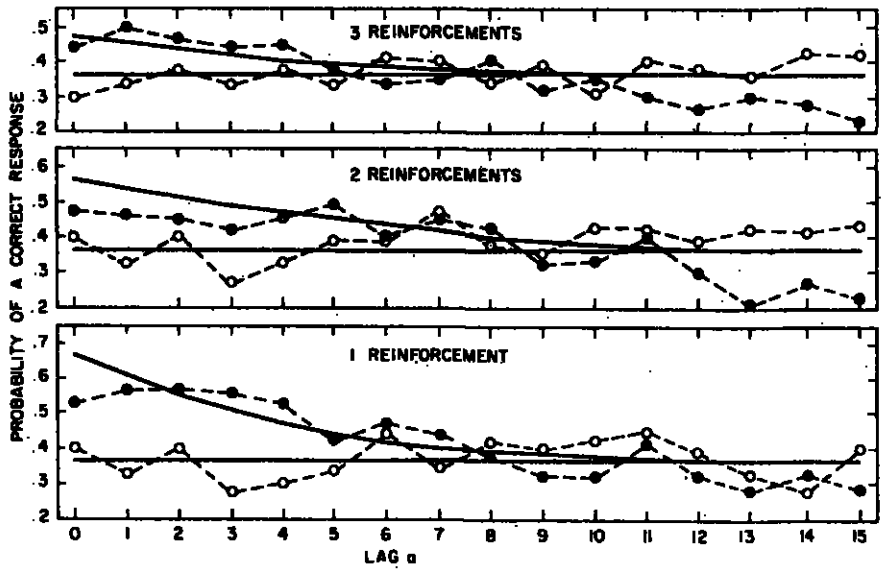
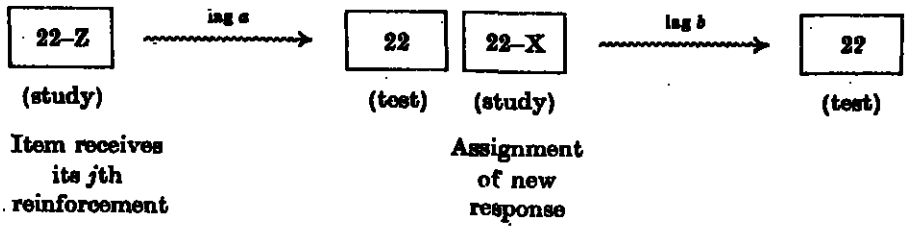


FIG. 16. Observed and theoretical probabilities of x correct response as a function of lag a conditionalized on errors or successes on the test at lag a (Experiment 3). ---●--- Correct data; ---o--- error data;—theory.

independent of lag a. Figure 16 gives the conditional curves and the predictions. The decreasing effect is fairly evident for the "correct" curves; as predicted, the "error" curves are quite flat over lags.[16] Conceivably one might argue that the effects are due to item selection, correct responses indicating easier stimuli and incorrect responses indicating more difficult ones. This objection, however, seems contraindicated in the present case. It is difficult to imagine how item selection could explain the crossing of the correct and error curves found in each of the three diagrams.[17] Indeed, the model does not explain the crossover —*the model predicts that the two curves should meet.* The model is in error at this point because it has not been extended to include negative transfer effects, an extension which would not be difficult to implement. An item responded to correctly at a long lag probably has a strong LTS trace; this strong trace would then interfere with the LTS trace of the new item which, of course, uses the same stimulus. All in all, these curves and predictions may be considered to provide fairly strong support for the details of the model, even to the extent of illuminating the one aspect omitted, albeit intentionally, from the assumptions.

The aspect left out is, of course, that of LTS response competition, or negative transfer. The model fails to take account of this effect because it fails to keep track of residual LTS strength remaining as a result of the previous items using the same stimulus. This lack is most clearly indicated by the occurrence of intrusion errors, particularly errors which were correct responses on the preceding occurrence of that stimulus. For example, consider the following sequence:



[16] The astute reader will have noticed that the predicted decrease becomes smaller as the number of reinforcements increases. The fact that the data support this prediction is quite interesting, for it sheds light upon the buffer replacement assumptions used in this model. The decreasing effect as reinforcements increase is predicted because the probability of entering the buffer is reduced for an item receiving its third reinforcement; remember, an item recovered from LTS is not entered into the buffer. Thus as reinforcements increase, the probability of being in the buffer decreases, and the normally increased probability of being in the buffer as a result of a short lag a is partially counterbalanced.

[17] Undoubtedly there are some selection effects in the data graphed in Fig. 16, but their magnitude is difficult to determine. Thus, these data should be regarded with some wariness.

Item 22-Z is studied for the jth time and then tested at lag a; on this trial 22 is paired with a new response X and next tested at lag b. By an intrusion error we mean the occurrence of response Z when 22 is tested at the far right of the diagram. The model predicts that these intrusion errors will be at chance level (1/25), independent of lag and number of reinforcements. In fact, these predictions fail. Figure 17 presents the

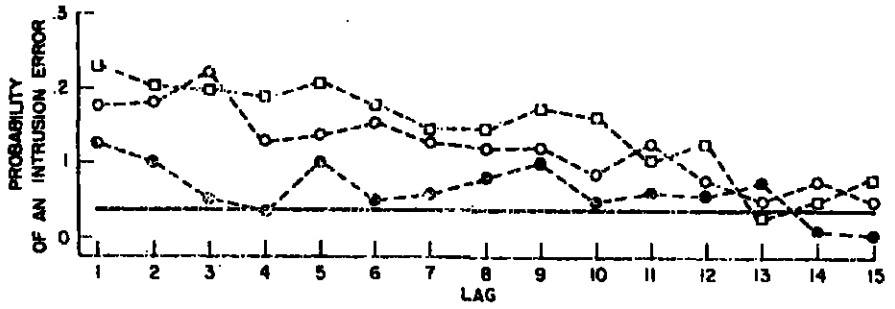


FIG. 17. Probability that the correct response for the preceding item using the same stimulus will be given in error to the present item (Experiment 3). --□-- Three reinforcements; --○-- two reinforcements; --●-- one reinforcement; —chance.

probability of intrusion errors as a function of lag b; where the data have been lumped over all values of lag a, three curves are plotted for $j = 1, 2,$ and 3. This failure of the model is not very distressing because it was expected: the model could be extended in a number of obvious ways to take account of competing LTS traces without appreciably changing any of the predictions so far presented. The extension has not been made because our interest in this study is centered upon short-term effects.

Judging by the agreement between theory and data for each of the effects examined, the accuracy of the model is extremely good. It is interesting to note that the multiple-reinforcement procedure is not sufficient by itself to cause the subjects to switch their strategies from rehearsal to coding. The major emphasis still appears to be on rehearsal manipulations in STS, a not entirely surprising result since the situation is identical to that used in Experiment I except for the number of reinforcements given. The comments previously made concerning the difficulty associated with LTS storage in Experiment I apply here also. Because the emphasis is upon short-term mechanisms, this experiment is not to be considered in any strong sense as a bridge to the usual paired-associate learning situation. Nevertheless, a number of long-term effects, such as intrusion errors and interference caused by previously learned items on new items with the same stimulus, demonstrate that LTS mechanisms cannot be ignored in the theory. In Section V we consider

experiments that are designed to provide a sharper picture of the workings of LTS; experimentally this is accomplished by systematically varying the number of items in LTS through which searches must be made. Before considering this problem, however, there are other features of the STS rehearsal strategy to be explored. We turn next to an experiment in which the probability of entering an item into the buffer is manipulated experimentally.

D. Overt versus Covert Study Procedures (Experiment 4)

The statistics considered in the previous section leave little doubt about the role of O items, N items, and the buffer entry parameter α. But one question we have not considered is whether α is amenable to experimental manipulation; if the process is really under the control of the subject, such manipulation would be expected. We now turn to a study by Brelsford and Atkinson (1968) which was designed to answer this question.

In Experiment 1, the proportions of O items and N items were varied by changing the size of the stimulus set, and the predicted differences were found. Manipulating α, however, is a somewhat more subtle task since it is the subject's strategy that must be affected. One experimental device which seems likely to increase the probability of an item's entering the buffer is to have the subject recite the item aloud as it is presented for study; this will be referred to as the "overt" study procedure. The "covert" study procedure is simply a replication of the procedure used in Experiment 1 where the subject was not required to recite the item aloud when it was presented for study, but simply told to study it.

1. *Method*

The method was identical to that used in Experiment 1 except for the following changes. The size of the stimulus set was fixed at six for all subjects and sessions. Each session consisted of 200 trials divided into four 50-trial blocks alternating between the overt and covert conditions. The initial 50-trial block was randomly chosen to be either an overt or a covert condition. The covert condition was identical in all respects to Experiment 1; when the word "study" and an S-R pair appeared on the CRT (the display screen) the subjects were told to silently study the item being presented. In the overt blocks, instead of the word "study" appearing on the CRT during the study portion of a trial, the word "rehearse" appeared. This was a signal for the subject to recite aloud twice the item then being presented for study. This was the only difference from the procedure used during the covert trials. It was hoped that the act of repeating the items aloud would raise the subject's probability of entering the item into his rehearsal buffer.

2. *Results*

In order to allow for the subject's acclimation to a change in study conditions, the first 15 trials of each 50-trial block are not included in the data analysis. Figure 18 presents the lag curves for the overt and covert conditions. It is evident that performance is superior in the overt condition. Furthermore, the overt lag curve is S-shaped in form, an

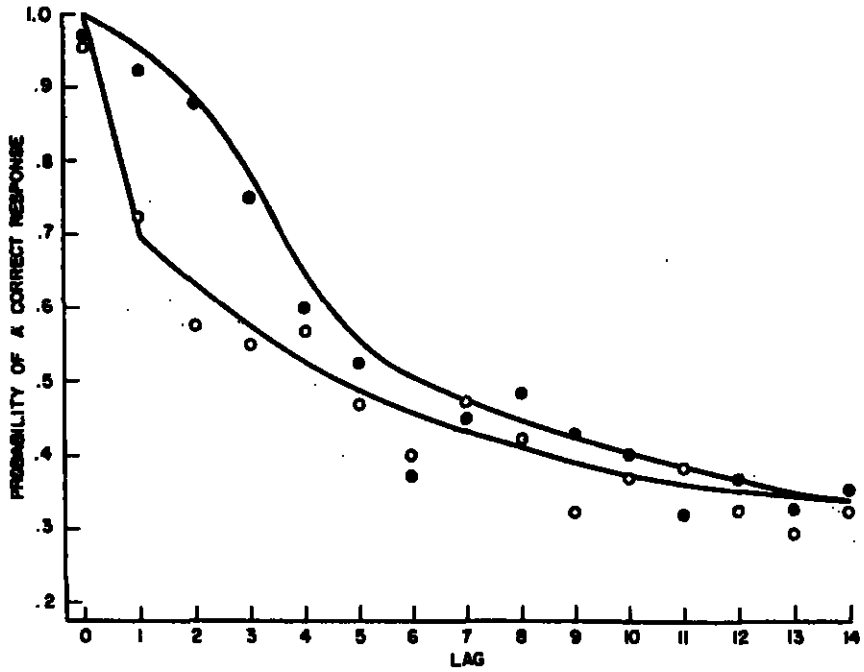


FIG. 18. Observed and theoretical probabilities of a correct response as a function of lag (Experiment 4). ● Overt; ○ covert;—theory.

effect not observed in earlier curves. Since the parameters of the models will be estimated from these curves, the model is presented before considering additional data.

The model for the covert condition is, of course, identical to that used in the analysis of Experiment 1. It has the four parameters r, α, θ, and τ. Since it was hypothesized that α would be raised in the overt condition, we might try estimating α separately for that condition. This version of the model will not fit the overt data, however, because of the pronounced S-shaped form of the lag curve. Although setting α equal to 1.0 will predict better performance in the overt condition, the lag curve will have the form of an exponentially decreasing function, which is clearly not found in the data. In order to account for the S-shaped curve, we

need to assume that in the overt condition the subject tends to knock the oldest items out of the buffer first. In the model for the covert case, an entering N item is said to knock out at random any item currently in the buffer. It will be assumed for the overt case that an entering N item tends to replace the oldest item in the buffer; remember O items are items whose stimulus is currently in the buffer and they automatically replace the item with that stimulus. This probability of knocking the oldest items from the buffer first is specified as follows: if there are r items in the buffer and they are numbered so that item 1 is the oldest and item r is the newest, then the probability that an entering N item will knock the jth item from the buffer is

$$\frac{\delta(1-\delta)^{j-1}}{1-(1-\delta)^{r}}.$$

This equation is derived from the following scheme. The oldest item is knocked out with probability δ. If it is not knocked out, then the next oldest is knocked out with probability δ. The process continues cyclically until an item is finally selected to be knocked out. When δ approaches zero, the knockout probabilities are random, as in the covert case. When δ is greater than zero there will be a tendency for the oldest items to be knocked out of the buffer first; in fact if δ equals one, the oldest item will always be the one knocked out. It should be clear that the higher the value of δ, the greater the S-shaped effect predicted for the lag curve.

The model for the curves in Fig. 18 is therefore structured as follows. The parameters r, θ, and τ will be assumed to be the same for the two conditions; the parameters α and δ will be assumed to be affected by the experimental manipulation. To be precise, in the covert case α will be estimated freely and δ will be set equal to zero, which is precisely the model used in Experiment 1. In the overt case, α will be set equal to 1.0, which means that every item enters the buffer, and δ will be estimated freely. The parameter values that provided the best χ^2 fit to the data in Fig. 30 were $r = 3$, $\theta = .97$, $\tau = .90$; for the covert condition the estimate of α was .58 (with δ equal to zero) and for the overt condition the estimate of δ was .63 (with α equal to one). The predictions for this set of parameter values are shown in Fig. 18 as smooth curves. The improvement in performance from the covert to overt conditions is well predicted; actually it is not obvious that variations in either α or δ should affect the overall level of performance. The principal reason for the improvement is due to the value of α; placing every item into the buffer means that an item entering the buffer will be expected to stay there for a shorter period than if some items did not enter the buffer. This shorter period in the buffer, however, is outweighed by the advantages resulting from the entry of every item in the first place. It is not

easy to find statistics, other than the gross form of the lag curve, which reflect changes in δ; thus the assumption that the oldest items are lost first is not easy to verify in a direct way. Nevertheless, it is quite common to find experiments that yield S-shaped recency curves and these results can be fit by assuming that the oldest items in the buffer tend to be knocked out first. Other examples will be presented in Section V.

A number of additional aspects of the data will now be examined. First we consider the "all-same" and "all-different" lag curves. Figure 19

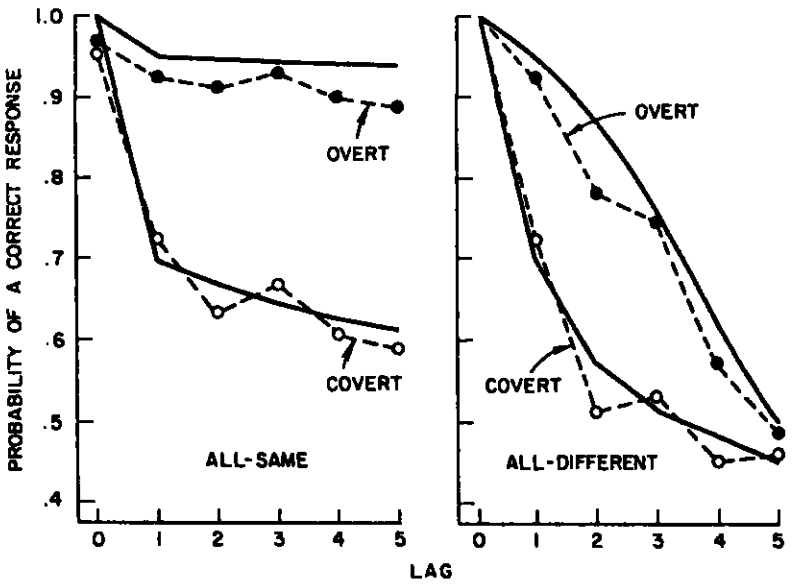


FIG. 19. Observed and theoretical probabilities of a correct response as a function of lag for the "all-same" and "all-different" conditions (Experiment 4).

gives the "all-same" lag curves for the overt and covert conditions. This curve gives the probability of a correct response for an item when all of the intervening items (between its study and test) have the same stimulus. This curve will be quite flat because the items following the first intervening item tend to be O items which will not knock other items from the buffer (for the overt case, *every* item following the first intervening item is an O item, since all items enter the buffer). Figure 19 also presents the "all-different" lag curves. This curve is the probability of making a correct response to a given item when the other items intervening between its study and test all involve different stimuli. The

predictions generated by the previous parameter values are given by the smooth curves; they appear to be quite accurate.

We now look for an effect that will be sharply dependent upon the value of α and hence differ for the overt and covert conditions. Such an effect is given in Fig. 20; graphed there is the probability of a correct response as a function of the number of immediately preceding items having the same stimulus as the item in question. This is the same

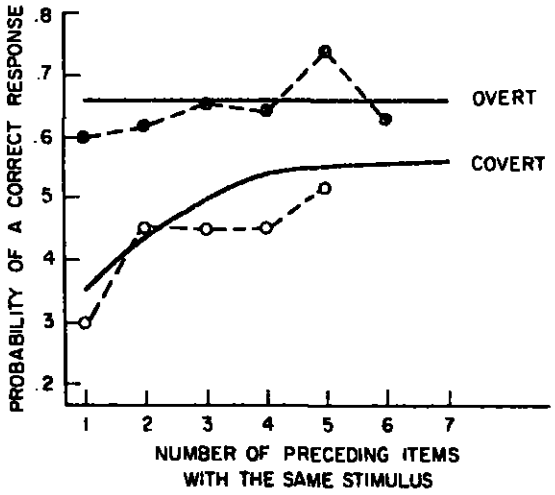


FIG. 20. Observed and theoretical probabilities of a correct response as a function of the number of consecutive preceding items all using the same stimulus (Experiment 4).

statistic that is plotted in Figs. 9 and 14; it is not a lag curve because the probability correct is given as an average over all possible lags at which the item was tested. If α is less than 1, then the length of the preceding sequence of items with the same stimulus will be an important variable; since any item in the sequence which enters the buffer will cause every succeeding item in the sequence to enter the buffer, the probability that the item in question enters the buffer will approach one as the length of the preceding sequence of items all using the same stimulus increases. For α equal to 1 (overt condition), every item enters the buffer and therefore no change would be expected. As indicated in Fig. 20, the data and theory are in good agreement. The slight rise in the data points for the overt condition may indicate that an estimate of α a little below 1.0 would improve the predictions, but the fit as it stands seems adequate.

E. Additional Variables Related to the Rehearsal Buffer (Experiments 5, 6, and 7)

1. *Known Items and the Buffer (Experiment 5)*

In this section we shall consider briefly a number of other variables that relate to the rehearsal buffer. The overt manipulation in the preceding section succeeded in raising to near 1.0 the probability of entering an item in the buffer. As an alternative, one would like an experimental manipulation which would cause the entry probability to drop to near zero for some items. W. Thomson at Stanford University has performed an experiment that satisfies this requirement. The experimental manipulation involves interspersing some extremely well-known items among a series of items never seen before. The assumption is that a well-known item will not enter the rehearsal buffer. The experiment was performed using a modification of the "all-different" stimulus procedure employed in Experiment 2. The stimuli were consonant-vowel-consonant trigrams and the responses were the digits 0–9. For each subject two stimuli were chosen at the start of the first session and assigned responses. These S-R pairs never changed throughout the series of sessions. Except for these two items all other items were presented just once. The size of the to-be-remembered set(s) was six which included the two "known" items. The presentation schedule was as follows: on each trial with probability .5 one of the two known items would be presented for test and then given yet another study period; otherwise one of the four items in the current to-be-remembered set would be tested and a new stimulus-response pair then presented for study. Thus, the task was like that used in Experiment 2, except that on half the trials the subject was tested on, and then permitted to study, an S-R pair which was thoroughly known. The data from the first session in which the known items were being learned will not be considered.

The simplest assumption regarding the two known items is that their probability of entering the buffer is zero. This assumption is the one used in the multiple-reinforcement study (Experiment 3); namely, that an item successfully recovered from LTS is not entered into the buffer.[18] In contrast to Experiment 3, in this study it is easy to identify the items that are known since they are experimentally controlled; for this reason we can look at a number of statistics depending upon the likelihood of entering known items into the buffer. The one of particular interest is presented in Fig. 21. Graphed there is the unconditional lag curve, the

[18] Underwood and Ekstrund (1967) have found that insertion of known items from a previously learned list into a succeeding list improves performance on the learning of unknown items on the second list, although list length was a confounded variable.

"all-known-intervening" lag curve and the "all-unknown-intervening" lag curve. By known items we mean the two S-R pairs that repeatedly are being studied and tested; by unknown items we mean those pairs that are studied and tested only once. The unconditional lag curve gives the probability correct for unknown items as a function of lag, independent of the type of items intervening between study and test; of

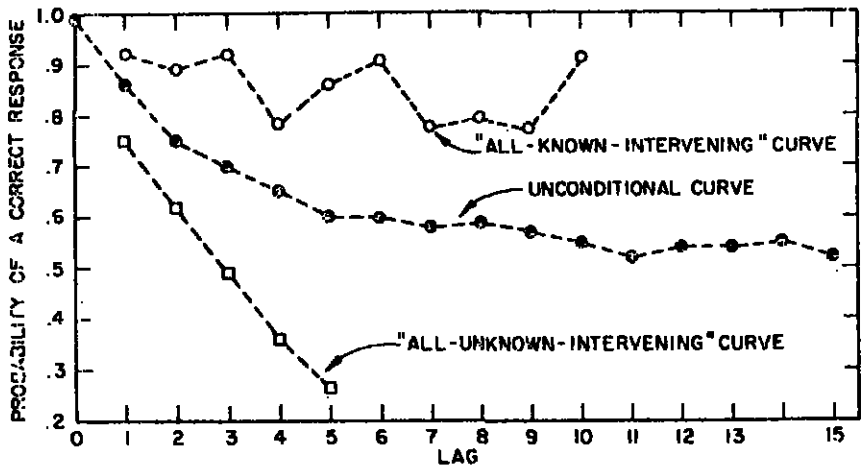


FIG. 21. Observed and theoretical probabilities of a correct response as a function of lag, for the overall condition and for the "all-known-intervening" and "all-unknown-intervening" conditions (Experiment 5).

course, the corresponding curve for known items would be perfect at all lags since subjects never make errors on them. The all-known-intervening curve gives the probability correct as a function of lag, when all of the items intervening between study and test are known items. If none of the known items enter the buffer, this curve should be level from lag 1 on and equal to α, the probability that the item entered the buffer when presented for study. At the opposite extreme is the all-unknown-intervening curve; when all the intervening items are new, the probability of knocking the item of interest from the buffer increases with lag and therefore the curve should decay at a rapid rate. It may be seen that this curve indeed drops at a more rapid rate than the unconditional lag curves. The marked difference between the all-known and all-unknown curves in Fig. 21 leads us to conclude that known and unknown items clearly have different probabilities for entering the rehearsal buffer. If the all-known curve were flat after lag 1, then the probability for entering a known item into the buffer would be zero. Another possibility is that

α is indeed zero for known items, but that the subject occasionally picks an item from LTS for additional rehearsal when a known item is presented.

2. *Response Time Measures (Experiment 6)*

We now turn to a consideration of some latency results. Potentially, latencies offer an avenue of analysis that could be more fruitful than the

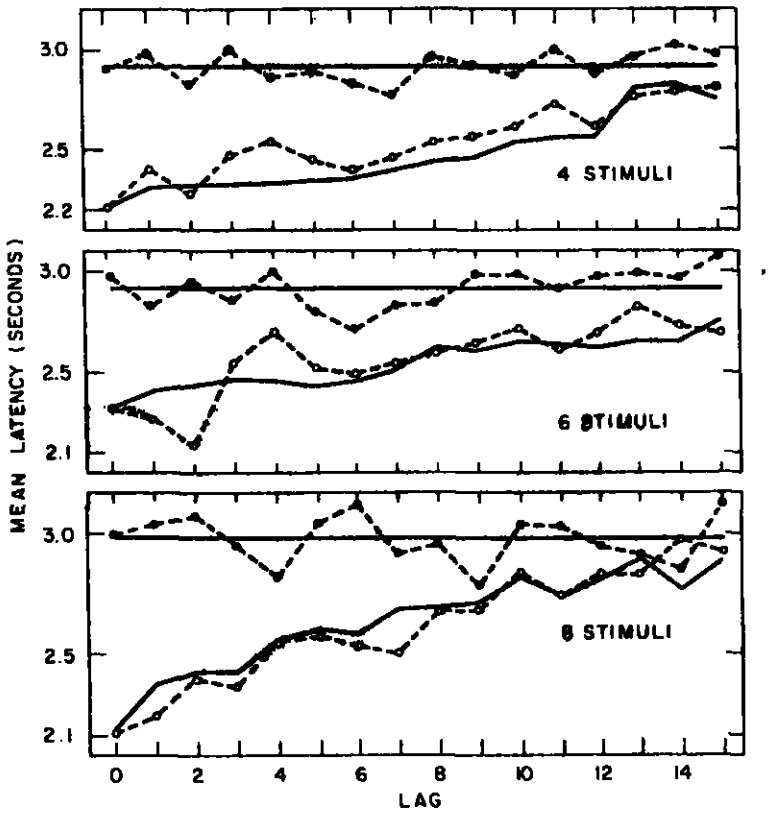


FIG. 22. Observed and theoretical mean latencies as a function of lag for correct and incorrect responses (Experiment 6). --●-- Error latencies; --○-- correct latencies;—predicted latencies.

analysis of choice response data; we say this because the latencies should reflect search and retrieval times from both STS and LTS. A detailed latency analysis is beyond the scope of this paper, but one simple result will be considered. Figure 22 presents the average latencies as a function of lag for correct and incorrect responses in a study by Brelsford *et al.* (1966). This experiment employed the same procedure described earlier

in our discussion of Experiment 1 except that only 6 rather than 26 responses were used. As in Experiment 1, this study used three different stimulus-set sizes; i.e., s equalled 4, 6, or 8. For each stimulus set in Fig. 22 it may be seen that the correct and incorrect latency curves converge at long lags. This convergence would be expected since the probability of a correct response is dropping toward chance at long lags. The theoretical curves are based on an extremely simple latency model which assumes that latencies for responses correctly retrieved from either LTS or STS have a fixed mean value λ, whereas a failure to retrieve and a subsequent guess has a fixed mean value of λ'. Thus error responses always have a mean latency λ'; however, a correct response may occur as a result of a retrieval from memory or a correct guess, and consequently its latency is a weighted average of λ and λ'. We can estimate λ' as the average of the points on the latency lag curve for errors, and λ can be set equal to the latency of a correct response at lag zero since all responses are due to retrievals from memory at this lag. In order to predict the remaining latency data, we make use of the observed probability of a correct response as a function of lag; these values are reported in Brelsford *et al.* (1966). If p_i is the observed probability of a correct response at lag i, then

$$p_i = x_i + (1 - x_i)\tfrac{1}{6}$$

where x_i is the probability of retrieving the response from memory and $(1 - x_i)\frac{1}{6}$ is the probability of making a correct response by guessing. Estimating x_i in this way, we predict that the mean latency of a correct response at lag i is simply $x_i\lambda + (1 - x_i)\lambda'$. Using this equation and estimating λ and λ' as indicated above, leads to the theoretical curves displayed in Fig. 22. The error latency curve is predicted to be equal to λ' for all lags, whereas the correct latency curve is λ at lag 0 and approaches λ' over lags as the estimate of x_i goes to zero. This latency model is of course oversimplified, and fails to take into account differences in latencies due to retrieval from STS as compared to retrieval from LTS; the results nevertheless indicate that further analyses along these lines may prove fruitful.

3. *Time Estimation (Experiment 7)*

One factor related to our model that has not been discussed is temporal memory. It seems clear that there is some form of long-term temporal memory; in a negative transfer paradigm, for example, there must be some mechanism by which the subject can distinguish between the most recent response paired with a stimulus versus some other response paired with that stimulus at an earlier time. This temporal memory undoubtedly involves the long-term store; somehow when an

event is stored in LTS it also must be given a time tag or stored in such a way that the subject can date the event (albeit imperfectly) at the time of retrieval. In addition to long-term temporal storage, there is evidence that a subject's estimate of elapsed time depends upon an item's length of residence in the buffer. An experiment by R. Freund and D. Rundis at Stanford University serves to illustrate the dependence of temporal memory upon the buffer.[19] The study employed essentially

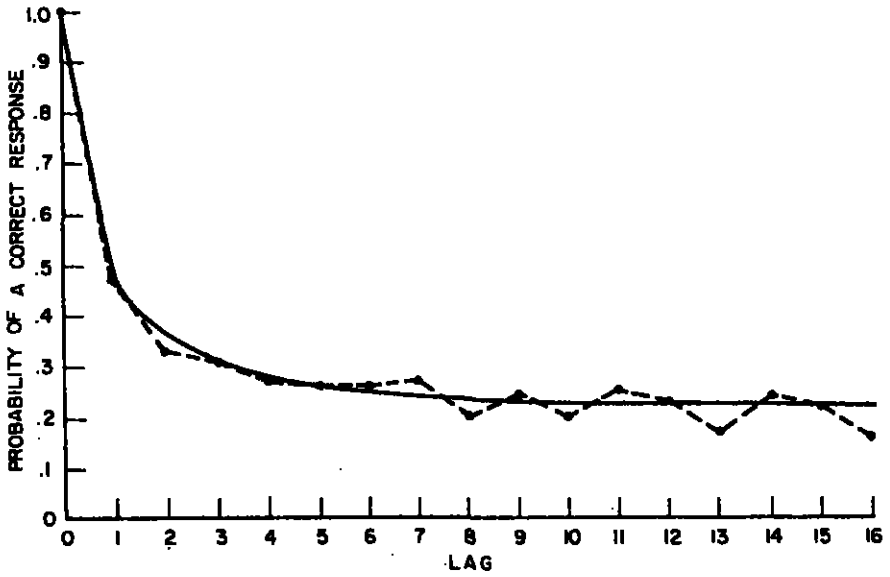


FIG. 23.. Observed and theoretical probabilities of a correct response as a function of lag (Experiment 7). - - - Observed;—predicted.

the same procedure used in Experiment 2. There was a continuous sequence of test-plus-study trials and the stimuli kept changing throughout each session; each stimulus appeared only once for study and test. The stimuli were consonant-vowel-consonant trigrams and the responses were the 26 letters of the alphabet; the size of the to-be-remembered set of items was fixed at eight. When a stimulus was tested the subject first gave his best guess of the response that had been previously studied with the stimulus and then gave an estimate of the number of trials that intervened between the item's initial study and final test; this estimate could range from 0 to 13; if the subject felt the lag was greater than 13 he responded by pressing a key labeled 14+.

The unconditional lag curve for the probability of a correct response is presented in Fig. 23. The solid line represents the predictions that were

[19] This study employs a time-estimation procedure similar to one developed by L. R. Peterson (personal communication).

generated by the model used to fit Experiment 2. The parameter values providing the best fit to the lag curve were $r = 2$, $\alpha = .57$, $\theta = .13$, $\tau = 1.0$. The data of interest is presented in Fig. 24. The average lag judgment is plotted as a function of the actual lag. The solid dots are the average lag judgments for those items to which a correct response was given; the open circles are the average lag judgments for those items to which an incorrect response was given. If lag judgments were perfect,

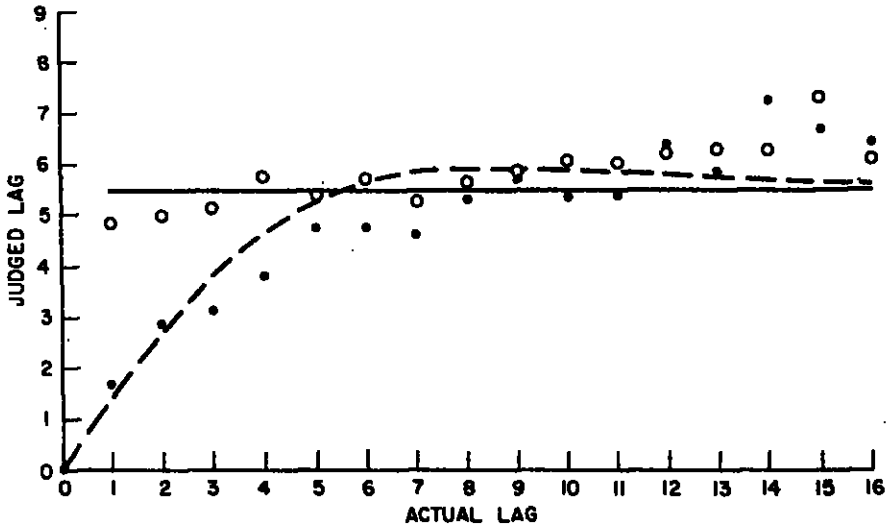


FIG. 24. Observed and theoretical mean lag judgments as a function of the actual lag (Experiment 7). ○ Error data;—error theory; ● correct data; – –correct theory.

they would fall on the 45° diagonal; it may be seen that the correct curve is fairly accurate to about lag 5 and then tails off. The lag judgments associated with incorrect responses seem to be virtually unrelated to the actual lag. This indicates that the retrieval of a correct response and temporal estimation are closely related. An extremely simple model for this data assumes that the mean lag judgment for an item in the buffer is the true lag value; any item not in the buffer is given a lag judgment at random from a distribution that is unrelated to the true lag. The predictions using the above parameter estimates are shown in Fig. 24. Freund and Rundis have developed more elaborate models which include both a long- and short-term temporal memory and have obtained quite accurate predictions; but these models will not be examined here. The point we want to make by introducing these data is that temporal memory may be tied to the short-term system even more strongly than to the long-term system.

V. Experiments Concerned with Long-Term Search and Retrieval

The major purpose of this section is to examine a series of experiments concerned with search and retrieval processes in LTS. These experiments differ from those of the preceding section in that the memory tasks are not continuous; rather, they involve a series of discrete trials which are meant to be relatively independent from one to the next. On each trial a new list of items is presented sequentially to the subject for study; following the presentation a test is made on some aspect of the list. Using this procedure, the size of the list, d, can be systematically manipulated. Variations in list size affect the size of the memory set through which the subject must search when tested, and consequently search and retrieval processes can be examined in more detail than was previously possible. The title of this section is not meant to imply, however, that the short-term processes involved in these experiments are different from those appearing in the continuous-presentation situations; in fact, the models used to describe the experiments of this section will be based upon the same STS rehearsal buffer introduced earlier. The difference is one of emphasis; the long-term processes will be elaborated and explored in greater depth in this section. This exploration of long-term models will by no means be exhaustive, and will be less extensive than that carried out for the short-term processes.

Prior to an examination of particular experiments, a few remarks need to be made about the separability of lists. In any experiment in which a series of different lists is presented, we may ask just what information in LTS the subject is searching through at test. The same problem arises, though less seriously, in experiments where the subject is tested on only one list. Clearly the information relevant to the current list of items being tested must be kept separate from the great mass of other information in LTS. This problem is accentuated when individual lists within a session must be kept separated. How this is managed is somewhat of a mystery. One possible explanation would call for a search along a temporal memory dimension: the individual items could be assumed to be temporally ordered, or to have "time tags." It is not enough to propose that search is made through all items indiscriminately and that items recovered from previous lists are recognized as such and not reported; if this were true, the duration and difficulty of the search would increase dramatically over the session. In fact, the usual result is that there is little change in performance over a session except for effects concentrated at the very start. On the other hand, judging from such factors as intrusion errors from previous lists, the subject is not able to restrict his search solely to the current list. In the experiments to follow, we will make the simplifying assumption, without real justification, that the

lists are entirely separated in LTS, and that the subject searches only through information relevant to the list currently being tested.

A. A Serial Display Procedure Involving Single Tests (Experiment 8)

This experiment involved a long series of discrete trials. On each trial a new display of items was presented to the subject. A display consisted of a random sequence of playing cards; the cards varied only in the

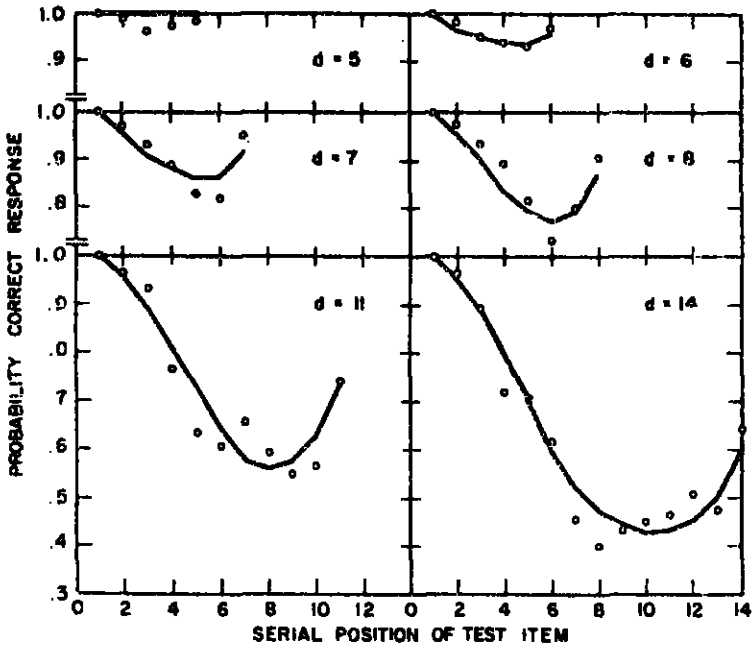


Fig. 25. Observed and theoretical probabilities of a correct response as a function of serial position (Experiment 8).

color of a small patch on one side; four colors (black, white, blue, and green) were used. The cards were presented to the subject at a rate of one card every 2 seconds. The subject named the color of each card as it was presented; once the color of the card had been named it was turned face down on a table so that the color was no longer visible, and the next card was presented. After presentation of the last card in a display, the cards were in a straight row on the table; the card presented first was to the subject's left and the most recently presented card to the right. The trial terminated when the experimenter pointed to one of the cards on the table and the subject attempted to recall the color of that card. The subject was instructed to guess the color if uncertain and to

qualify the response with a confidence rating. The confidence ratings were the numerals 1 through 4. The subjects were told to say 1 if they were positive; 2 if they were able to eliminate two of the four possible colors as being incorrect; 3 if one of the four colors could be eliminated as incorrect; and 4 if they had no idea at all as to the correct response.

It is important to note that only one position is tested in a display on each trial. The experiment involved 20 female subjects who participated in five daily sessions, each lasting for approximately 1 hour. Over the course of the five sessions, a subject was given approximately 400 trials. The display size, d, was varied from trial to trial and took on the following values: $d = 3, 4, 5, 6, 7, 8, 11$, and 14. Details of the experimental procedure are presented in Phillips, Shiffrin, and Atkinson (1967).

Figure 25 presents the probability of a correct response at each serial position for displays of size 5, 6, 7, 8, 11, and 14. For displays of sizes 3 and 4, the probability correct was 1.0 at all positions. The circles in the figure are the observed points; the solid lines are predicted curves which will be explained shortly. The serial positions are numbered so that item 1 designates the last item presented (the newest item), and item d designates the first item presented (the oldest item). The most apparent features of the curves are a fairly marked S-shaped recency portion and a smaller, quite steep primacy portion. For all display sizes, the probability of a correct response is 1.0 at serial position 1.

1. *Theory*

We must first decide whether a subject will set up and use a rehearsal buffer in this situation. Despite the fact that the continuous procedure has been dropped, it is still unlikely that the subject will engage in a significant amount of long-term coding. This is true because the task is still one of high "negative transfer"; the stimuli, which are the positions in the display, are constantly being re-paired with new responses as a session continues. Too much LTS encoding would undoubtedly lead to a high degree of interference among lists. It is only for a relatively weak and decaying LTS trace that a temporal search of long-term memory may be expected to keep the various lists separate. This difficulty in LTS transfer leads to the adoption of short-term strategies. Another reason for using a rehearsal buffer in this task depends upon the small list lengths employed; for small list lengths, there is a high probability that the item will be in the buffer at the moment of test. Thus the adoption of a rehearsal buffer is an efficient strategy. There is some question concerning just what the unit of rehearsal is in this situation. For example, the subject could assign numbers to positions in the display and then rehearse the number-color pairs. Most likely, however, the subject uses the fact that the stimuli always remain before her to

combine STS rehearsal with some form of visual mnemonic. That is, the unit of rehearsal is the response alone; as the subject rehearses the responses, she "mentally" places each response upon the appropriate card before her. This might therefore be a situation where the a-v-l and visual short-term stores are used in conjunction with each other. In any case, it seems reasonable that the units of rehearsal are the names (or perhaps the abbreviations) of the colors.

We might ask how the buffer will act in this situation. As noted earlier, in reference to the "overt-covert" experiment, the fact that items are read aloud as they are presented will tend to cause the subject to enter each item into the buffer. Furthermore, an S-shaped recency effect would not be unexpected. Indeed, if the units of rehearsal are the responses themselves, then the subject might tend to keep them in consecutive order to ease the visual memory task; if all items enter the buffer and are kept in consecutive order, then the oldest items will tend to be deleted first. That is, when a new item enters the buffer there will be a tendency to eliminate the oldest item from the buffer to make room for it. One other question that should be considered is the size of the buffer the subject would be expected to use in this task. There are a number of reasons why the buffer size should be larger here than in the continuous tasks of Section IV. First, the subject is not continually being interrupted for tests as in the previous studies; more of the subject's attention may therefore be allotted to rehearsal. Second, rehearsal of color names (or their abbreviations) is considerably easier than number-letter combinations. Equivalent to rehearsing "32-G, 45-Q" might be "Black, White, Black, Green" (or even a larger set if abbreviations are used). The magnitude of the difference may not be quite as large as this argument would lead us to expect because undoubtedly some time must be allotted to keeping track of which response goes on which position, but the estimate of the buffer size nevertheless should be larger in this situation than in the continuous tasks.

The STS part of the model for this experiment is similar to that used in the "overt" experiment in Section IV,D in that every item is entered in the buffer when it is presented. There is one new factor, however, that must be considered. Since each trial starts with the buffer empty, it will be assumed that the first items presented enter the buffer in succession, without knocking any item out, until the buffer is filled. Once the buffer is filled, each item enters the buffer and knocks out one of the items currently there. If the most recently presented item is in slot r of the buffer, and the oldest item is in slot 1, then the probability that the item in slot i of the buffer will be the one eliminated is

$$\frac{\delta(1-\delta)^{i-1}}{1-(1-\delta)^{r}}.$$

This is the same equation that was used to describe the knock-out process for the overt-covert study (Experiment 4). The larger δ, the greater the tendency to delete the oldest item in the buffer when making room for a new one.

The first set of long-term storage and retrieval assumptions that will be considered are essentially identical to those used in the previous sections. Information will be assumed to enter LTS during the entire period an item resides in the buffer at a rate θ per inter-item interval. This process must be qualified with regard to the first few items presented on each trial before the buffer is filled; it is assumed that the subjects divide their attention equally among the items in the buffer. Thus, if the rate of transfer is θ when there is only one item in the buffer, and the buffer size is r, then the rate of transfer will be θ/r when the buffer is filled. That is, since attention must be divided among r items when the buffer is full, each item receives only $1/r$th as much transfer as when the buffer only holds a single item. In general, information on each item will be transferred to LTS at rate θ/j during the interval in which there are j items in the buffer. The effect of this assumption is that more information is transferred to LTS about the items first presented in a list than about later items that are presented once the buffer is full.

The LTS decay and retrieval processes must now be examined. In earlier experiments we assumed that information decayed solely as a result of the number of items intervening between study and test; in other words, only the retroactive interference effect was considered. Because the previous tasks were continuous, the number of items preceding an item's presentation was effectively infinite in all cases. For this reason the proactive effects were assumed to be constant over conditions and did not need explicit inclusion in the model. In the present experiment the variation in list size makes it clear that proactive interference effects within a trial will be an important variable. The assumption that will be used is perhaps the simplest version of interference theory possible: each preceding and each succeeding item has an equal interfering effect. To be precise, if an amount of information I has been transferred to LTS for a given item, then every other item in the list will interfere with this information to the extent of reducing it by a proportion τ. Thus, if there were d items in the list, the item of interest would have an amount of information in LTS at the time of test equal to $I(\tau^{d-1})$. Clearly, the longer the list the greater the interference effect.

The model can now be completed by specifying the response process which works as follows. An item in the buffer at the time of test is responded to correctly. If the item is not in the buffer, then a search is made in LTS. The probability of retrieving the appropriate response is,

as in our other models, an exponential function of this information and equals $1 - \exp[-I(\tau^{d-1})]$; if a retrieval is not made, then the subject guesses.

2. *Data Analysis*

The parameter values that gave the best fit to the data of Fig. 25 using a minimum χ^2 criterion were as follows: $r = 5$, $\delta = .38$, $\theta = 2.0$, and $\tau = .85$.[20] Remember that r is the buffer size, δ determines the probability of deleting the oldest item in the buffer, θ is the transfer rate to LTS, and τ is the proportional loss of information caused by other items in the list. The theoretical curves generated by these parameter estimates are shown in Fig. 30 as solid lines. The predictions are quite accurate as indicated by a χ^2 value of 44.3 based on 42 degrees of freedom. It should be emphasized that the curves in the figure were all fit simultaneously with the same parameter values.

The primacy effect in the curves of Fig. 25 is predicted because more information is transferred to LTS for the first items presented on each trial. There are two reasons for this. First, the transfer rate on any given item is higher the fewer items there are in the buffer; thus the initial items, which enter the buffer before it is filled, accumulate more information in LTS. Second, the initial items cannot be knocked out of the buffer until the buffer is filled; thus the time period that initial items reside in the buffer is longer on the average than the time for later items. The recency effect is predicted because the last items presented in a list tend to be still in the buffer at the time of test; the S-shape arises because the estimate of δ indicates a fairly strong tendency for the oldest items in the rehearsal buffer to be eliminated first when making room for a new item.

Having estimated a set of parameter values that characterizes the data in Fig. 25, we now use these estimates to predict the confidence rating data. Actually, it is beyond the scope of this paper to analyze the confidence ratings in detail, but some of these data will be considered in order to illustrate the generality of the model and the stability of the parameter estimates. The data that will be considered are presented in Fig. 26; graphed is the probability of giving confidence rating R_1 (most confident) for each list size and each serial position. The observed data is represented by the open circles. It is clear that these results are similar in form to the probability correct curves of Fig. 25. The model used to fit these data is quite simple. Any item in the buffer is given an R_1. If the item is not in the buffer, then a search is made of LTS. If the amount of information in LTS on the item is $I(\tau^{d-1})$ then the probability of giving R_1 is an exponential function of that information: namely the

[20] For details on the method of parameter estimation see Phillips, Shiffrin, and Atkinson (1967).

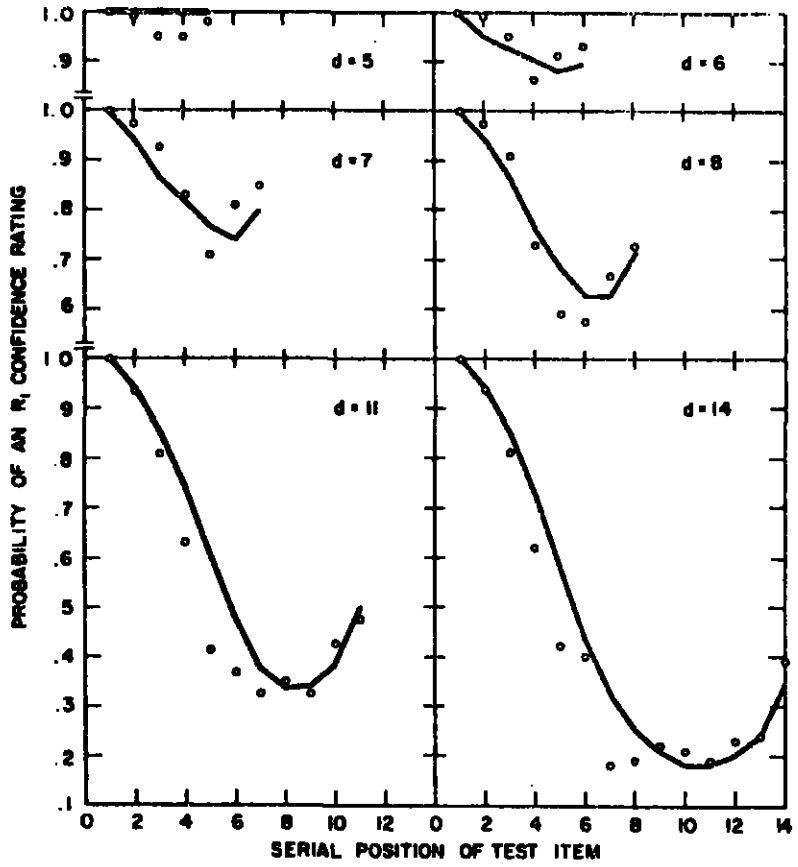


FIG. 26. Observed and predicted probabilities of confidence rating R_1 as a function of serial position (Experiment 8).

function $1 - \exp[-c_1 I(\tau^{d-1})]$, where c_1 is a parameter determining the subject's tendency to give confidence rating R_1. This assumption is consistent with a number of different viewpoints concerning the subject's generation of confidence ratings. It could be interpreted equally well as an assignment of ratings to the actually perceived amount of information in LTS, or as a proportion of the items that are recovered in an all-or-none fashion.[21] In any event, the predictions were generated using the previous parameter values plus an estimate of c_1. The predicted curves, with c_1 equal to .66, are shown in Fig. 26. The predictions are not as accurate as those in Fig. 25; but, considering that only one new parameter was estimated, they are quite good.

[21] The various possibilities may be differentiated through an analysis of conditional probabilities of the ratings given correct and incorrect responses, and through ROC curve (Type II) analyses (Bernbach, 1967, Murdock, 1966) but this will not be done here.

3. *Discussion*

In developing this model a number of decisions were made somewhat arbitrarily. The choice points involved will now be considered in greater detail. The assumption that the amount of transfer to LTS is dependent upon the number of items currently in the buffer needs elaboration. Certainly if the subject is engaged in coding or other active transfer strategies, the time spent in attending to an item should be directly related to the amount of transfer to LTS. On the other hand, the passive type of transfer which we assume can occur in situations where the subject makes use of a rehearsal buffer may not be related to the time spent in rehearsing an item per se, but rather to the total period the item resides in the buffer. That is, direct attention to an item in STS may not be necessary for some transfer to take place; rather a passive form of transfer may occur as long as the item remains in STS. Thus in situations where the rehearsal buffer is used and active transfer strategies such as coding do not occur, it could reasonably be expected that the amount of information transferred to LTS would be related solely to the total time spent in the buffer, and not to the number of items in the buffer at the time. In practice, of course, the actual transfer process may lie somewhere between these two extremes. Note that even if the transfer rate for an item is assumed to be a constant (unrelated to the number of items currently in the buffer) the first items presented for study still would have more information transferred to LTS than later items; this occurs because the items at the start of a list will not be knocked out of the buffer until it is filled and hence will reside in the buffer for a longer time on the average than later items. For this reason, the primacy effect could still be explained. On the other hand, the primacy effect will be reduced by the constant transfer assumption; in order to fit the data from the current experiment with this assumption, for example, it would be necessary to adjust the retrieval scheme accordingly. In modeling the free verbal-recall data that follows, a constant transfer assumption is used and accordingly a retrieval scheme is adopted which amplifies more strongly than the present one small differences in LTS strength.

We now consider the decay assumption in greater detail. The assumption is that the information transferred to LTS for a particular item is reduced by a proportion τ for every other item in the list. There are a number of possibilities for the form of this reduction. It could be actual physical interference with the trace, or it could be a reduction in the value of the current information as a result of subsequent incoming information. An example of this latter kind of interference will be helpful. Suppose, in a memory experiment the first item is GEX-5, and the subject stores "G_-5" in LTS. If tested now on GEX, the subject

would give the correct response 5. Suppose a second item GOZ-3 is presented and the subject stores "G_-3" in LTS. If he is now tested on either GEX or GOZ his probability of a correct response will drop to .5. Thus the actual information stored is not affected, but its value is markedly changed.

The assumption that every other item in a list interferes equally is open to question on two counts. First of all, it would be expected that an item about which a large amount of information is transferred would interfere more strongly with other items in LTS than an item about which little information is transferred. Certainly when no transfer occurs for an item, that item cannot interfere with other LTS traces. However, the equal interference assumption used in our analysis may not be a bad approximation. The second failing of the equal interference assumption has to do with separation of items. If the list lengths were very long, it might be expected that the number of items separating any two items would affect their mutual interference; the greater the separation, the less the interference. The list lengths are short enough in the present experiment, however, that the separation is probably not an important factor.

4. *Some Alternative Models*

It is worth considering some alternatives to the interference process of the model just presented, henceforth referred to as Model I in this subsection. In particular it is important to demonstrate that the effects of the interference-decay assumption, which could be viewed as a structural feature of memory, can be duplicated by simple search processes. For example, any limited search through the information in LTS will give poorer performance as the amount of that information increases. In order to make the concept of the search process clear, Model II will adopt an all-or-none transfer scheme. That is, a single copy of each item may be transferred to LTS on a probabilistic basis. If a copy is transferred, it is a perfect copy to the extent that it always produces a correct response if it is retrieved from LTS. The short-term features of the model are identical to those of Model I: each item enters the buffer; when the buffer is filled each succeeding item enters the buffer and knocks out an item already there according to the δ-process described earlier.

The transfer assumption for Model II is as follows. If an item is one of the j items in the buffer, then the probability that a copy of that item will be placed in LTS between one item's presentation and the next is θ/j. Therefore, the transfer depends, as in Model I, upon the number of other items currently in the buffer. No more than one copy may be placed in LTS for any one item. The retrieval assumptions are the

following. A correct response is given if the item is in the buffer when tested. If it is not in the buffer, then a search is made in LTS. If a copy of the item exists in LTS and is found, then a correct response is given; otherwise a random guess is made. As before, we assume that the information pertinent to the current list is distinguishable from that of earlier lists; thus, the search is made only among those copies of items in the current list. The central assumption of Model II is that exactly R selections are made (with replacement) from the copies in LTS; if the tested item has not been found by then, the search ends. The restriction to a fixed number of searches, R, is perhaps too strong, but can be justified if there is a fixed time period allotted to the subject for responding. It should be clear that for R fixed, the probability of retrieval decreases as the list length increases; the longer the list the more copies in LTS, and the more copies the less the probability of finding a particular copy in R selections. Model II was fit to the data in the same fashion as Model I. The parameter values that gave the best predictions were $r = 5$, $\delta = .39$, $\theta = .72$, and $R = 3.15$. The theoretical curves generated by these parameters are so similar to those for Model I that Fig. 25 adequately represents them, and they will not be presented separately. Whereas the χ^2 was 44.3 for Model I, the χ^2 value for Model II was 46.2, both based on 42 degrees of freedom. The similarity of the predictions serves to illustrate the primary point of introducing Model II: effects predicted by search processes and by interference processes are quite similar and consequently are difficult to separate experimentally.

The search process described above is just one of a variety of such mechanisms. In general there will be a group of possible search mechanisms associated with each transfer and storage assumption; a few of these processes will be examined in the next section on free-verbal-recall. Before moving on to these experiments, however, we should like to present briefly a decay and retrieval process combining some of the features of interference and search mechanisms. In this process the interference does not occur until the search begins and is then caused by the search process itself. The model (designated as Model III) is identical in all respects to Model II until the point where the subject begins the search of LTS for the correct copy. The assumption is that the subject samples copies with replacement, as before, but each unsuccessful search may disrupt the sought-after copy with probability R'. The search does not end until the appropriate copy is found or until all copies in LTS have been examined. If the copy does exist in LTS, but is disrupted at any time during the search process, then when the item is finally retrieved, the stored information will be such that the subject will not be able to recall at better than the chance level. The parameter values giving the best fit for this model were $r = 5$, $\delta = .38$, $\theta = .80$, and

$R' = .25$. The predicted curves are again quite similar to those in Fig. 25 and will not be presented. The predictions are not quite as accurate, however, as those of Models I and II, the χ^2 value being 55.0.[22]

B. Free-Verbal-Recall Experiments

The free-verbal-recall situation offers an excellent opportunity for examining retrieval processes, because the nature of the task forces the subject to engage in a lengthy search of LTS. The typical free-verbal-recall experiment involves reading a list of high-frequency English words to the subject (Deese & Kaufman, 1957; Murdock, 1962). Following the reading, the subject is asked to recall as many of the words as possible. Quite often list length has been a variable, and occasionally the presentation time per item has been varied. Deese and Kaufman, for example, used lists of 10 and 32 items at 1 second per item. Murdock ran groups of 10, 15, and 20 items at 2 seconds per item, and groups of 20, 30, and 40 items at 1 second per item. The results are typically presented in the form of serial position curves: the probability of recall is plotted against the item's position in the list. The Murdock (1962) results are representative and are shown in Fig. 27. It should be made clear that the numbering of serial positions for these curves is opposite from the scheme used in the previous section; that is, the first item presented (the oldest item at the time of test) is labeled serial position 1. This numbering procedure will be used throughout this section to conform with the literature on free-verbal-recall; the reader should keep this in mind when comparing results here with those presented elsewhere in the paper. The primacy effect in Fig. 27 is the rise on the left-hand portions of the curves and the recency effect is the larger rise on the right-hand portions of the curves. The curves are labeled with the list length and the presentation rate per item. Note that the curves are quite similar to those found in Experiment 8 of the previous section; an effect not seen in Experiment 8 (because of the short list lengths used) is the level asymptotic portions of the curves which appear between the primacy and recency effects for the longer lists.

The form of the curves suggests that a buffer process could explain the results, with the words themselves being the units of rehearsal. The recency effect would be due to the probability that an item is still in the buffer at test; this probability goes to near zero after 15 items or so and the recency effect accordingly extends no further than this. The primacy effect would arise because more information accrued in LTS for the first few items presented in the list. Whether a buffer strategy is reasonable in the free-recall situation, however, is worth further discussion. It can hardly be maintained that high-frequency English words are difficult to

[22] For a more detailed account of Models I, II, and III, and a comparison among models, see Atkinson and Shiffrin (1965).

code; on the other hand, the task is not a paired-associate one and cues must be found with which to connect the words. One possibility is that upon seeing each word the subject generates a number of associates (from LTS) and tries to store the group of words; later during testing a search which retrieves any of the associates might in turn retrieve the desired word. We tend to doubt that this strategy, used by itself, will greatly improve performance.[23] To the extent that coding occurs, it

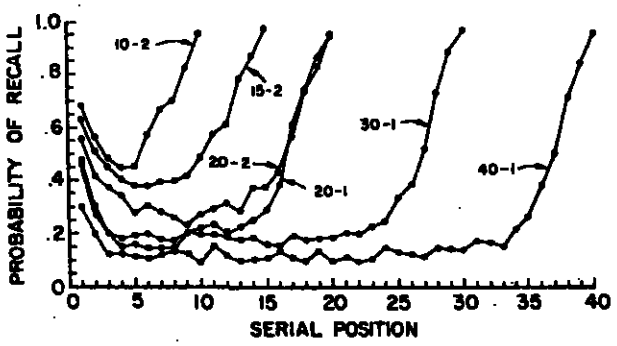


FIG. 27. Probability of correct recall as a function of serial position for free verbal recall. After Murdock (1962).

probably involves connecting words within the presented list to each other. This technique would of course require the consideration of a number of words simultaneously in STS and therefore might be characterized reasonably well by a buffer process. Whether or not coding occurs in the free-recall situation, there are other reasons for expecting the subjects to adopt a buffer strategy. The most important reason is undoubtedly the improvement in performance that a rehearsal buffer will engender. If the capacity of the buffer is, say, 4 or 5 words, then the use of a buffer will assure the subjects of a minimum of four or five items correct on each list (assuming that all of the items may be read out of the buffer correctly). Considering that subjects report on the average only about 8 or 9 items, even for long lists, the items stored in the buffer are an important component of performance.

It will be assumed, then, that the subjects do adopt a rehearsal strategy. The comparability of the curves in Fig. 25 to those in Fig. 27

[23] B. H. Cohen (1963) has presented free-recall lists containing closely related categories of words, e.g., north, east, south, west. Indeed, the recovery of one member of a category usually led to the recovery of other members, but the total number of categories recalled did not exceed the number of separate words recalled from noncategorized lists.

might indicate that a model similar to any of the models presented in the previous section could be applied to the current data. There are, however, important differences between the two experimental paradigms which must be considered: the free-recall situation does not involve pairing a response with a stimulus for each list position, and has the requirement of multiple recall at the time of test. The fact that explicit stimulus cues are not provided for each of the responses desired would be expected to affect the form of the search process. The multiple-response requirement raises more serious problems. In particular, it is possible that each response that is output may interfere with other items not yet recalled. The problem may be most acute for the case of items still in the buffer; Waugh and Norman (1965) have proposed that each response output at the time of test has the same disrupting effect upon other items in the buffer as the arrival of a new item during study. On the other hand, it is not clear whether a response emitted during test disrupts items in LTS. It might be expected that the act of recalling an item from LTS would raise that item's strength in LTS; this increase in strength is probably not associated, however, with the transfer of any new information to LTS. For this reason, other traces will most likely not be interfered with, and it shall be assumed that retrieval of an item from LTS has no effect upon other items in LTS.

Because there is some question concerning the effects of multiple recall upon the contents of the buffer, and because this section is primarily aimed at LTS processes, the part of the free-recall curves that arise from the buffer will not be considered in further analyses. This means that the models in this section will not be concerned with the part of the curve making up the recency effect; since the data in Fig. 27 indicate that the recency effect is contained in the last 15 items (to the right in the figure) of each list, these points will be eliminated from the analyses. Unfortunately, the elimination of the last 15 items means that the short list lengths are eliminated entirely. The problem of obtaining data for short list lengths not contaminated by items in the buffer at the time of test has been circumvented experimentally by a variation of the counting-backward technique. That is, the contents of the buffer can be eliminated experimentally by using an interfering task inserted between the end of the list and the start of recall. We now turn to a consideration of these experiments.

A representative experiment is that by Postman and Phillips (1965). Words were presented at a rate of one per second in all conditions. In one set of conditions three list lengths (10, 20, and 30) were used and recall was tested immediately following presentation. This, of course, is the usual free recall procedure. The serial position curves are shown in the top panel of Fig. 28 in the box labeled "0 second." The same list lengths

were used for those conditions employing an intervening task; immediately following presentation of the list the subjects were required to count backwards by three's and four's for 30 seconds. Following this intervening task, they were asked to recall the list. The results are shown in the lower panel in Fig. 28. If the intervening task did not affect the contents of LTS but did wipe out all items in the buffer, then the recency

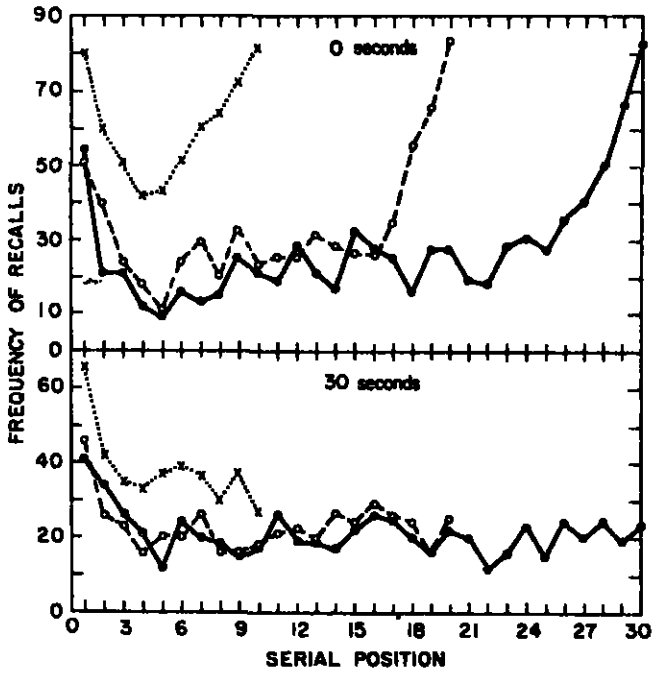


FIG. 28. Probability of correct recall as a function of serial position for free verbal recall with test following 0 seconds and 30 seconds of intervening arithmetic. After Postman & Phillips (1965).

effects would be expected to disappear with the curves otherwise unchanged. This is exactly what was found. The primacy effects and asymptotic levels remain unchanged while the recency effect disappears. It is clear, then, that normal free-recall curves (without intervening arithmetic) from which the last 15 points have been deleted should be identical to curves from experiments using intervening arithmetic. The following data have therefore been accumulated: Murdock's data with the last 15 points of each list deleted; data reported by Deese and Kaufman (1957) using a free-recall paradigm, but again with the last 15 points of each list deleted; the data reported by Postman and Phillips (1965); and some data collected by Shiffrin in which an intervening task

was used to eliminate the contents of the buffer.[24] All of these serial position curves have the same form; they show a primacy effect followed by a level asymptote. For this reason the results have been presented in Table 1. The first three points of each curve, which make up the primacy

TABLE 1

OBSERVED AND PREDICTED SERIAL POSITION CURVES FOR VARIOUS FREE-VERBAL-RECALL EXPERIMENTS

List	Point 1		Point 2		Point 3		Asymptote		Number of points
	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	
M-20-1	.46	.45	.27	.37	.20	.29	.18	.22	2
M-30-1	.38	.35	.30	.28	.21	.22	.19	.17	12
M-20-2	.55	.61	.42	.51	.37	.41	.31	.32	2
M-40-1	.30	.29	.20	.23	.13	.18	.12	.14	22
M-25-1	.38	.39	.23	.32	.21	.25	.15	.19	7
M-20-2.5	.72	.66	.61	.56	.45	.46	.37	.35	2
D-32-1	.46	.33	.34	.27	.27	.21	.16	.16	14
P-10-1	.60	.62	.42	.52	.35	.42	.34	.32	7
P-20-1	.47	.45	.27	.37	.23	.29	.22	.22	17
P-30-1	.41	.35	.34	.28	.27	.22	.20	.17	27
S-6-1	.71	.74	.50	.64	.57	.52	.42	.40	3
S-6-2	.82	.88	.82	.79	.65	.66	.66	.52	3
S-11-1	.48	.60	.43	.50	.27	.40	.31	.31	8
S-11-2	.72	.76	.55	.66	.52	.54	.47	.42	8
S-17-1	.55	.49	.33	.40	.26	.32	.22	.24	14
S-17-2	.68	.66	.65	.56	.67	.45	.43	.35	14

effect, are given in the table. The level portions of the curves are then averaged and the average shown in the column labeled "asymptote." The column labeled "number of points" is the number of points which have been averaged to arrive at the asymptotic level.[25] The column labeled "list" gives the abbreviation of the experimenter, the list length, and the presentation rate for each of the serial position curves. (M = Murdock, 1962; D = Deese and Kaufman, 1957; P = Postman and Phillips, 1965; S = Shiffrin.)

[24] The Shiffrin data are reported in more detail in Atkinson and Shiffrin (1965).

[25] For the Postman-Phillips and Shiffrin lists the number of points at asymptote are simply list length, d, minus 3. For the Murdock and the Deese-Kaufman lists the number of points is $d - 15 - 3$ because the last 15 points in these lists have been eliminated.

1. *Theoretical Analysis*

Having accumulated a fair amount of parametric data in Table I, we should now like to predict the results. The first model to be considered is extremely simple. Every item presented enters the subject's rehearsal buffer. One by one the initial items fill up the buffer, and thereafter each succeeding item knocks out of the buffer a randomly chosen item. In conditions where arithmetic is used following presentation, it is assumed that the arithmetic operations knock items from the buffer at the same rate as new incoming items. This is only an approximation, but probably not too inaccurate. Information is assumed to be transferred to LTS as long as an item remains in the buffer, in fact as a linear function of the total time spent in the buffer (regardless of the number of other items concurrently in the buffer). If an item remains in the buffer for j seconds, an amount of information equal to θ times j is transferred to LTS. Call the amount of information transferred to LTS for an item its *strength*. When the subject engages in a search of LTS during recall it is assumed that he makes exactly R searches into LTS and then stops his search (the number of searches made might, for example, be determined by the time allowed for recall). On each search into LTS the probability that information concerning a particular item will be found is just the ratio of that item's strength to the sum of the strengths of all items in the list. Thus, items which have a greater LTS strength will be more likely to be found on any one search. The probability that the information in LTS will produce a correct recall, once that information has been found in a search, is assumed to be an exponential function of the strength for that item.

There are just three parameters for this model: r, the buffer size; θ, the parameter determining the rate per second at which information on a given item is transferred to LTS while the item resides in the rehearsal buffer; and R the number of searches made.[26] The probability of a correct response from the buffer is zero for the results in Table I because the contents of the buffer have been emptied experimentally by intervening arithmetic, or because the recency data (which represents recovery from the buffer) has been omitted. The parameters giving the best fit to the data were as follows: $r = 4$, $\theta = .04$, and $R = 34$. The predictions also are presented in Table I. The predictions are rather remarkable considering that just three parameters have been used to predict the results from

[26] It is important to remember that θ for this model is defined as the *rate per second* of information transfer, and thus the time measures listed in Table I need to be taken into account when applying the model. For example, an item that resides in the buffer for three item presentations will have 3θ amount of information in LTS if the presentation rate is one item per second, and 7.5θ if the presentation rate is 2.5 seconds per item.

four different experiments employing different list lengths and different presentation rates. Some of the points are not predicted exactly but this is largely due to the fact that the data tends to be somewhat erratic; the predictions of the asymptotic values (where a larger amount of data is averaged) are especially accurate.

2. *Some Alternative Models*

A number of decisions were made in formulating the free-recall model that need to be examined in greater detail. First consider the effect of an arithmetic task upon items undergoing rehearsal. If the arithmetic caused all rehearsal and long-term storage operations to cease immediately, then the probability of recalling the last item presented should decrease toward chance (since its LTS strength will be negligible, having had no opportunity to accumulate). The serial position curve, however, remains level and does not drop toward the end of the list. One possible explanation is that all transfer to LTS takes place when the item first enters the buffer, rather than over the period the item remains in the buffer; in this case the onset of arithmetic would not affect the formation of traces in LTS. While this assumption could handle the phenomenon under discussion, we prefer to consider the LTS trace as building up during the period the item remains in the buffer. Recall that this latter assumption is borne out by the accuracy of the earlier models and, in particular, the U-shaped functions presented in Fig. 12 for the multiple-reinforcement experiment. The explanation of the level serial position curve implied by our model is that the arithmetic operations remove items from the buffer in a manner similar to that of new entering items. Two sources give this assumption credibility. First, Postman and Phillips (1965) found that short periods of arithmetic (15 seconds) would leave some of the recency effect in the serial position curve, suggesting that some items remained in the buffer after brief periods of arithmetic. Second, the data of Waugh and Norman (1965) suggest that output operations during tasks such as arithmetic act upon the short-term store in the same manner as new incoming items.

Another choice point in formulating the model occurred with regard to the amount of LTS transfer for the first items in the list. The assumption used in an earlier model let the amount of transfer depend upon the number of other items concurrently undergoing rehearsal, as if the attention allotted to any given item determines the amount of transfer. An alternative possibility is that the amount of transfer is determined solely by the length of stay in the buffer and is therefore independent of the number of items currently in the buffer. Another assumption resulting in this same independence effect is that the subject allots to items in the buffer only enough attention to keep them "alive"; when

the number of items in the buffer is small, the subject presumably uses his spare time for other matters. A free-verbal-recall experiment by Murdock (1965) seems to support a variant of this latter assumption. He had subjects perform a rather easy card-sorting task during the presentation of the list. The serial position curve seemed unaffected except for a slight drop in the primacy effect. This would be understandable if the card-sorting task was easy enough that the buffer was unaffected, but distracting enough that extra attention normally allotted to the first few items in the list (before the buffer is filled) is instead allotted to the card-sorting task. In any case, it is not clear whether the transfer rate should or should not be tied to the number of items concurrently in the buffer. The model that we have proposed for free-recall (henceforth referred to as Model I in this subsection) assumed a constant transfer process; a model using a variable transfer assumption will be considered in a moment.

The search process used in Model I is only one of many possibilities. Suppose, for example, that the strength value for an item represents the number of bits of information stored about that item (where the term "bits" is used in a nontechnical sense). A search might then be construed as a random choice of one bit from all those bits stored for all the items in the list. The bits of information stored for each item, however, are associated to some degree, so that the choice of one bit results in the uncovering of a proportion of the rest of the information stored for that item. If this proportion is small, then different searches finding bits associated with a particular item will result in essentially independent probabilities of retrieval. This independent retrieval assumption was used in the construction of Model I. On the other hand, finding one bit in a search might result in all the bits stored for that item becoming available at once; a reasonable assumption would be that this information is either sufficient to allow retrieval or not, and a particular item is retrieved the first time it is picked in a search or is never retrieved. This will be called the dependent retrieval assumption.

It is interesting to see how well the alternate assumptions regarding transfer and search discussed in the preceding paragraphs are able to fit the data. For this reason, the following four models are compared:[27]

Model I: Transfer to LTS is at a constant rate θ regardless of the number of other items concurrently in the buffer, and independent retrieval.

Model II: Transfer to LTS is at a variable rate θ/j where j is the number of other items currently in the buffer, and independent retrieval.

Model III: Constant LTS transfer rate, and dependent retrieval.

[27] These models and the related mathematics are developed in Atkinson and Shiffrin (1965).

Model IV: Variable LTS transfer rate, and dependent retrieval. Model I, of course, is the model already presented for free-verbal-recall. The four models were all fit to the free-verbal-recall data presented in Table I, and the best fits, in terms of the sums of the squared deviations, were as follows: Model I: .814; Model II: 2.000; Model III: .925; Model IV: 1.602 (the lowest sum meaning the best predictions). These results are of interest because they demonstrate once again the close interdependence of the search and transfer processes. Neither model employing a variable transfer assumption is a good predictor of the data and it seems clear that a model employing this assumption would require a retrieval process quite different from those already considered in order to fit the data reasonably well.

Perhaps the most interesting facet of Model I is its ability to predict performance as the presentation rate varies. A very simple assumption, that transfer to LTS is a linear function of time spent in the buffer, seems to work quite well. Waugh (1967) has reported a series of studies which casts some light on this assumption; in these studies items were repeated a variable number of times within a single free-recall list. The probability of recall was approximately a linear function of the number of repetitions; this effect is roughly consonant with an assumption of LTS transfer which is linear with time. It should be noted that the presentation rates in the experiments we analyzed to not vary too widely: from 1 to 2.5 second per item. The assumption that the subject will adopt a buffer strategy undoubtedly breaks down if a wide enough range in presentation rates is considered. In particular, it can be expected that the subject will make increasing use of coding strategies as the presentation rate decreases. M. Clark and G. Bower (personal communication) for example, have shown that subjects proceeding at their own pace (about 6–12 seconds a word) can learn a list of 10 words almost perfectly. This memorization is accomplished by having the subject make up and visualize a story including the words that are presented. It would be expected that very slow presentation rates in free-recall experiments would lead to coding strategies similar to the one above.

One last feature of the models in this section needs further examination. Contrary to our assumption, it is not true that successive lists can be kept completely isolated from each other at the time of test. The demonstration of this fact is the common finding of intrusion errors: items reported during recall which had been presented on a list previous to the one being tested. Occasionally an intrusion error is even reported which had not been reported correctly during the test of its own list. Over a session using many lists, it might be expected that the interference from previous lists would stay at a more or less constant level after the presentation of the first few lists of the session. Nevertheless,

the primacy and asymptotic levels of the free-recall serial position curves should drop somewhat over the first few lists. An effect of this sort is reported by Wing and Thomson (1965) who examined serial position curves for the first, second, and third presented lists of a session. This effect is undoubtedly similar to the one reported by Keppel and Underwood (1962); namely, that performance on the task used by Peterson and Peterson (1959) drops over the first few trials of a session. The effects in both of these experiments may be caused by the increasing difficulty of the search process during test.

C. Further Considerations Involving LTS

The models presented in the last section, while concerned with search and retrieval processes, were nevertheless based primarily upon the concept of a rehearsal buffer. This should not be taken as an indication that rehearsal processes are universally encountered in all memory experiments; to the contrary, a number of conditions must exist before they will be brought into play. It would be desirable at this point then to examine some of the factors that cause a subject to use a rehearsal buffer. In addition, we want to consider a number of points of theoretical interest that arise naturally from the framework developed here. These points include possible extensions of the search mechanisms, relationships between search and interference processes, the usefulness of mnemonics, the relationships between recognition and recall, and coding processes that the subject can use as alternatives to rehearsal schemes.

Consider first the possible forms of search mechanisms and the factors affecting them. Before beginning the discussion two components of the search process should be emphasized: the first component involves locating information about an item in LTS, called the "hit" probability; the second component is the retrieval of a correct response once information has been located. The factor determining the form of the search is the nature of the trace in long-term store. The models considered thus far have postulated two different types of traces. One is an all-or-none trace which allows perfect recall following a hit; the other is an unspecified trace which varies in strength. The strength notion has been used most often because it is amenable to a number of possible interpretations: the strength could represent the "force" with which a particular bond has been formed, the number of bits of information which have been stored, or the number of copies of an item placed in memory. It should be emphasized that these different possibilities imply search processes with different properties. For example, if the strength represents the force of a connection, then it might be assumed that there is an equal chance of hitting any particular item in a search, but the

probability of giving a correct answer following a hit would depend upon the strength. On the other hand, the strength might represent the number of all-or-none copies stored in LTS for an item, each copy resulting in a correct response if hit. In this case, the probability of a hit would depend upon the strength (the number of copies) but any hit would automatically result in a correct answer. A possibility intermediate to these two extremes is that partial copies of information are stored for each item, any one partial copy allowing a correct response with an intermediate probability. In this case, the probability of a hit will depend on the number of partial copies, and the probability of a correct response following a hit will depend on the particular copy that has been found. A different version of this model would assume that all the partial copies for an item become available whenever any one copy is hit; in this version the probability of a correct answer after a hit would depend on the full array of copies stored for that item. In all the search processes where the retrieval probability following a hit is at an intermediate level, one must decide whether successive hits of that item will result in independent retrieval probabilities. It could be assumed, for example, that failure to uncover a correct response the first time an item is hit in the search would mean that the correct response could not be recovered on subsequent hits of that item.[28] This outline of some selected search processes indicates the variety of possibilities; a variety which makes it extremely difficult to isolate effects due to search processes from those attributable to interference mechanisms.

Other factors affecting the form of the search are at least partially controlled by the subject; a possible example concerns whether or not the searches are made with replacement. Questions of this sort are based upon the fact that all searches are made in a more or less ordered fashion; memory is much too large for a completely random search to be feasible. One ordering which is commonly used involves associations: each item recovered leads to an associate which in turn leads to another associate. The subject presumably exercises control over which associates are chosen at each stage of the search and also injects a new starting item whenever a particular sequence is not proving successful.[29] An alternative to the associate method is a search along some partially ordered dimension. Examples are easy to find; the subject could generate letters of the

[28] For a discussion of partial and multiple copy models, see Atkinson and Shiffrin (1965).

[29] Associative search schemes have been examined rather extensively using free-recall methods. Clustering has been examined by Deese (1959), Bousfield (1953), Cofer (1966), Tulving (1962), and others; the usual technique is to determine whether or not closely associated words tend to be reported together. The effect certainly exists, but a lack of parametric data makes it difficult to specify the actual search process involved.

alphabet, considering each in turn as a possible first letter of the desired response. A more general ordered search is one that is made along a temporal dimension; items may be time-tagged or otherwise temporally ordered, and the subject searches only among those items that fall within a particular time span. This hypothesis would explain the fact that performance does not markedly deteriorate even at the end of memory experiments employing many different lists, such as in the free-verbal-recall paradigm. In these cases, the subject is required to respond only with members of the most recent list; if performance is not to degenerate as successive lists are presented, the memory search must be restricted along the temporal dimension to those items recently stored in LTS. Yntema and Trask (1963) have demonstrated that temporal information is available over relatively long time periods (in the form of "time-tags" in their formulation) but the storage of such information is not well understood.

We now turn to a brief discussion of some issues related to interference effects. It is difficult to determine whether time alone can result in long-term interference. Nevertheless, to the extent that subjects engage in a search based upon the temporal order of items, interference due to the passage of time should be expected. Interference due to intervening material may take several forms. First, there may be a reduction in the value of certain information already in LTS as a result of the entry of new information; the loss in this case does not depend on making any previous information less accessible. An example would be if a subject first stores "the simulus beginning with D has response 3" and later when another stimulus beginning with D is presented, he stores "the stimulus beginning with D has response 1." The probability of a correct response will clearly drop following storage of the second trace even though access to both traces may occur at test. Alternatively, interference effects may involve destruction of particular information through interaction with succeeding input. This possibility is often examined experimentally using a paired-associate paradigm where the same stimulus is assigned different responses at different times. DaPolito (1966) has analyzed performance in such a situation. A stimulus was presented with two different responses at different times, and at test the subject was asked to recall both responses. The results indicated that the probability of recalling the first response, multiplied by the probability of recalling the second response, equals the joint probability that both responses will be given correctly. This result would be expected if there was no interaction of the two traces; it indicates that high strengths of one trace will not automatically result in low strengths on the other. The lack of an interaction in DaPolito's experiment may be due to the fact that subjects knew they would be tested on both responses. It is

interesting to note that there are search mechanisms that can explain this independence effect and at the same time interference effects. For example, storage for the two items might be completely independent as suggested by DaPolito's data; however, in the typical recall task the subject may occasionally terminate his search for information about the second response prematurely as a result of finding information on the first response.

Within the context of interference and search processes, it is interesting to speculate about the efficacy of mnemonics and special coding techniques. It was reported, for example, that forming a visual image of the two words in a paired-associate item is a highly effective memory device; that is, one envisages a situation involving the two words. Such a mnemonic gains an immediate advantage through the use of two long-term systems, visual and auditory, rather than one. However, this cannot be the whole explanation. Another possibility is that the image performs the function of a mediator, thereby reducing the set of items to be searched; that is, the stimulus word when presented for test leads naturally to the image which in turn leads to the response. This explanation is probably not relevant in the case of the visual-image mnemonic for the following reason: the technique usually works best if the image is a very strange one. For example, "dog-concrete" could be imaged as a dog buried to the neck in concrete; when "dog" is tested, there is no previously well-learned association that would lead to this image. Another explanation involves the protection of the stored information over time; as opposed to the original word pairs, each image may be stored in LTS as a highly distinct entity. A last possibility is that the amount of information stored is greatly increased through the use of imagery—many more details exist in the image than in the word pair. Since the image is highly cohesive, the recovery of any information relevant to it would lead to the recovery of the whole image. These hypotheses are of course only speculations. At the present time the relation of the various search schemes and interference processes to mnemonic devices is not well understood. This state of affairs hopefully will change in the near future since more research is being directed toward these areas; mediation, in particular, has been receiving extensive consideration (e.g., Bugelski, 1962; Runquist & Farley, 1964).

Search processes seem at first glance to offer an easy means for the analysis of differences between recognition and recall. One could assume, for example, that in recall the search component which attempts to locate information on a given item in LTS is not part of the recognition process; that is, one might assume that in recognition the relevant information in LTS is always found and retrieval depends solely on matching the stored information against the item presented for test.

Our analysis of free-verbal recall depended in part upon the search component to explain the drop in performance as list length increased. Thus if the free recall task were modified so that recognition tests were used, the decrement in performance with list length might not occur. That this will not be the case is indicated by the position-to-color memory study (Experiment 8) in which the number of responses was small enough that the task was essentially one of recognition; despite this fact, the performance dropped as list length increased. One possible explanation would be that search is necessary even for recognition tasks; i.e., if the word "clown" is presented, all previous times that that word had been stored in LTS do not immediately spring to mind. To put this another way, one may be asked if a clown was a character in a particular book and it is necessary to search for the appropriate information, even though the question is one of recognition. On the other hand, we cannot rule out the possibility that part of the decrement in performance in free recall with the increase of list length may be due to search changes, and part to other interference mechanisms. Obviously a great deal of extra information is given to the subject in a recognition test, but the effect of this information upon search and interference mechanisms is not yet clear.

We now turn to a consideration of LTS as it is affected by short-term processes other than the rehearsal buffer. It has been pointed out that the extent and structure of rehearsal depends upon a large number of factors such as the immediacy of test and difficulty of long-term storage. When rehearsal schemes are not used in certain tasks, often it is because long-term coding operations are more efficacious. These coding processes are presumably found in most paired-associate learning paradigms; depending upon conditions, however, the subject will probably divide *his attention* between coding and rehearsal. Atkinson and Shiffrin (1965) have presented a paired-associate learning model based upon a rehearsal-buffer. Whether a rehearsal strategy would be adopted by the subject in a given paired-associate learning experiment needs to be determined in each case. The answer is probably no for the typical fixed-list learning experiment, because the items are usually amenable to coding, because the test procedure emphasizes the importance of LTS storage, and because short study-test intervals are so infrequent that maintenance of an item in STS is not a particularly effective device. If these conditions are changed, however, then a paired-associate model based upon a rehearsal buffer might prove applicable.

It is important to note the distinction between coding models and rehearsal models. Rehearsal models actually encompass, in a rough sense, virtually all short-term processes. Coding, for example, may be considered as a type of rehearsal involving a single item. The buffer

process is a special type of rehearsal in which a fixed number of items are rehearsed for the primary purpose of maintaining them in STS. A pure coding process is one in which only a single item is considered at a time and in which the primary purpose is the generation of a strong LTS trace; almost incidentally, the item being coded will be maintained in STS through the duration of the coding period, but this is not a primary purpose of the process. These various processes, it should be emphasized, are under subject control and are brought into play as he sees fit; consequently there are many variations that the subject can employ under appropriate conditions. One could have a coding model, for example, in which more than one item is being coded at a time, or a combination model in which several items are maintained via rehearsal while one of the items is selected for special coding.

At the other extreme from the buffer strategy, it might be instructive to consider a coding process that acts upon one item at a time. Although such a process can be viewed as a buffer model with a buffer containing only one item, the emphasis will be upon LTS storage rather than upon the maintenance of the item in STS. The simplest case occurs when the presentation rate is fairly slow and the subject attempts to code each item as it is presented for study. However, the case that seems most likely for the typical paired-associate experiment, is that in which not every item is coded, or in which it takes several presentation periods to code a single item. The first case above could be conceptualized as follows: each item is given a coding attempt during its presentation interval, but the probability of finding a code is ξ. The second case is a bit more complex. One version would have a single item maintained in STS over trials until a code is found. It could be supposed that the probability of a code being found during a single presentation interval is ξ; having once coded an item, coding attempts are focused on the next presented item. This model has something in common with the buffer models in that some items will remain in STS over a period of several trials. This will produce a short-term decay effect as the interval between presentation and test is increased.

It is worth considering the form of the usual short-term effects that are found in a paired-associate learning. Figure 29 presents data from a paired-associate experiment by Bjork (1966). Graphed is the probability of a correct response for an item prior to its last error, as a function of the number of other items intervening between its study and subsequent test. The number of intervening items that must occur before this curve reaches the chance level can be taken as a measure of the extent of the short-term effect. It can be seen that the curve does not reach chance level until after about 20 items have been presented. If the coding model mentioned above were applied to this data, a short-term effect would be

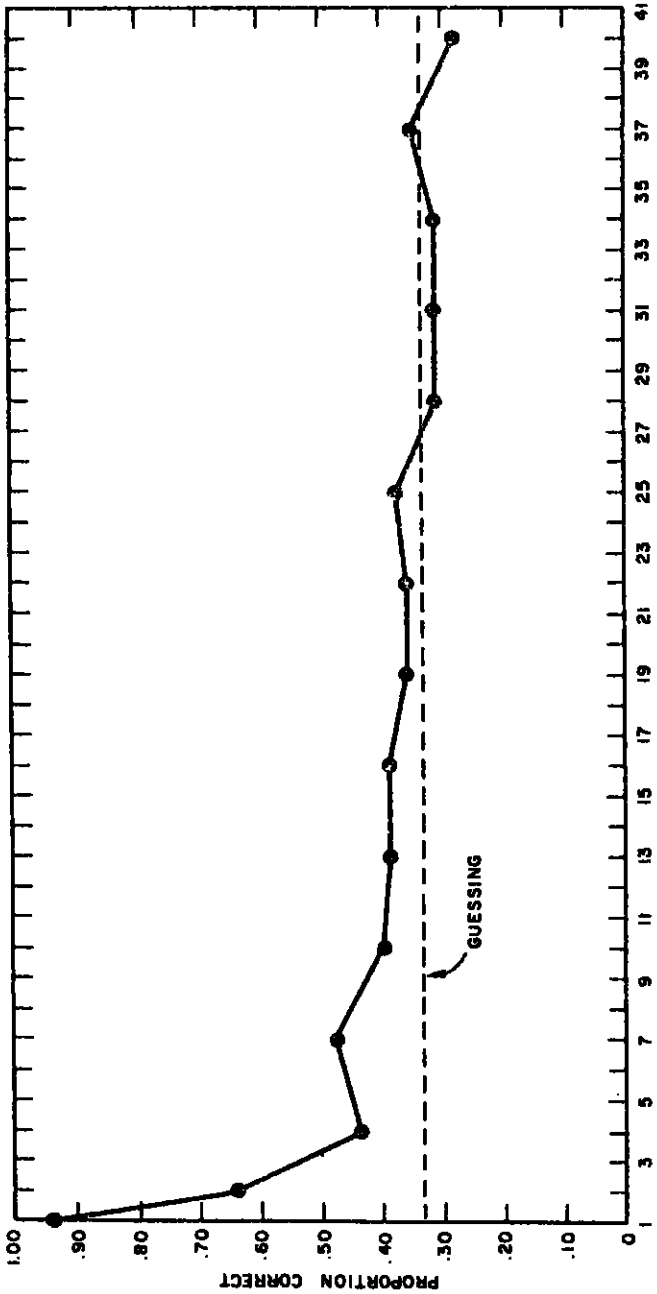


FIG. 29. Probability of a correct response prior to the last error as a function of lag. After Bjork (1966).