Psychological Review
1969, Vol. 76, No. 2, 179–193

STORAGE AND RETRIEVAL PROCESSES IN LONG-TERM MEMORY[1]

R. M. SHIFFRIN[2] AND R. C. ATKINSON

Stanford University

A theory of human memory is described in which a distinction is made between three memory stores: the sensory register, the short-term store, and the long-term store. Primary emphasis is given to the processes by which information is stored in and retrieved from the long-term store, a store which is considered to be a permanent repository for information. Forgetting and related phenomena are attributed to a failure of the retrieval process, in which the search through some memory area becomes less efficient as new information is placed in it. Storage and retrieval in the long-term store are conceived of as parallel processes, one mirroring the other, and each is divided into three stages for conceptual clarity. The memory trace is viewed as an ensemble of information stored in some memory location, the location of storage determined largely by the components of the ensemble itself. The ability of the system to cope with diverse phenomena is demonstrated by a consideration of a number of selected experimental paradigms.

Theories of human long-term memory have given primary emphasis either to the organization of that memory, in terms of the dimensions of storage and the associations between the stored information (e.g., Cofer, 1965; Deese, 1966; Mandler, 1968; Osgood, 1963), or to the characteristics of temporal decay from that memory as in the interference theories (e.g., Keppel, 1968; Melton, 1963; Postman, 1961; Underwood, 1957). The processes by which information is stored in, and retrieved from, long-term memory have been relatively neglected. An example of this type of process is the memory search during retrieval; during the search, a succession of memory codes is examined, each examination followed by a decision process in which the search is either terminated or continued, and in which information recovered is either accepted as that desired and output, or rejected. It is the intention of this paper to elaborate the memory input and output processes. As will be indicated later, our view of storage and retrieval eliminates the necessity for assuming decay of information from long-term memory. It will be assumed that long-term memory is permanent; decrements in performance over time are ascribed to an increasingly ineffective search of the stored information.

We begin by describing the overall conception of the memory system. The system follows that described in Atkinson and Shiffrin (1965, 1968a), and is similar to those proposed by Feigenbaum (1966) and Norman (1968). The major components of the system are diagrammed in Figure 1: the sensory register, the short-term store (STS) and the long-term store (LTS). The solid arrows in the diagram represent directions in which information is transferred from one part of the system to another. Note that transfer is not meant to imply the removal of information from one store and the placing of it in the next; rather, transfer is an operation in which information in one store is "copied" into the next without affecting its status in the original store. It should be emphasized that our hypotheses about the various memory stores do not require any assumptions regarding the physiological locus of these stores; the system is equally con-

[1] Preparation of this article was supported by a grant from the National Aeronautics and Space Administration, and is an outgrowth of ideas first developed in two earlier reports (Atkinson & Shiffrin, 1968b; Shiffrin, 1968). This paper, in combination with papers by Atkinson and Shiffrin (1965, 1968a), represents an attempt to formulate a general schema within which to analyze memory and learning.

[2] Now at Indiana University. Requests for reprints should be sent to Richard C. Atkinson, Institute for Mathematical Studies in the Social Sciences, Ventura Hall, Stanford University, Stanford, California 94305.

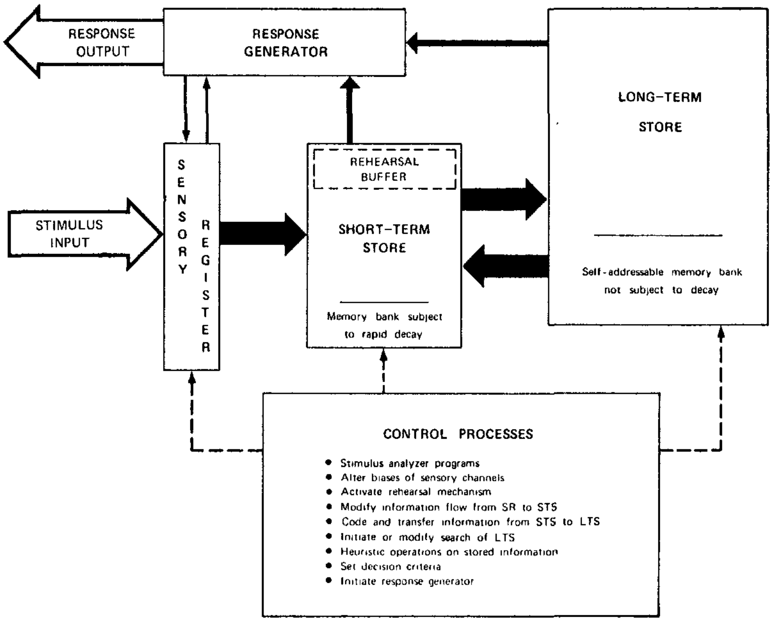

FIG. 1. A flow chart of the memory system. (Solid lines indicate paths of information transfer. Dashed lines indicate connections which permit comparison of information arrays residing in different parts of the system; they also indicate paths along which control signals may be sent which activate information transfer, rehearsal mechanisms, etc.)

sistent with the view that the stores are separate physiological structures as with the view that the short-term store is simply a temporary activation of information permanently stored in the long-term store. The control processes listed in Figure 1 are a sample of those which the subject (*S*) can call into play at his discretion, depending upon such factors as the task and the instructions. Control processes govern informational flow, rehearsal, memory search, output of responses, and so forth.

The sensory register is a very short-lived memory store which temporarily holds incoming sensory information while it is being initially processed and transferred to the short-term store. In the visual modality, for example, information will decay from the sensory register in a period of several hundred milliseconds (Sperling, 1960). Information in the short-term store, if not attended to by *S*, will decay and be lost in a period of about 30 seconds or less, but control processes such as rehearsal can maintain information in STS for as long as *S* desires (the buffer process in Figure 1 is one highly organized rehearsal scheme). While information resides in STS, portions of it are transferred to LTS. The long-term store is assumed to be a *permanent* repository of information; we realize that factors such as traumatic brain damage, lesions, and deterioration with extreme age must lead to memory loss, but such effects should be negligible in the types of experiments considered in this paper. Thus it is hypothesized that information, once stored in LTS, is never thereafter destroyed or eliminated. Nevertheless, the ability to retrieve information from LTS varies considerably with time and interfering material.

The short-term store serves a number of useful functions. On the one hand it decouples the memory system from the external environment and relieves the system from the responsibility of moment-to-moment attention to environmental changes. On the other hand, STS provides a working memory in which manipulations of information may take place on a temporary basis. Because STS is a memory store in which information can be maintained if desired, it is often used as the primary memory device in certain types of tasks; in these tasks the information presented for retention is maintained in STS until the moment of test and then emitted. Tasks in which STS is utilized for this purpose, and the mechanisms and control processes that may come into play, have been examined extensively in Atkinson and Shiffrin (1968a). In this report we are primarily interested in STS as a temporary store in which information is manipulated for the purposes of storage and retrieval from LTS, rather than as a store in which information is maintained until test. In the remainder of this paper, discussion is limited to that component of memory performance which involves LTS retrieval, and the components arising from STS and the sensory register will not be considered.

LONG-TERM STORE

In describing the structure of LTS, an analogy with computer memories is helpful. The usual computer memory is "location addressable"; if the system is given a certain location it will return with the contents of that location. When given the contents of a word (a "word" refers to a single computer memory location), such a system must be programmed to examine each location in turn in order to find the possible locations of these contents in the memory. It seems untenable that an exhaustive serial search is made of all of LTS whenever retrieval is desired. An alternative type of memory may be termed "content-addressable"; if the system is given the contents of a word it will return with the locations in memory containing those contents. One way in which such a memory may be constructed utilizes a parallel search through all memory locations; the system then returns with the locations of all matches. If this view is adopted, however, an additional process is needed to select the desired location from among the many returned by the parallel search. Thus, if we feed the system the word "red," it would not be useful for the system to return with all references or locations of "red"; there are simply too many and the original retrieval problem would not be significantly reduced in scale. There is, however, an alternative method for forming a content-addressable memory; in this method, the contents to be located themselves contain the information necessary to specify the storage location(s). This can occur if the information is originally stored in locations specified by some master plan dependent upon the contents of the information. Such a system will be termed "self-addressing." A self-addressing memory may be compared with a library shelving system which is based upon the contents of the books. For example, a book on "caulking methods used for 12th century Egyptian rivercraft" will be placed in a specific library location (in the Egyptian room, etc.). If a user desires this book, it may be located by following the same shelving plan used to store it in the first place. We propose that LTS is to a large degree just such a self-addressable memory. An ensemble of information presented to the memory system will define a number of memory areas in which that information is likely to be stored; the memory search will therefore have certain natural starting points. The system is assumed to be only partially self-addressing in that the degree to which the storage locations are specified will vary from one ensemble to the next and one moment to the next, in much the way as proposed in stimulus sampling theory (Estes, 1959). Thus it may be necessary to embark upon a memory search within the specified locations, a search which may proceed serially from one location to the next. This conception of LTS leads to a number of predictions. For example, a recognition test of memory will not proceed via exhaustive scanning of all stored codes, nor will a recognition test eliminate in all cases the necessity for an LTS search. If information is presented and *S* must indicate whether this information has been presented previously, then the likely storage location(s) is queried. To the degree that the information has highly salient characteristics which precisely identify the storage location, the extent of the LTS search will be reduced. Thus, for items with highly salient characteristics, *S* should be able to identify quickly and accurately whether the item was presented previously, and the identification might not require a memory search which interrogates more than a single storage location. The less well-specified the storage location, the greater the memory search needed to make an accurate recognition response. This view is in many respects similar to that proposed by Martin (1967), in which he suggests that certain stimuli tend to give rise to identical encoding responses from one instance to the next, that is, on both study and test trials. From the present viewpoint "identical encoding responses" is taken to be equivalent to identical storage locations.

Although it is assumed that information ensembles are sorted into LTS locations according to a master plan, it is beyond the scope of this paper to attempt to outline this organizational structure; other workers have been dealing with this problem (e.g., Mandler, 1968; Pollio, 1966). Undoubtedly this organization is highly complex, but the form of the organization will not be crucial to our discussions of input and output mechanisms. There is one dimension of organization that must be mentioned here, however; this is the "temporal" dimension. We are assuming that this dimension is like the other organizational dimensions in that information may be stored along it, and that retrieval may be based upon it. In the remaining sections of this paper, the temporal dimension will often be singled out for special mention, not because it differs in substance from other organizational dimensions, but because it is virtually impossible to eliminate it as a systematic variable in most memory experiments.

The term "location" is used in relation to the organizational schema; an LTS location is defined by the place in the organizational structure occupied by an information ensemble. The location will be defined in terms of the modality of the information (e.g., visual versus auditory), the level of analysis, (e.g., spelling versus syntactic structure), and all other dimensions of organization that may be relevant. Two locations will be said to be "close" if the information in them tends to be retrieved together. In particular, we shall refer to a *code,* or an *image,* as an ensemble of information that is closely related and very likely to be retrieved together. We do not wish to imply that there is some unitary atom of storage called a code or an image. The information making up a code in one task may be considered to be several codes in a different task. For example, an entire sentence may be considered a code if we are comparing the meaning of that sentence with others; however, the same sentence might be considered to be made up of a series of codes if we are comparing it with sentences of the same meaning but different grammatical form. Nevertheless, for most tasks the concept of a code or image as representing a cohesive array of information in a single storage location proves useful. This view of the memory trace is consistent with the analysis given by Bower (1967). In his model, a trace, or image, is represented by a vector of attributes which serves to identify both the information contained in the trace and the location of the trace in LTS. The values of the attributes in the vector serve to indicate the position of the trace on the various organizational dimensions; for example, one of the positions in the vector might indicate whether the trace is auditory or visual. It should be apparent that this quantitative view is compatible with our description of an image or code, and also with the description of LTS as a self-addressable store.

STORAGE AND RETRIEVAL

Since LTS is self-addressing, storage and retrieval have many features in common, one process mirroring the other. Storage is assumed to consist of three primary mechanisms: *transfer, placement,* and *image-production.* The transfer mechanism includes those control processes by which *S* decides what to store, when to store, and how to store information in LTS. The placement mechanism determines the locations in which the ensemble of information under consideration will be stored. To a large degree, the components of the ensemble itself will determine the location of storage. That is, in the action of encoding the desired information for storage, *S* may supplement the information currently in STS with pertinent information retrieved from LTS; the resultant ensemble in STS determines the storage location. The image-production mechanism determines what proportion of the current ensemble of information in STS will be placed in the designated LTS location(s). The proportion stored should be a function of the duration of the period that the ensemble is maintained in STS. Retrieval, like storage, is assumed to consist of three primary mechanisms: *search, recovery,* and *response generation.* The search process is a recursive loop in which locations or images are successively selected for examination. As each image is examined, the recovery process determines how much information will be recovered from the image and placed in STS. The response generation process then examines the recovered information and decides whether to continue the search or terminate and emit a response. If the search does not terminate, the selection of the next location or image for examination may depend upon information already uncovered during the search.

Although storage and retrieval are treated separately in this paper, we do not wish to imply that these processes are separated in time, one following the other. Rather, long-term storage is continually occurring for the information residing in short-term store. In addition, retrieval is continually occurring during storage attempts by *S;* for example, *S* may try to store a paired-associate by searching LTS for prominent associations to the stimulus, associations which could then be used as mediators.

The remaining portions of this report will be directed toward a delineation of these input and output processes. It is beyond the scope of this paper to do more than briefly describe the major mechanisms of the theory. Nevertheless, an attempt will be made to indicate how these processes work, what evidence supports our assumptions, what predictions may be derived, and how the theory may be applied to a number of selected tasks.

Storage Processes

Transfer. The decisions involving what to store, when to store, and how to store information are under a high degree of control by *S,* and therefore can result in striking performance changes from one task to another. Although we shall be discussing experimental situations in which storage takes place largely during designated study periods, we should note that storage in general occurs whenever information is cycled through the short-term store; for example, in reflective thinking, in daydreaming, and in moment-to-moment consideration of the day's happenings.

The decision concerning when to store information is especially important in situations where a large amount of information is being input rapidly to STS. Such a situation taxes the capacity of the system and forces *S* to select a subset of the presented information for special attention and coding. There are a number of factors that determine the information so selected. Important or easily stored information is likely to be given preference. For example, Harley (1965) has shown in a mixed-list design that paired-associate items given high monetary payoffs are selectively attended in preference to items given low payoffs. Information selection will also be governed by the degree to which the incoming material is already known. For example, there is considerable evidence that *S*s tend to store more information on a given item if that item's presentation is followed by other items that are well-known as opposed to being followed by items that are not known (Atkinson, Brelsford, & Shiffrin, 1967; Thompson, 1967). Finally, note that the decision when to store may be determined by strategies associated with the list or task as a whole, rather than with the individual item. For example, in an experiment employing a paired-associate list with only the two responses A and B, *S* might decide to store only information about those stimuli paired with Response A, and always guess Response B if the answer is not known at test.

The decisions concerning how to store information will also affect performance: storage via visual images may be more effective than auditory storage (Schnorr & Atkinson, 1969); overt and covert rehearsal methods may result in very different effects (Brelsford & Atkinson, 1968); and mediating versus nonmediating instructions may give rise to considerable performance differences (Runquist & Farley, 1964). On a somewhat different level, *S* may engage in organizational storage strategies such that different items are stored in locations determined by some fixed organizational framework (Tulving, 1962).

What information is stored, given the presentation of a particular information ensemble, will also be highly dependent upon the control processes utilized by *S*. In some cases, as Underwood (1963) points out in making a distinction between the nominal and functional stimulus, not all the information contained in the presented item is necessary for correct responding (e.g., if the stimuli are all nonsense syllables differing only in the first letter, then only the first letters need be stored). In these cases only the relevant characteristics of the input need be stored. In most cases *S* will select relevant characteristics of the presented information and add to this additional coding information from LTS; for example, a paired-associate plus a mediator might be stored in many cases. One type of information which *S* often attempts to store is information indicating that a particular response (usually one just given in error) is not correct; Millward (1964) has examined evidence for this process in simple paired-associate tasks. This process of tagging responses as incorrect is particularly important in studies of negative transfer; the higher the probability that the first response assigned to a stimulus will be tagged as incorrect (when the response changes), the less will be the proactive interference effect observed in the data (Shiffrin, 1968).

The examples of transfer mechanisms given above are by no means exhaustive, but they serve to indicate the pervasiveness and importance of these processes. Relatively simple decisions to select one transfer scheme rather than another can and do lead to large performance effects. These facts emphasize the need for carefully establishing the transfer mechanisms used before extending analysis to other aspects of the data.

Placement. The location in which an image will be stored is determined by the contents of that image; *S* therefore controls the storage location by manipulating the information complex in STS. For any given ensemble of information, however, there is a certain amount of randomness in placement. For this reason, a search for an image may have to be undertaken at test even if the entire information ensemble originally present in STS during study is presented for consideration. Information to be remembered may be stored in images in more than one location; for example, a paired-associate may be encoded by the use of two entirely different mnemonics. This notion has been given quantitative form in "multiple-copy" models for the memory trace (Atkinson & Shiffrin, 1965).

The primary mechanism determining storage location is in all cases an organizational framework. The self-addressing characteristic of input information depends upon the prior, already established, LTS organization; each item is sent to locations that depend upon the preexisting organizational framework. However, in many cases (especially situations involving free-recall learning and paired-associate learning in which the same list of items is presented over a series of trials) it is an effective technique for *S* to form new organizational structures. In free-recall learning, for example, output on successive trials becomes increasingly organized and similar to output on the preceding trial (Cofer, 1965; Tulving, 1962); it may be inferred that a consistent new organization has been imposed upon a set of words which were disorganized at the start of learning. This point regarding growth of organization is especially important with regard to list-structured paired-associate tasks, that is, a task in which a list of paired-associates is presented over and over on successive trials (possibly in a new random order from trial to trial). In such a task, organizational effects between items will tend to occur over trials, and care must be taken in inferring that effects seen by averaging data over all items in the list will also apply to the individual items. For example, there is evidence indicating that interference effects found by averaging over items in a list do not apply to the individual items making up the list (DaPolito, 1966; Greeno, 1967). For this reason we shall often consider in the rest of this paper results from what is termed a "continuous" paired-associate task (Bjork, 1966; Brelsford, Shiffrin, & Atkinson, 1968; Rumelhart, 1967). This type of task employs a long series of presentations, each involving first a test and then a study on a paired-associate item. The character of the experiment is essentially homogeneous over time because new items are continually being introduced and old items deleted; a new item may appear at any point in the sequence, receive a fixed number of presentations distributed over a subset of trials, and then be dropped and replaced by yet another item. In this way the list-structure feature of the typical paired-associate experiment is eliminated, and it is extremely difficult for *S* to develop special schemes for interitem organization.

In most cases, *S*'s placement strategy involves choosing one of many preexisting organizational dimensions for storage. In these cases, an organizational clue contained in the experimental design may prove useful. In categorized free-recall, for example, *S* will be induced to store (and retrieve) in the given categories (Bousfield & Cohen, 1956; Cofer, 1965; Cohen, 1963). Thus the location in which the word *division* will be stored will be quite different if preceded by *multiplication, addition,* and *subtraction* than if preceded by *platoon, regiment,* and *battalion.*

In the interests of effective memory, the most important requirement of the placement process is that it results in a storage location which will later be searched during test. If organizational schemes are used for placement, the schemes themselves will have to be stored, recovered, and utilized in order for retrieval to be optimal.

Image-production. When an ensemble of information is present in STS, some portion of it will be stored in a designated location in LTS as a permanent image. The proportion of information that is stored will be a function of the duration of time that the ensemble stays in STS, or of the number of times that ensemble is cycled through STS. It would be most natural to look for evidence of this mechanism in experiments varying study time for particular items. However, improved performance with longer study times can be attributed to a higher chance of finding a good mnemonic. Better evidence is found in experiments in which *S*s are induced to utilize rote rehearsal rather than coding, and in which longer periods of rote verbal rehearsal lead to improved performance (Brelsford & Atkinson, 1968; Hellyer, 1962).

To conclude the discussion of storage, we consider the content of the image: the range and form of the stored information. A single image may contain a wide variety of information, including characteristics of the item presented for study (its sound, meaning, color, size, shape, position, etc.) and characteristics added by *S* (such as codes, mnemonics, mediators, images, associations, etc.). In addition, an image usually will contain links to other images (other information which was in the short-term store at the same time); these links can be regarded as a set of directions to the locations of related images in LTS.

Retrieval Processes

The retrieval mechanism forms the crux of the present theory, since it enables a permanent long-term store to exhibit the characteristics of a failing memory. The basic mechanism by which memory loss occurs involves a partially random search through an increasingly large set of images in some local set of memory locations. This local area may be defined by one or more dimensions in the organizational structure of LTS; as the number of images in the local area increases, the search for the desired information will become increasingly ineffective. For example, consider two areas, one consisting of three images, the other consisting of 10 images, and both containing the desired code. A random search through the smaller set will result in successful retrieval in a shorter period of time than a search through the larger set.

The retrieval process begins with the presentation of an information complex which places constraints on the response desired and also provides a number of clues, or delimiting information, concerning that desired output (i.e., a stimulus). On the basis of this presented information, or as the result of an external search strategy, *S* is led to look in some local memory area and select (possibly randomly) an image for examination. The process by which information is recovered from this image is called *recovery.* The recovered information will be placed in the short-term store which may also contain other information such as the search strategy being employed, salient information recovered previously in the search, the LTS locations that have been examined already, and some of the links to other images that have been examined already, and some of the links to other images that have been noted in the search but not yet examined. The short-term store thus acts as a "window" upon LTS, allowing *S* to deal sequentially with a manageable amount of information. The current contents of STS are now examined and various decisions made concerning whether the desired response has been found, whether to emit it, whether to terminate the search unsuccessfully, or whether to continue the search. These decisions and the generation of the response are called the *response-generation* process. If a decision is made to continue the search, then a new location is selected either randomly, on the basis of information just recovered, or in accord with an overriding external search strategy. The process which continues cycling in this manner until termination is called the *search* process.

Search. Each cycle in the search recursion begins with a mechanism which locates the next image for examination. This mechanism may be separated into *directed* and *random* components. The directed component includes strategies controlled by *S* and depends upon the input information and the self-addressing nature of the system. As a result of the directed component, a number of locations in some area of memory are marked for examination. The locations and images so marked will be referred to as the *examination subset,* and the directed component of the search may be characterized by the probability that the sought-after code is in the examination subset.

The directed component of the search can be either under a high degree of *S* control or relatively automatic. For example, the control factor is high in a situation where *S* engages in a first-letter alphabetic search, or where *S* attempts to remember lunch of 2 days previously by reconstructing the events of that day. In cases like this, the search assumes many of the aspects of problem solving. At the other extreme is the almost completely automatic direction of search which occurs, say, in the attempted recognition of a word as having been presented earlier in the session. In such a case, the word itself serves to direct the search via the self-addressing feature of the system.

In its most general form the search process can be viewed as a series of stages in which searches are made successively in different examination subsets. That is, after choosing codes randomly for a time from one subset, information recovered during the search may cause *S* to change the memory area currently being examined for one quite different. This type of search could occur, for example, in a categorized free-recall task, in which the various categories are searched successively (Cohen, 1963). In most applications, however, a restricted search model should be adequate, the restriction allowing only a single examination subset to be searched during any retrieval period. This model should be particularly applicable to tasks in which only a short response period is allowed.

The random component of the search specifies the selection of codes in the examination subset from one cycle of the search to the next. It is assumed that the likelihood of any particular code being selected at some point will depend upon the amount of information contained in it, as well as the total number of codes and their temporal ordering in the examination subset. It will be seen that the extent to which the search depends upon the temporal ordering of the items is an important variable which, among other things, determines interference effects.

The division of search into directed and random components leads to somewhat different memory models, depending upon which component is selected for elaboration. Feigenbaum (1966) and Hintzman (1968), for example, have elaborated upon the directed component in computer simulation models. In these formulations there is a mechanism termed a "discrimination net," which enables *S* to sort through the organizational structure of LTS to reach the local memory area where desired information is stored. On the other hand, there are models which emphasize the probabilistic search through an examination subset of items in some memory area (Atkinson & Shiffrin, 1965). It is upon this latter type of model that attention will be focused in this paper.

Recovery. Once an image has been located, it is appropriate to ask what information contained in the image will be entered into the short-term store. This process is called recovery. The amount of information recovered from an image is assumed to be probabilistic, depending upon the current noise level in the system and the amount of information in the image. In particular, the amount of information recovered should be an increasing function of the amount of information in the image.

Response generation. Having recovered information from LTS, *S* is faced with decisions as to whether to terminate the search and respond or to continue to search. These decisions must first of all depend upon the consistency of the recovered information with that indicated by the test information. Inconsistent information can be ruled out at once. If information consistent with the test stimulus is found, and a recognition response is desired, then the response will be given when temporal or contextual information is recovered indicating that the image was stored recently. In a paired-associate test, three types of information may be recovered: that associated with the stimulus, that associated with the response, or associative information linking the two. Recovery of stimulus or response information will serve to lead to recognition of either; if both are recovered in nearby locations in memory, then the response may be emitted. On the other hand, a response might be output following recovery of associative information linking the response to the stimulus. Since the ability to output a response depends upon the amount of response information recovered, the process may often be represented by a decision-theoretic model in which *S* is attempting to filter information through a noisy background (Bernbach, 1967; Kintsch, 1967; Wickelgren & Norman, 1966). In situations where a long time period exists for responding, a likely response may be recovered but not emitted; in these cases the search will be continued in the hope of finding a better response. An extended search of this kind leads naturally to predictions that *S* will be able to rank responses in the order of their probability of being correct, with responses ranked after the first being correct at an above-chance level (Binford & Gettys, 1965).

A decision can be made to terminate the search unsuccessfully if response time has run out, or if response time is expected to run out and *S* wishes to make a guess before it does, or if *S* decides that further search would not be useful. Termination schemes of the latter type are quite varied. One simple rule would terminate the search when all images in the examination subset have been interrogated unsuccessfully. Other schemes would end the search when some fixed time limit expired, or some fixed number of items examined. These schemes are described in more detail in Atkinson and Shiffrin (1965).

The extensive decision structure that has been outlined may make it seem unlikely that time is available during the search for examination of a very large number of images. This is indeed the point of view we adopt: in most cases it is assumed that only a few images will be examined in the search before termination. For example, in a continuous paired-associate task analyzed by Shiffrin (1968), the estimated number of codes examined prior to termination was from about one to five.

APPLICATIONS OF THE SYSTEM

Forgetting

Decrements in performance occur in the system as a result of the input of additional information to LTS. These decrements result from three related mechanisms. First is a mechanical effect; information sufficient to respond correctly at one point in time may prove inadequate after additional information has been added. For example, a paired-associate GAX-4 may be stored as G**-4, and this code will be sufficient for correct responding (if recovered) when GAX is tested. Suppose, however, that GEK-3 is now presented and stored as G**-3. When either of these stimuli is tested later, both codes may be retrieved from LTS and therefore S will have to guess whether the correct response is 3 or 4.

The second cause of forgetting arises from a breakdown in the directed component of the search mechanism. That is, correct retrieval requires that the same memory area be searched at test as was used for storage during study. This may not occur, however, if only a portion of the input information is used to direct storage during study, for a different portion might be utilized to locate the storage area during retrieval. This process could be viewed within the framework of stimulus sampling theory (Estes, 1959) if the stimulus elements are taken to represent dimensions of organization. For clarity, let us denote the image which encodes the correct response for the current test as a "*c* code," and denote the other codes as "*i* codes." Thus the *i* codes are irrelevant codes which should lead to intrusion errors, whereas the *c* code, if examined, should lead to a correct response. Then the directed component of search can be characterized by the probability that the *c* code is in the examination subset, called p_c. In experiments in which clues are available to denote the organizational dimensions to be searched, p_c may be close to 1.0. In other situations, such as continuous tasks with randomly chosen stimuli and responses, p_c will be lower and dependent upon such factors as the amount of information in the *c* code and its age (where "age" denotes the position of the code on the temporal dimension). Although the breakdown in the directed component can provide a reasonable degree of forgetting, we shall focus primarily upon the third mechanism of forgetting: the increasing size of the examination subset.

When searching the examination subset, there are a number of possible results. The *c* code may be examined and give rise to a correct response, one of the *i* codes may be examined and produce an intrusion response, or none of the codes may give rise to a response and the search terminates. If the search through the examination subset is at least partially random, then the following conclusions may be reached. When the size of the subset is increased (i.e., the number of *i* codes is increased), then the probability of giving an intrusion will increase, the average time until the *c* code is examined will increase, and the probability of giving a correct response will decrease. When we say that the order of search is *partially* random, we mean to imply that the order in which codes in the examination subset are selected for consideration may depend upon both the amount of information in the code and the age of the code. Clearly, as the amount of information in a code tends toward zero, or as the age of a code increases, the probability of examining that code early in the search should decrease.

In order to make the sequence of events in the search clear, an example of a search is presented in Figure 2 for a continuous paired-associate task. In this task, suppose that on successive trials stimulus-response pairs are presented; on each trial the stimu-

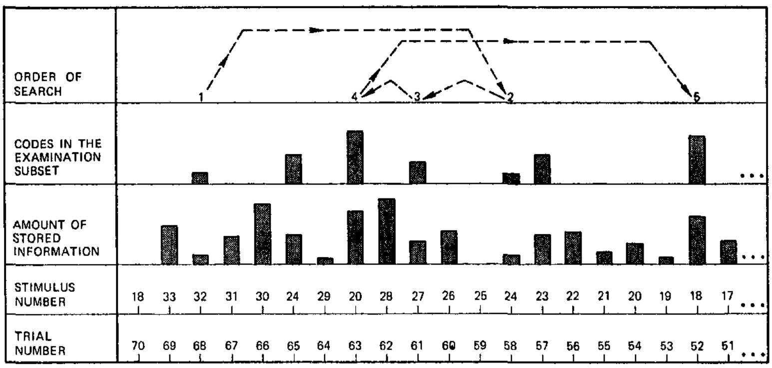

FIG. 2. A schematic representation of an LTS search in a continuous memory task.

lus is tested first and then the pair is studied. A new pair may be presented on any trial, and given several reinforcements and tests at varying intervals. The bottom row in the figure gives the trial number for a section of the task from Trials 51–70. The number of the stimulus-response pairs tested and studied is given in the second row. The third row shows which pairs have had codes stored in LTS, and the height of the bar indicates the amount of information in the code. The fourth row indicates those codes that are in the examination subset on Trial 70 when Stimulus 18 is presented for test. For example, the stimulus tested may have begun with a vowel, and the items in the examination subset could have stimulus components which also begin with a vowel. Note that in this example the code for Stimulus 18 was stored on Trial 52, and happens to be in the examination subset. The top row of the table gives the order of search through the subset; the first four codes were examined and rejected but the fifth code examined was the *c* code and enabled the search to end with a correct response. It may be seen how forgetting occurs if it is imagined that Item 18 had not been tested until Trial 90 or 100. In this event, more *i* codes would be present in the examination subset and the probability would be greater that an intrusion would occur, or that the search would terminate before the *c* code was examined.

Interference

Various interference phenomena are readily predicted by the system. Although in general the order of search through the examination subset will depend upon the age and amount of information in the codes, suppose for simplicity that the search order is entirely random. Then both nonspecific proactive and retroactive interference effects are predicted, and in a sense are predicted to be equal. That is, extra *i* codes in the examination subset added either temporally before or after the *c* code will cause the correct retrieval probability to drop; in the case of a random search the probability decrease will be the same whether caused by an *i* code preceding or following the *c* code. The drop occurs because the extra codes increase the amount of time required to find the *c* code. Therefore, if the size of the examination subset is increased, it is more likely that either response time will run out or an intrusion response will occur. Obviously the greater the degree to which the search is ordered temporally backwards from the most recent item, the less the proactive, and the greater the retroactive interference effect. Thus if codes are examined strictly in temporal order, the average amount of time until the *c* code is examined will be independent of the number of codes which are older than the *c* code and hence no proactive effect will be expected. One of the best places to examine nonspecific proactive and retroactive interference effects is in the study of free-verbal recall as a function of list length (Murdock, 1962). In this task a list of words is read to *S*, who attempts to recall as many of them as possible following their presentation. The data are usually graphed as a serial-position curve which gives the probability of correct recall as a function of the presentation position (and hence as a function of the number of preceding and succeeding items in the list). As the list length is varied, the number of preceding and succeeding items is systematically varied and it is possible to apply the theory to the resultant data. The application of the theory is made particularly easy in this case because *S* is trying to recall all of the list, and hence the examination subset can be assumed to consist of all codes that have been stored. In fact, a model derived from the theory has been applied to free-verbal recall data as a function of list length and has proved remarkably successful (Atkinson & Shiffrin, 1965, 1968a). The model assumes that performance decreases with list length because more codes are missed in the memory search at longer list lengths. This model also predicts that some of the missed codes should be retrieved if a second recall test is given following the first. Just such an effect was found by Tulving (1967), and its magnitude is predicted accurately by the model. Three successive recall tests were given following a single list presentation and only 50% of the items recalled were recalled on all three tests, even though the actual number recalled remained constant over the three tests.

Item-specific interference is also readily predicted via the search mechanism. Item interference refers to that condition arising when a stimulus originally paired with one response is later paired with a second, different response. In this case two different codes with the same stimulus may be placed in LTS. Thus, the amount of proactive or retroactive interference will depend upon the number of times the wrong code is examined and accepted prior to the correct code. In particular, the degree of temporal ordering in the search will affect the relative amounts of proactive and retroactive effects. That is, the greater the degree that the search is ordered temporally backwards from the most recent item, the less proactive and the greater retroactive effect is predicted. The reasoning is similar to that in the previous paragraph for nonspecific effects. This rather simple view of interference is complicated by at least two factors. First, if *S* is aware that he will eventually be tested for both responses, he may link them in nearby codes, or in a single code, and thereby reduce interference effects (Dallet & D'Andrea, 1965). Second, when the first response is changed, *S* may tag the first code with the information that the response is now wrong. If the first code is later recovered during search, then this information will enable him to inhibit an intrusion and continue the search; an effect like this was found by Shiffrin (1968).

We might ask how this view of interference phenomena compares with traditional theories such as the various "two-factor" interference theories (Melton & Irwin, 1940; Postman, 1961; Underwood, 1957). In a number of respects, the present system differs sharply with the traditional views; for example, in the assumption of a permanent store. It might be expected in the present framework, where memory is permanent, that interference effects which appear under one form of test (say recall) would be reduced under less stringent tests (say recognition) whenever the less stringent test succeeds in making both the old and new codes available. Evidence of this sort has been found by McGovern (1964). Note that the present search system does not necessitate the introduction of such processes as proactive and retroactive inhibition and spontaneous recovery, each with associated changes over time. The effects accounted for by these processes are easy to handle within the search framework, at least in continuous tasks of the type presented in Figure 2. In list-structured tasks, however, there is room for considerable complication, in that learning of lists allows for organization and retrieval schemes based on the list as a whole. Thus, when the first-list responses to stimuli are all changed in a second list, the organizational strategy or the retrieval scheme, (in the present terms, the "directed" search component) may be the mechanism which is disturbed, and it may be this disturbance which is described by the traditional interference theories. Indeed, there is evidence from list-structured tasks that interference effects found over lists as a whole may *not* be related to individual stimulus-response assignments within those lists (DaPolito, 1966; Greeno, 1967). In any event, when quantitative models derived from the present theory have been applied to data (Atkinson & Shiffrin, 1968a; Shiffrin, 1968), the search scheme outlined here handled forgetting and interference effects in a parsimonious and accurate manner.

Intrusions

Another useful feature of the model is its natural prediction of intrusions, and of variations in intrusion rates over differing conditions. In a paired-associate task, an intrusion occurs when the response contained in an *i* code is recovered and emitted. Actually, the intrusion process has not yet been specified clearly, since both the probability of being in the examination subset and the probability of accepting the recovered response will be smaller for an *i* code than a *c* code containing an equal amount of information. It may be assumed that the likelihood of an *i* code being in the examination subset will be a function of its similarity to the test stimulus, since storage is carried out primarily on the basis of stimulus information.

The probability of accepting an *i* code as being correct will similarly depend upon the generalization from the test stimulus to the stimulus information encoded in the *i* code. Given that the *i* code is examined and accepted, however, the probability that a response will be recovered and emitted should depend directly upon the response information encoded, just as for a *c* code. The above statements allow intrusion probabilities to be predicted for various conditions. In the situation of Figure 2, for example, an increase in intrusions will be predicted over the course of the session (since the number of *i* codes in the examination subset on tests during the course of the session will increase). This increase has been found in such a task, and a model based on the present theory predicts the increase accurately (Shiffrin, 1968).

Another phenomenon predicted by the theory is that of second-guessing, where second-guessing refers to the giving of a second response after *S* has been told that his first response is incorrect. A variety of assumptions can be made about this process, the simplest of which postulates that *S* continues his search of the examination subset from the point where the intrusion occurred. This assumption predicts that the level of second-guessing will be above chance, an effect found by Binford and Gettys (1965). If the search is temporally ordered to any degree, then strong predictions can be made concerning the second-guessing rate depending upon whether the response given in error was paired in the sequence with a stimulus occurring before or after the tested stimulus (assuming that the task utilizes a set of unique responses). In fact, examination of this effect is one method of determining the temporal characteristics of the search.

Latency of Responses

Another variable which may be predicted from the theory in a straightforward way is the latency of responses. The basic assumption requires latency to be a monotonic increasing function of the number of images examined before a response is emitted. Among the implications of this assumption are the following. Latencies of correct responses should increase with increases in the number of intervening items. This prediction holds whenever there is some temporal component to the search, or whenever the number of items preceding the tested item is large. If the reasonable assumption is made that codes containing more information are examined earlier in the search, then a decrease in correct response latency is expected as the number of reinforcements increase, since the item will gain stored information over reinforcements and therefore tend to be examined earlier in the search. This effect has been found by Rumelhart (1967) and Shiffrin (1968) in a continuous paired-associate task. In general, any manipulation designed to vary the number of codes examined, whether by instructions, by organization of the presented material, or by other means should affect the response latencies in a specifiable way.

Recognition and Recall

In terms of the present system the search proceeds in a similar manner whether recognition or recall is the mode of test; the difference lies in the size of the examination subset in the two cases. Once information is recovered from LTS, however, the decision process involved in response generation may be somewhat different for recognition and recall. In a paired-associate design, the search will begin with an attempted recognition of the stimulus, with the decision whether to continue the search dependent upon a positive stimulus recognition (Martin, 1967). Hypotheses which ascribe different retrieval mechanisms for recognition and recall are not necessary. In both recognition and recall the presented stimulus will be sorted into an LTS area, and a search initiated there. In the case of recognition, this search can be quite limited, perhaps consisting of an examination of a single image. In the case of recall, the stimulus may be recognized with little search needed, but the necessity for recovering the response may entail a larger search, although "larger" might imply only examination of two to five additional items (Shiffrin, 1968).

CONCLUSIONS

The theory outlined here is descriptive; we have attempted to present a theory of memory in fairly general terms and to demonstrate for certain commonly studied variables how the theory can be applied. It is beyond the scope of this paper to present specific quantitative models that follow from the general theory and apply them to data, but such models have been set forth elsewhere and applied successfully: in continuous paired-associate learning experiments where the variables examined included the number of intervening items, rankings of responses, second-guessing, proactive interference effects, intrusions, and latencies (Shiffrin, 1968); in free-verbal recall where the variables examined included list length and presentation time (Atkinson & Shiffrin, 1968a); and in paired-associate memory tasks where the variables include list length, confidence ratings, and response times (Atkinson & Shiffrin, 1965; Phillips, Shiffrin, & Atkinson, 1967). Despite these successes, we wish to emphasize that the theory is still in an early formative stage, and awaits application to a wider range of problems. For example, it is not yet known whether the theory can be extended in an elegant way to account quantitatively for the interference phenomena observed in a typical list-structured task. Whatever the fate of such applications, the present theory serves the purpose of providing a general framework within which many of the specific quantitative models known to the authors may be placed, including all of our own work. In addition, we hope that this report will lead to a more detailed consideration of memory input and output mechanisms, especially the memory-search process.

REFERENCES

ATKINSON, R. C., BRELSFORD, J. W., JR., & SHIFFRIN, R. M. Multiprocess models for memory with applications to a continuous presentation task. *Journal of Mathematical Psychology,* 1967, **4**, 277–300.

ATKINSON, R. C., & SHIFFRIN, R. M. Mathematical models for memory and learning. Technical Report 79, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1965. (Republished: D. P. Kimble (Ed.), *Proceedings of the third conference on learning, remembering and forgetting.* New York: New York Academy of Science, in press.)

ATKINSON, R. C., & SHIFFRIN, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory,* Vol. 2. New York: Academic Press, 1968. (a)

ATKINSON, R. C., & SHIFFRIN, R. M. Some speculations on storage and retrieval processes in long-term memory. Technical Report 127, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1968. (b)

BERNBACH, H. A. Decision processes in memory. *Psychological Review,* 1967, **74**, 462–480.

BINFORD, J. R., & GETTYS, C. Nonstationarity in paired-associate learning as indicated by a second guess procedure. *Journal of Mathematical Psychology,* 1965, **2**, 190–195.

BJORK, A. Learning and short-term retention of paired-associates in relation to specific sequences of interpresentation intervals. Technical Report 106, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1966.

BOUSFIELD, W. A., & COHEN, B. H. Clustering in recall as a function of the number of word-categories in stimulus-word lists. *Journal of General Psychology,* 1956, **54**, 95–106.

BOWER, G. H. A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation.* Vol. 1. New York: Academic Press, 1967.

BRELSFORD, J. W., JR., & ATKINSON, R. C. Short-term memory as a function of rehearsal procedures. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7**, 730–736.

BRELSFORD, J. W., JR., SHIFFRIN, R. M., & ATKINSON, R. C. Multiple reinforcement effects in short-term memory. *The British Journal of Mathematical and Statistical Psychology,* 1968, **21**, 1–19.

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

COHEN, B. H. Recall of categorized word lists. *Journal of Experimental Psychology,* 1963, **66**, 227–234.

DALLET, K. M., & D'ANDREA, L. Mediation instructions versus unlearning instructions in the A-B, A-C paradigm. *Journal of Experimental Psychology,* 1965, **69**, 460–466.

DAPOLITO, F. J. Proactive effects with independent retrieval of competing responses. Unpublished doctoral dissertation, Indiana University, 1966.

DEESE, J. *The structure of associations in language and thought.* Baltimore: Johns Hopkins University Press, 1966.

ESTES, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 2. New York: McGraw-Hill, 1959.

FEIGENBAUM, E. A. Information processing and memory. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and proba-*

bility, 1966. Vol. IV. Berkeley: University of California Press, 1966.

GREENO, J. G. Some attempts to locate associative information. Technical Report, Department of Psychology, Indiana University, 1967.

HARLEY, W. F., JR. The effect of monetary incentive in paired-associate learning using an absolute method. *Psychonomic Science*, 1965, **3**, 141–142.

HELLYER, S. Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. *Journal of Experimental Psychology*, 1962, **64**, 650.

HINTZMAN, D. L. Explorations with a discrimination net model for paired-associate learning. *Journal of Mathematical Psychology*, 1968, **5**, 123–162.

KEPPEL, G. Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

KINTSCH, W. Memory and decision aspects of recognition learning. *Psychological Review*, 1967, **74**, 496–504.

McGOVERN, J. B. Extinction of associations in four transfer paradigms. *Psychological Monographs*, 1964, **78** (16, Whole No. 593).

MANDLER, G. Association and organization: Facts, fancies, and theories. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968.

MARTIN, E. Stimulus recognition in aural paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 272–276.

MELTON, A. W. Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 1–21.

MELTON, A. W., & IRWIN, J. M. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology*, 1940, **53**, 173–203.

MILLWARD, R. An all-or-none model for noncorrection routines with elimination of incorrect responses. *Journal of Mathematical Psychology*, 1964, **1**, 392–404.

MURDOCK, B. B., JR. The serial position effect of free recall. *Journal of Experimental Psychology*, 1962, **64**, 482–488.

NORMAN, D. A. Toward a theory of memory and attention. *Psychological Review*, 1968, **75**, 522–536.

OSGOOD, C. E. Psycholinguistics. In S. Koch (Ed.), *Psychology: A study of a science*. New York: McGraw-Hill, 1963.

PHILLIPS, J. L., SHIFFRIN, R. M., & ATKINSON, R. C. Effects of list length on short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 303–311.

POLLIO, H. R. *The structural basis of word association behavior*. The Hague: Mouton, 1966.

POSTMAN, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961.

RUMELHART, D. E. The effects of interpresentation intervals in a continuous paired-associate task. Technical Report 116, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1967.

RUNQUIST, W. N., & FARLEY, F. H. The use of mediators in the learning of verbal paired-associates. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 280–285.

SCHNORR, J., & ATKINSON, R. C. Repetition versus imagery instructions in the short- and long-term retention of paired associates. *Psychonomic Science*, 1969, in press.

SHIFFRIN, R. M. Search and retrieval processes in long-term memory. Technical Report 137, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1968.

SPERLING, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, **74** (11, Whole No. 498).

THOMPSON, W. J. Recall of paired-associate items as a function of interpolated pairs of different types. *Psychonomic Science*, 1967, **9**, 629–630.

TULVING, E. Subjective organization in free recall of 'unrelated' words. *Psychological Review*, 1962, **69**, 344–354.

TULVING, E. The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 175–184.

UNDERWOOD, B. J. Interference and forgetting. *Psychological Review*, 1957, **64**, 49–60.

UNDERWOOD, B. J. Stimulus selection in verbal learning. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill, 1963.

WICKELGREN, W. A., & NORMAN, D. A. Strength models and serial position in short-term recognition memory. *Journal of Methematical Psychology*, 1966, **3**, 316–347.

(Received February 2, 1968)