Depth of processing in recognition and recall: Some aspects of a general memory system.

Robert S. Lockhart, Fergus I.M. Craik and Larry L. Jacoby

University of Toronto

Chapter to appear in Recognition and Recall, John Brown (Ed.) London: Wiley, 1975.

This research was supported by grants from The National Research Council of Canada, No. A0355 to R. S. Lockhart and No. A8261 to F. I. M. Craik. We would like to thank P. Kolers, B. B. Murdock, Jr., and E. Tulving for their critical and helpful comments.

INTRODUCTION

Most attempts to characterize the essential differences between recognition and recall have made use, in one form or another, of the distinction between "storage" and "retrieval." A straightforward example is the two-process theory of Kintsch (1970) in which it is claimed that whereas recall involves both search and decision, recognition entails only the latter. Compared with earlier accounts such as those which distinguished recall from recognition in terms of a different threshold, the two-process theory appears to have several advantages. Its additional degree of theoretical freedom allows it to account readily for results (see McCormack, 1972) which show that certain variables influence recall but not recognition. Further, if a two-process theory is accepted, recognition can be taken as a relatively pure measure of storage, and recognition/recall comparisons used to establish whether a given variable exerts its effect via the retrieval or storage phase. Murdock (1968), for example, concluded that since differences between auditory and visual presentation exist in both recognition and recall, such effects represent storage differences.

Apart from its simplicity, a major factor contributing to the persistence of the two-process theory is undoubtedly the elasticity of the terms, especially terms such as "retrieval" and "search." Efforts to falsify the theory have done so by attempting to demonstrate that recognition does, after all, involve "retrieval." But the meaning of this term (and thus the status of the evidence) is left quite vague. It is, of course, a simple matter to define retrieval in operational terms. Such a definition would identify retrieval effects with performance differences which occur as a function of conditions at time of test, all conditions prior to the test phase having been kept constant. This definition is satisfied by paradigms such as those which vary the context of recognition test items, or the availability of cues for recall. As straightforward as this operational definition is, it should be noted that it encompasses a concept of retrieval as broad as that of Melton's (1963) general notion of trace utilization. Moreover, it fails to incorporate the important point (stressed by Tulving and Thomson (1973) that item retrieval, properly understood, involves an <u>interaction</u> between encoding processes and test conditions. For example, whether or not recall is facilitated by the presence of certain "retrieval cues" may depend critically on the form of the item's initial encoding and a failure to recall might just as well be viewed as a consequence of inappropriate initial coding as due to an inadequate retrieval cue.

If retrieval is defined as broadly as the operational definition above, disproving the two-process theory becomes a trivial matter; it need only be shown that recognition accuracy depends on the context in which the test item is presented. It seems likely then, that the advocates of the two-process theory are using the term retrieval in a different, more restricted sense, but the exact nature of this restriction is far from clear. Attempts to pin down the term more precisely have most frequently introduced the concept of search and the retrieval problem is that of locating the to-be-recalled item, somewhat like locating a book in the library. There are several similar analogies that have been coined to make the same point. In a somewhat different context, Miller (1963) drew an analogy with a junk box; inability to recover an object

might be a consequence either of its not being in the box or being in the box but effectively lost amongst the countless other objects, just as an incorrectly shelved library book can no longer be easily located. In its turn this type of analogy draws heavily on computer models of information retrieval for which notions of search and location can be given unambiguous meaning. The usefulness of this analysis was demonstrated by Tulving and Pearlstone (1966) who used the terms "availability" to refer to the existence of the intact trace in the memory store and "accessibility" to refer to the problem of locating it. A retrieval cue is then viewed as anything which serves to locate the available but hitherto inaccessible trace.

This conceptualization of retrieval has been most extensively studied within the context of categorical organization, the assumption being that the facilitating effects of such organization operate via the retrieval component. It is not surprising then that many attempts to test the two-process theory's account of recognition and recall differences, have centred on the question of whether or not organization influences recognition. The results have varied from study to study, but it seems safe to conclude that there are certain conditions under which recognition increases with degree of organization (Mandler, 1972). However, it is quite gratuitous to infer that these results say anything about retrieval processes in recognition. Notice that the typical experiment reported by Mandler (1972) does not conform to even the broad operational definition given above. It could be argued with equal plausibility that the task of sorting items into categories involves processing of a kind which establishes a more durable or distinctive trace and that the more categories employed, the deeper and more discriminating that processing will be. This issue is discussed in greater detail in a later section of this chapter.

The persistent difficulty with attempting to answer a question such as "does recognition entail retrieval?" is that as soon as one steps beyond the broad operational definition, it becomes apparent that the meaning of the question varies with one's theory of the recognition process and that from some points of view it may not be a particularly sensible one. Under these circumstances it is preferable to set such questions aside until a more detailed analysis of the recognition process has been made. In the next section such an analysis is attempted and using this analysis we then examine a number of basic questions of current interest. Naturally, many elements of our system have been suggested by previous workers; we make no great claims for originality, but we believe that the specific proposals advanced here fit the known facts reasonably well. In essence we suggest that recognition of an event depends on (a) the depth of initial encoding, and (b) the similarity of presentation and test encodings. We retain Tulving's (1972) distinction between semantic and episodic memory, although we believe that these two aspects of the perceptual/memory system are more closely interrelated than Tulving supposed. In our scheme, semantic memory acts as part of a pattern recognition system whose function is to interpret incoming stimuli by means of . complex analysing and encoding operations. The product of these operations is the memory trace which forms the latest addition to episodic memory -- that part of the system comprising the temporally-ordered collection of all encoded episodes and events. The deeper and more elaborately a stimulus is analysed by the perceptual system, the

richer and more detailed will be the episodic memory trace. At the time of the recognition test, the probe stimulus is again encoded by the pattern recognition system and the resulting encoding is used to specify the initial trace. For recent items, retrieval information is used to select the required trace; while for remote events the encoded probe information is used as the basis for reconstruction of the initial stimulus -- this reconstructive activity is constrained and guided by information contained in the memory trace. The purpose of the next few sections is to examine in greater detail what is meant by "depth" of processing and to suggest ways in which the recognition process may function.

Encoding of the presentation stimulus

Levels of processing

Craik and Lockhart (1972) described the processing of a stimulus in terms of a continuum of analysing operations. First the physical and structural features of the stimulus are analysed, then the stimulus is subjected to progressively more elaborate semantic analyses. The phrase "greater depth" referred to these later semantic-associative operations. It was further postulated that the memory trace was a by-product of the analysing operations and that the durability of the trace was a function of depth -- deeper initial processing yields a longer-lasting trace.

We would like to modify some of these ideas. For example, if "shallow levels of processing" comprise simple physical analyses and "deep levels" consist of complex semantic analyses, it is difficult to see how these very different analytic operations could possibly lie on a continuum. There is no sense in which the physical analysis of visually-presented words (type-face, letter form, size, colour, etc.) shades off into the meaning which the words convey. Rather, the physical, phonemic and semantic characteristics of words exist in different dimensions or domains to use Sutherland's (1972) expression. We suggest that "greater depth" may refer to two somewhat distinct changes in processing. First, the domains themselves may be thought of as a hierarchical organization proceeding from shallow, structural domains to deep, semantic domains. Second, at one depth in this sense, the stimulus may be further analysed or elaborated by carrying out additional operations within one qualitatively coherent domain. Although the term "depth" will be retained to describe all cases where many analytic operations have been carried out, the two senses in which the word is used should be borne in mind. Some further comments on the two kinds of depth are offered below.

In Sutherland's (1972) discussions of object recognition he suggests that a structural description of the input pattern is formed within each analytic domain. At each level of analysis, the structural descriptions formed within one domain serve as the input to the next domain via a set of mapping rules. A slightly different way of phrasing these ideas is to say that each level of analysis provides evidence which is used to confirm (or reject) the structural description of the hypothesized patterns at the next level. This second way of describing the process may be preferable in that it stresses the notion that structural descriptions at any level are as much a product of expectancies and past learning as they are products of the current stimulus input. Further, descriptions of very probable events will require only minimal confirming evidence from preceding levels, since their structural descriptions have been largely preformed in anticipation. These ideas are essentially identical to Norman's (1968)

discussion of the roles of sensory input and "pertinence" in perception, and to Treisman's (1964) notion of levels of analysis, where each level is seen as a signal-detection test whose function is to allow or prevent the passage of information from one level to the next.

This analysis has several implications for practice and repetition effects. Since practice has the effect of making stimuli, or stimulus-response sequences, more probable, highly practised encoding operations will run off with a minimum of processing in each domain. That is, fewer encoding operations take place at each qualitative level of analysis. With extended practice, it may even be possible to bypass a complete domain -- thus while young children and poor readers may be forced to sound out letter sequences to understand written text, the practised reader may bypass the phonemic stage entirely, for simple material at least (Hardyck & Petrinovitch, 1970).

Thus while we believe that processing typically proceeds through a fixed series of qualitatively distinct stages or domains, this does not mean that all possible analyses are laboriously carried out in each domain. Only those analyses required to provide critical evidence for deeper levels of processing are carried out. Our views may be contrasted with the notion that practice has the effect of speeding up the same encoding sequence. We suggest that, with practice, fewer operations are carried out.

In general, processing proceeds until the domain relevant to the present task is reached and, quite often, it is only at the "target" domain that sufficient processing occurs for conscious awareness of the processing operations to take place. By this latter proposition we mean that encoding operations for familiar stimuli can be effected with minimum evidence from shallower domains and that the initial analysing operations can be run off automatically, without the involvement of consciousness. However, we can become conscious of some early processing operations by directing attention to the relevant domain; in this case, further analyses are carried out within that domain and conscious awareness ensues. The difference in these two processing strategies is illustrated by skimming an easy text for gist as opposed to reading the text for typing errors. There are also exceptions to the general statement that processing proceeds only as far as the target domain -- if a stimulus is highly compatible with deeper levels of analysis it may be impossible to prevent processing occurring at a deeper level even when it is not in the subject's interests to do so; the Stroop effect may be described in these terms.

We do not believe that it is very useful to talk about verbal encoding as "automatic" except, perhaps, in certain limited contexts (Posner & Warren, 1972; Wickens, 1970). While it is true that a well practised encoding may be accomplished by little or no conscious effort, this accomplishment seems to be a matter of degree rather than an all or none event. Words in a half-learned foreign language often require conscious processing before they are understood; with further practice they require less analysis until subjectively the processing feels "automatic." But if the context is misleading or the subject is set to perceive a different word, or if he is fatigued or distracted, he may again be forced to invoke conscious processing before the word is understood. The apparent automaticity of encoding depends on the amount of analysis which must be performed on the word before its meaning is extracted -- this in turn will depend on such factors as the material, practice, context and set.

The preceding discussion of the amount of processing carried out in each domain has important implications for our view of memory. We believe

that the episodic memory trace results from those encoding operations carried out in the pattern recognition and cognitive systems. Thus a reasonably familiar pattern or stimulus-response sequence will be analysed and encoded by a moderate number of analytic operations and will result in a moderately rich memory trace. When a stimulus becomes highly practised or expected, it is analysed more readily be fewer operations and impoverished memory trace will result. Conversely, if the stimulus is novel, or difficult to process, or important, more analyses are carried out and a richer memory trace results.

The Order of Analysis

Savin and Bever (1970) showed that target syllables could be detected faster than target phonemes in acoustically presented syllables. They concluded, therefore, that syllables were not in any sense constructed from phonemes — syllable analysis seemed to be prior to phoneme analysis. In the same vein, Mistler-Lachman (1974) has shown that deep structural ambiguity is apparently resolved before surface level lexical ambiguity. These demonstrations cast doubt on the validity of a processing sequence which necessarily proceeds from simple to complex. Two comments are offered on these results. First, McNeill and Lindig (1973) report further results showing that words can be detected faster than syllables and even that short sentences are detected faster than words. It is certainly difficult to believe that sentences are completely processed before any phonemes or syllables are perceived. McNeill and Lindig show that minimum reaction times occur whenever the linguistic level of the target and the search list is the same; further, given a match between target and list items, latencies tend to be faster for shallower linguistic levels. Thus they argue that the relatively long latencies shown by Savin and Bever's subjects when they detected phonemes, were artefactually caused by the mismatch between the target and list items.

A second reason underlying the faster detection of apparently deeper features may be that such deeper levels of linguistic analysis (words and phrases for example) are the levels at which conscious decision-making typically operates. We speculatively suggest that the amount of processing necessary to yield conscious perception of a spoken word may be less than that necessary to yield conscious perception of its constituent phonemes, even although evidence from the phonemic domain is a necessary input for word perception. That is, the amount of processing at the phonemic level which is necessary to provide confirming evidence for word perception, is less than the processing necessary to construct a conscious phoneme percept. This shift in the level at which consciousness operates, from shallower to deeper levels, seems typical of all practised skills — a beginning piano player can run off a well practised phrase more quickly than he can play a single note contained in that phrase; we perceive the figures and objects in a pointilliste painting faster than we perceive the constituent coloured dots — even though the deeper more meaningful levels must depend on some analysis at shallower levels. Reaction-time studies are not a good source of evidence for the order in which encoding operations are performed.

Further elaborations within a domain

In the same way that the processing of a well-practised stimulus proceeds rather automatically from domain to domain, further processing of a common item or event is carried out within one domain. Those

analyses are performed which enable the perceiver to distinguish the event from other possible events; again if the event is common or expected, rather little confirming evidence may be required and few analyses need be carried out. If the event is uncommon, unexpected or ambiguous, further processing within the domain is necessary to specify the event unambiguously. Again it is suggested that well practised analyses are carried out with little conscious involvement; to the extent that the task requires further processing of the stimulus input, conscious effort is involved. As a further suggestion, automatic, practised encodings make use of the existing analysing routines in semantic memory; when further conscious analyses are required, the structure of semantic memory itself is altered by these new analyses. Thus, if two things appear together for the first time (a particular man and woman, let us say) it may require some conscious effort to construct a new encoding (what do they have in common?). After several such appearances, however, our encoding system contains a new "couple" concept and the subsequent analysis of each constituent is altered accordingly.

In summary, when a stimulus is presented or an event occurs, it is analysed by the pattern-recognition processes and other "cognitive structures" which comprise semantic memory. The essential purpose of these analytic operations is to interpret and understand environmental events -- the memory trace is seen as the by-product of these cognitive operations. The bundle of analysed features which comprises the memory trace, forms the latest addition to episodic memory. The richness of the memory trace depends on the number and the nature of features analysed. In turn, the operations carried out on the stimulus depend on the interpretive task given to the subject, the context and the material. If the stimulus is expected and commonplace, few features need be noted to resolve cognitive ambiguity; if the stimulus is rare, unexpected or must be discriminated from similar stimuli, many features are analysed and a rich memory trace results. Semantic memory is seen as the system concerned with storing our knowledge of the world; it thus provides the means to analyse and interpret events as they occur. Apart from shallow sensory "primitives" the analysing processes of semantic memory have been formed from recurring commonalities of past events -- rules and relations have been abstracted from previous episodes. To the extent that further episodes conform to past rules and regularities the structure of semantic memory remains unchanged -- our interpretation of the world is confirmed. When events occur in new groupings, however, or demand different responses for their resolution, the structure of semantic memory must be amended. In this sense "memory" can be identified as the bundle of encoded features in episodic memory while "learning" corresponds to a change in the structure of semantic memory.

Episodic Memory

Before moving on to a consideration of trace utilisation, we would like to make a few comments concerning the episodic nature of the encoded trace. While we agree with Tulving (1972) that the distinction between episodic and semantic memory is a useful one, we question a number of the statements he makes about the nature of episodic memory. First, we would place greater emphasis on the inter-dependence of the two systems; on the one hand the structure of semantic memory has been formed from a series of episodes and, on the other, an event in episodic memory can be accessed by repeating the encoding

operations in semantic memory. This suggestion is amplified below. Second, Tulving suggests that episodic memory is structured in temporal and spatial terms. We would like to offer the more radical suggestion that episodic memory has no inherent structure. Spatial and other attributes of events are analysed and encoded at input, while the temporal properties of retrieved events are inferred from contextual features encoded as part of the episodic trace and from the ordinal position of the event relative to others. Thus episodic memory is not seen as a system with a built-in time marker, but a rather structureless system which maintains the order in which episodes occur but does little else. The fact that we can perceive temporal rhythms may necessitate some modification of this statement, at least as far as recent events are concerned. In the same way that literal properties of visual and auditory events are apparently preserved for short times, so a literal record of temporal sequences may be available for a limit[ed] time after their occurrence.

We see semantic memory as a system whose function is essentially one of interpreting incoming stimuli as a basis for deciding on some relevant action. Thus episodic memory contains copies of these interpretations or "prescriptions for action." While it is undoubtedly not worthwhile debating whether the memory trace is the end result of perceptual input processes or is the "efferent copy," a copy of output instructions, some emphasis on the latter view is probably useful since it stresses the active, dynamic role of memory as opposed to the view of the memory trace as a dead inactive residue.

After an episodic trace has been laid down, it can still be accessed from its "home base" in semantic memory. That is, when the event recurs, and if it is processed in a similar way, this repetition of the encoding operation helps to evoke the memory of the original event. In our scheme the repeated encoding operations serve as a basis for "reconstruction" of the encoding induced by the original event. Formation of the test encoding is constrained on the one hand by the test stimulus and on the other by information from the episodic trace. It is also suggested that traces may be accessed in a quite different way, by searching back through recent episodic memory, using the retrieval information to select the relevant items from other recent items. Thus episodic traces can be accessed either directly from semantic memory or by a scanning procedure through recent episodic memory.

Utilisation of the Stored Trace

We suggest that the memory system has two modes of retrieval available. The first of these may be described as a process of reconstruction and the second as a search or scanning operation in which recent episodic traces are examined for the presence of some salient feature of the retrieval probe. These notions are explored further below.

Retrieval by Reconstruction -- The "Resonance" Metaphor

What happens when a stimulus pattern is recognised? Phenomenologically a feeling of familiarity is experienced, but we can also locate the previous occurrence of the event in some specific temporal and spatial context. That is, conscious experience contains features induced by the stimulus pattern itself plus contextual features which were present during the first occurrence of the event. Recognition can thus be described as a process in which some approximation to the initial encoding of the event is reconstructed in the perceptual/cognitive system. The reconstruction is guided and constrained by the recognition stimulus on the one hand and information from the episodic trace on the other.

If good recognition depends on achieving a highly similar encoding to that induced on the item's first presentation, then it is easy to see why memory performance improves as the stimulus-plus-context presented at the time of test, approximates more closely to the stimulus-plus-context presented when the item was learned initially. By this view, recognition is generally superior to recall since the recognition stimulus can evoke a similar encoding more easily. It is more difficult to see how the specific context surrounding the item's first presentation is re-achieved. Initial context cannot be induced by the test context and must be derived from the episodic trace. How does the present stimulus contact traces of its previous occurrences? We have no complete answer to this central question, but offer an analogy which may help in conceptualising the process.

It is suggested that when a particular pattern of encoding operations is induced by the test stimulus, all episodic traces of the pattern are contacted and help to constrain further reconstructive encoding of the stimulus. If there are many traces of the pattern, a "familiar" encoding will be achieved easily (since it is guided by many previous traces and thus easily encoded), but the stimulus will not elicit recognition of a specific previous instance (since many different contexts are competing for conscious awareness). If the pattern was unique or distinctive, however, only one or a few traces of such a pattern exist -- now if the retrieval information (the test stimulus) is specific enough, the previous trace will be contacted and the present encoding will contain contextual features from the initial encoding.

One metaphor for the basic mechanism of such a model of the recognition process is the phenomenon of resonance. The test encoding is viewed as analogous to the application of a specific frequency which tends to elicit, as feedback, sympathetic vibrations in all the episodic "tuning-forks" set to that frequency. However, the parallel to physical resonance breaks down at this point since we want to say that if many episodes share the features of the test encoding, each episode will be stimulated only slightly. Thus the mixed metaphor is a "limited energy resonance model" -- if there are few tuning-forks (episodes) at that frequency, each will be stimulated to a greater extent. Of course, the resonance model is substantially more passive than the notion we wish to convey. Rather than the stimulus simply "evoking" or "activating" the trace, we prefer to think of the trace shaping the encoding of the test stimulus as it develops, the resonance mechanism providing the necessary feedback. It follows from these notions that deeper, more elaborate encodings will be recognized better since such encodings are more unique and distinctive. By this view, too, there is nothing magical about the fact that rare words are recognized better than common words -- rare words are more unique, they have occurred less often and fewer competing traces are contacted by the retrieval information. This in turn means that it is easier for the retrieval system to reconstruct the specific episode of "occurrence in the presentation list". If a common word is made more unique, for example if it is qualified by an unusual adjective, then it should also be well recognized -- thus "dog" is poorly recognized while "crimson dog" may be well recognized. The interaction of word frequency with recognition and recall is discussed further below.

To this principle of distinctiveness we add two further components of good recognition. The first is the "richness" of the encoding -- the number of features analysed at input and thus laid down in the episodic trace -- and the second is a "set to recognize" factor. By the first point we mean that good recognition is partly a function of the number of features which

the presentation and test encodings have in common; other things being equal, "richer" traces will be better recognized. A trace is enriched if features of the context are encoded with it, or if it is modified or qualified in some specific way -- if the same features are encoded at test, then recognition has a higher chance of success. By the second point we mean that some further factor may facilitate "trace contact" and thus the reconstruction of an encoding similar to the initial encoding. This factor may be described as a "set to recognize." Although undoubtedly recognition occurs when the subject does not expect a recurrence of the event, recognition (and perhaps false alarms) may be enhanced if the subject encodes the test stimulus with a view to making a recognition decision. "Set" in this situation may take the form of active construction of plausible contexts in an attempt to contact previous traces.

Thus, subjective feelings of recognition may accompany two types of situation -- first when a common event is presented it is encoded easily since many previous traces facilitate the present encoding. In this case recognition is a function of "repetition of operations" as suggested by Kolers (1973). In the second case -- repetition of an uncommon or distinctive event -- the feeling of recognition is based on the reconstruction of the initial presentation context. One way of stating these suggestions is to say that common events will typically give rise to strong feelings of familiarity in semantic memory -- the subject is extremely familiar with the meaning and connotations of the word "dog," but cannot recall where or when he heard the word uttered on previous occasions. Unique or distinctive events have a much greater chance of being recognized as specific episodes even although they are not so familiar in the semantic sense. In point of fact, a common event may not even give rise to a feeling of recognition in semantic memory. If the event is highly probable and expected in a particular context, it can apparently be assimilated by the system with a minimal amount of analysis — thus highly familiar objects in your living room do not typically give rise to strong feelings of recognition whenever you glance round the room, although if conscious attention is brought to bear on the object for some reason, it is felt to be familiar. Paradoxically, it is when the object or event is missing or out of context that consciousness is involved. This is the phenomenon of dishabituation, described by Sokolov (1963) in terms of a mismatch between the encoded stimulus and some internal representation. In our terms, the context induces the system to construct expectancies which are usually confirmed by a minimal analysis of the incoming stimulus. When the new analysis clearly does not confirm the expected patterns, conscious attention must be invoked to carry out the much more elaborate processing now necessary to identify the unexpected event or to ascertain why the expected event is missing. In the expected case, memory for the event will be poor since few features were analysed while, by the converse argument, memory for unexpected events will be good.

In this system recall is also viewed as a reconstructive process. Again the information presented at retrieval provides the basis for the reconstruction, but since this information is typically rather meagre, more information must now be contributed by the system — either from well-practised encoding routines or from the episodic trace itself. Thus recall, like recognition, is conceived of as "guided reconstruction." We believe that recall and recognition do not differ in any crucial way — they are

different only in the sense that in recognition re-presentation of the stimulus provides better information from which the initial encoding can be reconstructed. The initial encoding contains features induced by the nominal stimulus and features induced by the context — in recognition, the stimulus is re-presented and the system has to reconstruct the context; in recall some aspects of the context are re-presented or referred to and the system has to reconstruct the stimulus. Thus again, recall involves contacting the episodic trace and it is postulated that the reconstructive processes are constrained and guided by episodic information — if reconstruction is proceeding on the right lines, further reconstruction of the same type is encouraged by feedback from the episodic trace.

We would like to distinguish our suggested recall process from two other suggestions made by previous workers — the "search" analogy and "generate-recognise" models. First, the idea of recall as a search process implies a blind procedure in which stored events are examined either systematically or at random until the desired information is located. In contrast, we believe that the actual process of constructing the encoding at retrieval is guided by feedback of an increasingly specific kind from the trace. Second, our suggestion differs from "generate-recognise" models in that the present system deals with partial information, not necessarily with totally reconstructed events. Thus, if the reconstruction is accurate in general terms, "positive feedback" will result and further reconstruction along the same lines is encouraged. This "guided reconstruction" is seen as a servo-mechanism in which feedback from the target controls the reconstructive processes in semantic memory. This system is quite different from a generate-recognise model in which complete items are constructed and then matched with episodic events.

Both recall and recognition are superior when "deeper," "richer," "more semantic" traces are formed at input. There is much evidence to support this claim at an empirical level (e.g., Hyde & Jenkins, 1969; Craik, 1973). By the present view the beneficial effect of depth of encoding is that deeper, richer encodings are also more distinctive and unique; thus their resulting episodic traces are more easily contacted (the resonance metaphor) and the richer information may then provide more adequate feedback to guide further reconstruction.

This section has dealt with the retrieval of stored events by means of a repetition of encoding operations. It should be pointed out that this notion is radically different from a scheme which postulates one memory store with "time-tags" to denote occurrence information. By the present view, when a probe item is encoded, it does not "point" to an area of <u>time</u> in episodic memory, but rather to similarly encoded <u>events</u>. In this system, time is inferred from various cues (e.g., the ordinal storage of events; linking the event to well-established times and dates, etc.) in much the same way that in visual perception, depth is inferred from input cues. The suggestion that the time or dating of an event is not coded directly but is inferred from other stored events, is illustrated by the findings of Warrington and Sanders (1972); when subjects were asked to date various events which had been in the news, they were only able to estimate the date of an event in relation to other well established dates in their lives.

<u>The "Selection" function of cues -- looking back in recent memory</u>

So far we have talked about using retrieval information as a means of contacting the episodic trace, and as a basis for reconstructing an encoding

of the original event. However, we believe that retrieval cue information (in both recognition and recall) can be used in a different fashion when the target trace has been laid down recently. In this second retrieval method, some salient aspect of the probe information is held and recent episodic traces are scanned for the presence of this salient feature. The scanning process is analogous to Neisser's (1964) visual scanning procedure in which the subject rapidly scans a visual array of letters for some target letter. In the recent memory situation, the subject knows from the probe item that he is searching for an episode with some specific feature -- it rhymes with TRAIN or is a synonym of GOOD. When the probe information is not particularly distinctive, it may be more efficient to use the information as a selector device, rather than as a means of gaining access to all previous episodes which share the probe features.

When retrieval information is used in this fashion, different factors affect the success of the recognition operation. Now the most important factor is not the similarity of probe and target attributes, but the salience of the information used to pick out the target event from other recent events. Just as a target K can be very easily picked out from a context of round letters (e.g., O,P,Q,C) in Neisser's paradigm, so a distinctive event can be located in recent memory -- the demonstration by Kroll <u>et al</u>. (1970) that a visual letter is better retained than an auditory letter over a period of auditory shadowing might be explained in this way. Thus the critical feature of the scanning operation is the ease or difficulty of <u>discriminating</u> between the target and other recent traces. In this sense, "surface structure" such as visual or phonemic features may well be as good or even better than deeper semantic features when used as selector information. This state of affairs may be contrasted with information used in the reconstruction mode -- in that case semantic information is much more effective in achieving correct recognition since more features are involved and each collection of features is highly distinctive. These two ways in which retrieval information may operate are to be thought of as optional alternatives in the same system; they are modes of retrieval, much as suggested by Tulving (1968) but with the difference that in the present system, the same type of information may be used in either mode -- we have no wish to equate the selector mode with phonemic information and reconstruction with semantic functioning.

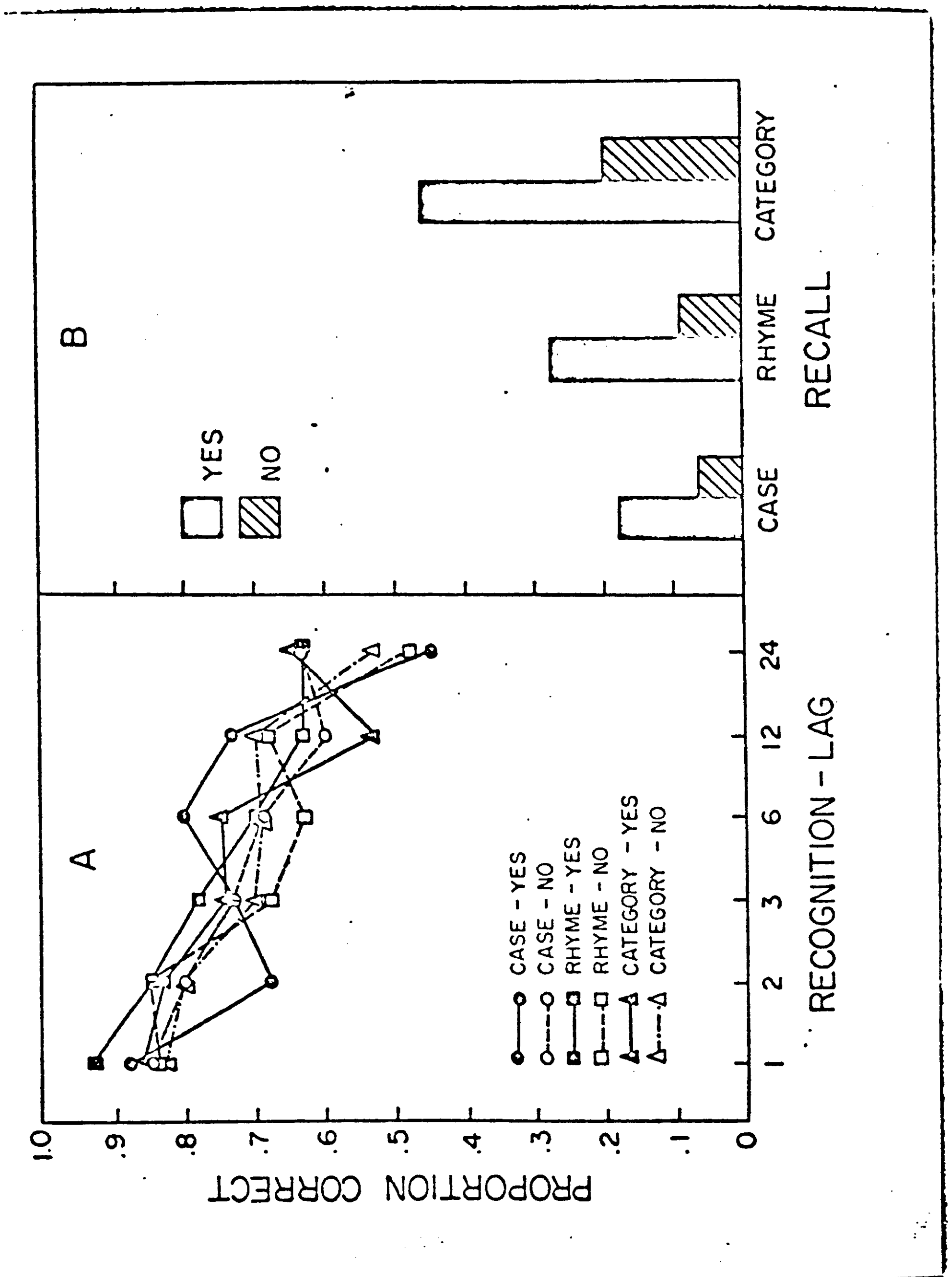
Why postulate two retrieval modes? In our view, a number of experimental results make the notion a necessary one. Several recent studies have shown clearly that if a word is encoded in terms of its semantic features, it is much better recalled and recognized that if it is encoded in terms of the phonemic features; similarly phonemic encoding is better than structural encoding (Craik, 1973; Hyde & Jenkins, 1969; Schulman, 1971). On the basis of these and similar studies, Craik and Lockhart (1972) suggested that deeper, more semantic coding yielded a more durable trace. However, there are some other studies whose results are curiously at odds with the notion that semantic encoding gives rise to a longer-lasting trace. Bregman (1968), Shulman (1970), and Jacoby (1974) all presented lists of words which were then tested by several types of probe -- semantic, phonemic, structural, etc. -- all these authors found the "decay rate" of recently presented words was unaffected by the type of probe used.

The apparent contradiction in these two sets of studies can be resolved by the suggestion that in the Hyde and Jenkins type of study, subjects are using the retrieval information to reconstruct the event, while in the Bregman type of study, the information is used to select the target from other recent items. The factors which induce subjects to use retrieval information in one way or another have not yet been elucidated, but one strong possiblity is that subjects use the information to select the target if scanning back through recent events is feasible -- that is, if the subject judges that the target item occurred fairly recently. A study by Craik and Jacoby (1975) bears out this conjecture. Subjects were induced to encode words either structurally (is the word in upper case letters?), phonemically (does the word rhyme with TRAIN?), or semantically (is the word a piece of furniture?). Interspersed with these encoding trials were recognition trials -- target words were re-presented for recognition after 0-23 intervening trials. Figure 1a shows that while recognition declined with increasing lag, the type of encoding question made no difference to the recognition rate. The study thus replicates the essential features of the experiments by Bregman, Shulman and Jacoby. However, after the encoding and recognition phase of the study was complete, subjects were unexpectedly asked to recall all the words they could; the results are shown in Figure 1b. Now there is a dramatic difference in the patterns of results -- semantic encoding is better than phonemic which exceeds structural, and at each level, the questions which yielded "yes" responses on the encoding trial, gave rise to better recall than those which yielded "no" responses. This latter pattern of results replicates the study reported by Craik (1973) and is similar to Hyde and Jenkins' (1969) and Shulman's

Insert Figure 1 (a and b) about here

(1971) findings. Thus, in the same study, different types of encoding may or may not yield differential memory effects. Our suggestion is that no coding differences are found when subjects use the scanning or selector retrieval strategy, but that coding differences emerge when the reconstruction strategy is used. It seems likely that at least some short-term memory phenomena may be attributed to the use of the selector retrieval strategy (Craik & Jacoby, 1975).

A final point of interest is the observation that there is at least one other example of two "retrieval modes" in nature. We have in mind the manner in which honey bees convey information about a food source to their hive mates, (Von Frisch, 1953). If the source of honey is near the hive, the bee returns and performs a particular series of movements or dance, other bees follow her and smell the odour of the honey she has brought back. In this case, the finder does not pass on location information, she simply gives the other bees the chance to smell the honey which in turn enables them to determine the kind of flower to look for. When they leave the hive they fly in all directions but many soon find the new food source since they have the information to <u>select</u> appropriate flowers. If the food source is far from the hive, however, a second type of dance is performed which gives information about distance and direction as well as information about the kind of flowers to look for. There is an interesting parallel here to our two suggested modes of retrieval from memory — if the event was recent, the retrieval information is used to select the desired episode from other recent events; if the event was not recent, however, retrieval information is used to locate the items directly in the memory system.



Craik and Lockhart argued that deeper, semantic processing yielded a more "durable" code. An alternative view, which fits more readily into the present scheme, is that all encoded events are equally durable — traces are not lost from the system — but that some traces become impossible to access because they are not distinctive, they are similar to many other events. Thus only very recent events of this type can be retrieved — and they are accessed by means of the scanning operation. Events which are encoded in a richer, semantic fashion, may be accessed by either retrieval mode. If the event is recent or if the subject is under time pressure, he may choose to locate the item by scanning, otherwise he may use more of the retrieval information to gain access to the trace; in the first case, semantic information will not yield superior performance or evidence of a more durable trace (since semantic distinctiveness in this surface sense is not a better basis for discrimination than phonemic or structural distinctiveness) but in the second case, semantic information will be superior since the encoded event is more distinctive and the episodic trace can be contacted more readily by semantic retrieval information. This speculative account leads to the interesting prediction that if an event is encoded semantically, it may or may not yield superior performance depending on the retrieval mode used. Evidence for this is provided in the studies of Jacoby (1974) and in the data shown in Figure 1.

Finally, it should be stressed that the scanning or selector retrieval mode is just as applicable to recognition as it is to recall. If a word is presented for recognition, and the subject believes that the word has been presented recently he may abstract some salient feature of the probe word and search back through recent episodes using the abstracted features as

selector information. In this sense, short-term probe situations (where some feature is given as the retrieval information) may yield as good performance as short-term recognition where the whole item is provided — an example is Shulman's (1970) study in which homonym probes yielded as good performance as identical probes.

Some Empirical Issues Re-examined

Before considering certain theoretical questions, we should like to examine several direct consequences of the preceding analysis for some empirical issues of current interest.

Organization and Recognition

Our analysis permits a relatively straightforward account of the role played by organization in the recognition process. The term "organization" is being used in its most general sense to refer to the incorporation of a nominal item into a larger perceptual or conceptual unit. Thus we wish to include the incorporation of an item into a superordinate category, the grouping or association of subsets of items within a list, and the interaction between nominal items (such as head and light) to form a functional item which may bear no simple relation to its component parts. It is doubtful whether these three forms of organization represent any important theoretical differences, and elements of all three might be operative in the encoding of any particular item.

The first point to be noted is that there can be no general answer to the question of whether or not organization facilitates recognition; it may facilitate, hinder, or have no effect at all, depending on the relationship between such organization and the conditions at the time of the test. Insofar as organization represents increased depth and distinctiveness of processing, and test conditions are such to encourage identical processing of the probe, recognition should be enhanced. For one retrieval mode, such conditions would lead to a rapid reconstruction of the episodic trace, while if the scanning mode were adopted, they would ensure the selection of an effective cue. However, in recognition, the test-cue conditions are very much under the experimenter's control, and it is this fact that permits the deliberate manipulation of the similarity between the organizational processing of the item as originally presented and when re-presented as a probe.

It has become popular in recent years to demonstrate that the manipulation of probe processing can result in substantial decrements in recognition performance. Such results are scarcely surprising and the techniques used are similar to thoseemployed by Gestalt psychologists to demonstrate the effects of context, stimulus segmentation, embedding, etc. Thus ambiguous figures are replaced by ambiguous words (Light & Carter-Sobell, 1970; Tulving & Thomson, 1971) or the test word represents a component of the original item whose separate identity was embedded in a larger whole as in Kohler's (1947) demonstration that elements of a previously seen figure may not be recognized if subsequently shown in isolation (Horowitz & Manelis, 1973; Tulving, 1968b).

It has been more difficult to demonstrate positive effects of organization on recognition. One reason for this is that the similarity between item and probe encoding is likely to be very high, even in a control or "neutral" condition in which the experimenter does not explicitly set out to manipulate the relationship between the encodings. Several experiments have shown that encouraging organization during list presentation enhances

recognition (D'Agostino, 1969; Jacoby, 1972; Bower, Clark, Lesgold, & Winzenz, 1969; Mandler, 1972). These effects probably reflect the deeper processing necessary to achieve the categorization, such processing being effective for recognition unless the encoding of the test item is explicitly biased in the manner described in the previous paragraph. There is some evidence that if test order is varied, recognition is highest when the order most closely approximates study order (Jacoby, 1972; Jacoby & Hendricks, 1973; Light, 1973). This facilitation is interpreted in terms of the reinstatement of the study context at the time of test, thereby inducing greater similarity between the two encoding operations.

A further reason for the difficulty of showing positive effects is that recognition can be enhanced by organization only to the extent that other sources of information do not provide an adequate basis for recognition. Retention of physical or sensory characteristics may provide an alternative basis for recognition and may, in certain circumstances, prove to be more reliable. For example, an experiment by Horowitz and Manelis (1973) found that a context change between study and test was ineffective when items were presented visually rather than auditorally. This elimination of context effects was attributed to the additional source of physical information gained from visual presentation. Recognition on the basis of physical characteristics would also help explain the results reported by Davis, Lockhart and Thomson (1972) who found that recognition performance increased with item repetition, even when such repetitions occurred in a semantically misleading context. If physical or sensory characteristics become ineffective as a basis for recognition more rapidly than do semantic elements, then the influence of semantically based organization should become more evident with longer retention intervals. Such a result has been reported by Mandler (1972).

Optimal encoding

An issue which has received recent attention is the question of test-appropriate strategies -- that is, the notion that subjects process items differently depending on the type of test anticipated. This idea leads to the concept of optimal encoding which refers to that encoding of the stimulus which anticipates (and thus makes maximum use of) those retrieval cues available at the time of test. (See, for example, Monk & Cooper; this volume). Most discussions of retrieval processes have considered the differential effectiveness of retrieval cues relative to a fixed, stored trace; but it is possible to consider the matter in reverse -- the relative effectiveness of encoding operations for a given set of retrieval conditions. Insofar as the nature and availability of retrieval cues is fixed by the experimenter, performance may be enhanced by encoding operations which take these conditions into account.

Because recognition and free recall represent very different retrieval conditions, these paradigms provide a natural starting point for the investigation of test-appropriate strategies. Evidence that subjects adopt different encoding strategies depending upon whether a recall or recognition test is anticipated is still rather scant. However, there are now a few experiments showing that recognition is more rapid (Frost, 1972) and accurate (Carey & Lockhart, 1973; Tversky, 1973) when recognition rather than free recall testing is anticipated. If cued recall represents an intermediate retrieval situation relative to free recall and recognition, then an intermediate form of encoding should be optimal. This idea was confirmed by

Jacoby (1973); cued recall performance was higher when subjects anticipated cued recall rather than free recall on recognition. Thus there is evidence that subjects develop encoding strategies that are appropriate for cued recall and recognition tests.

In free recall studies, the evidence is less clear. A frequent finding has been that the overall probability of free recall is independent of the form of retention test anticipated (Carey & Lockhart, 1973; Jacoby, 1973; Tversky, 1973, Exp. I). This overall similarity may mask subtler encoding differences, however. For example, in Jacoby's (1973) study, further analyses revealed that subjects who had been led to expect a cued recall test, recalled fewer categories but more instances of each category than the subjects who correctly anticipated the free recall test. In the Carey and Lockhart experiment, more detailed analyses showed that the probability of free recall declined across intracategory serial positions when subjects were anticipating a free recall test but remained stable across positions when a recognition test was anticipated. Thus, it appears that the same overall level of free recall was attained in different ways dependent upon the form of retention test that was anticipated.

In contrast to the studies cited above, Tversky (1973, Exps. II & III) has demonstrated an overall free-recall advantage for subjects anticipating a free recall rather than a recognition test. However, in Tversky's experiments, the words used were unrelated, not members of a limited number of categories with category instances presented together as in the studies by Carey and Lockhart (1973) and Jacoby (1973). In addition, Tversky's subjects who were told to expect a free recall test were encouraged to organise the study material. Thus it seems possible that Tversky's different results are due to the greater lattitude her subjects had in organising the material; in both the Jacoby and the Carey and Lockhart experiments, organisation was heavily constrained by the presentation conditions.

As well as anticipating the retrieval environment, optimal encoding must also take into account the length and nature of the retention interval. This aspect is particularly important under the conditions where study time is limited, and some strategy of distributing study time and effort must be adopted. There is good evidence that subjects adopt different encoding strategies as a function of the anticipated retention interval. When retention was tested after a long delay, performance was higher for those subjects who expected the delay than for subjects anticipating a shorter retention interval (Gotz & Jacoby, 1974). Also, Evans and Jacoby (1975) have shown that the provision of additional study time is more beneficial for memory over the long term if a delayed rather than a shorter retention test is anticipated; in the "delayed" case subjects may indulge in deeper, semantic processing.

Similarly, several recent studies have shown that "primary memory" items are poorly retained in a second, delayed recall test (Craik, 1970; Craik and Watkins, 1973) unless a delayed test is anticipated (Jacoby & Bartz, 1972). The processing which maximizes long-term retention may actually be less than optimal for immediate or short-term recall — that is, if subjects are induced to form deeper, elaborative codes, such encoding appears to be less efficient for immediate recall than shallower phonemic encoding (Mazuryk & Lockhart, 1974). In the case

of negative recency in free recall, the poorer long-term retention of terminal items appears to be a direct consequence of an encoding strategy which optimizes immediate recall.

Much research remains to be done before a detailed account of optimal processing can be given. In the meanwhile, however, the possibility of test-appropriate encoding strategies should be considered before it is assumed that nominally identical study conditions are also functionally equivalent.

The effects of recognition on subsequent retrieval

We have suggested that recognition is executed via one of two possible retrieval modes. Even within each of these modes, however, the amount of processing necessary to make the recognition decision can vary considerably. In a very short-term recognition test, for example, it may be possible to make a "same" decision on the basis of a few, relatively superficial (sensory) features and thus to make the decision quite rapidly. At longer retention intervals such features may no longer be effective, and deeper, more extensive analysis of the probe may be necessary. Thus the act of recognition itself can demand varying depths of processing, depending on such factors as the retention interval, the similarity of surrounding items, and the distinctiveness of the original episodic trace.

One method of indexing this increased processing is to measure decision latencies (Okada, 1971; Murdock & Duffy, 1972). For our present purposes this method has a number of disadvantages. It is frequently difficul to distinguish the time occupied by processing within the memory system fron that spent in making the actual decision, and this division of time is likely to be sensitive to trade-off effects between speed and accuracy.

Moreover, it cannot distinguish between the quality, as opposed to the quantity of processing and in the context of the present discussion, this distinction is crucial.

A preferred method of evaluating the process involved in recognition is to examine the effects of such processing on subsequent retrieval. That is, it should be possible to index the depth of processing involved in a given act of recognition by examining its effects on a subsequent retrieval. This approach is tantamount to treating recognition as an incidental orienting task and using it in the manner suggested by Craik and Lockhart (1972).

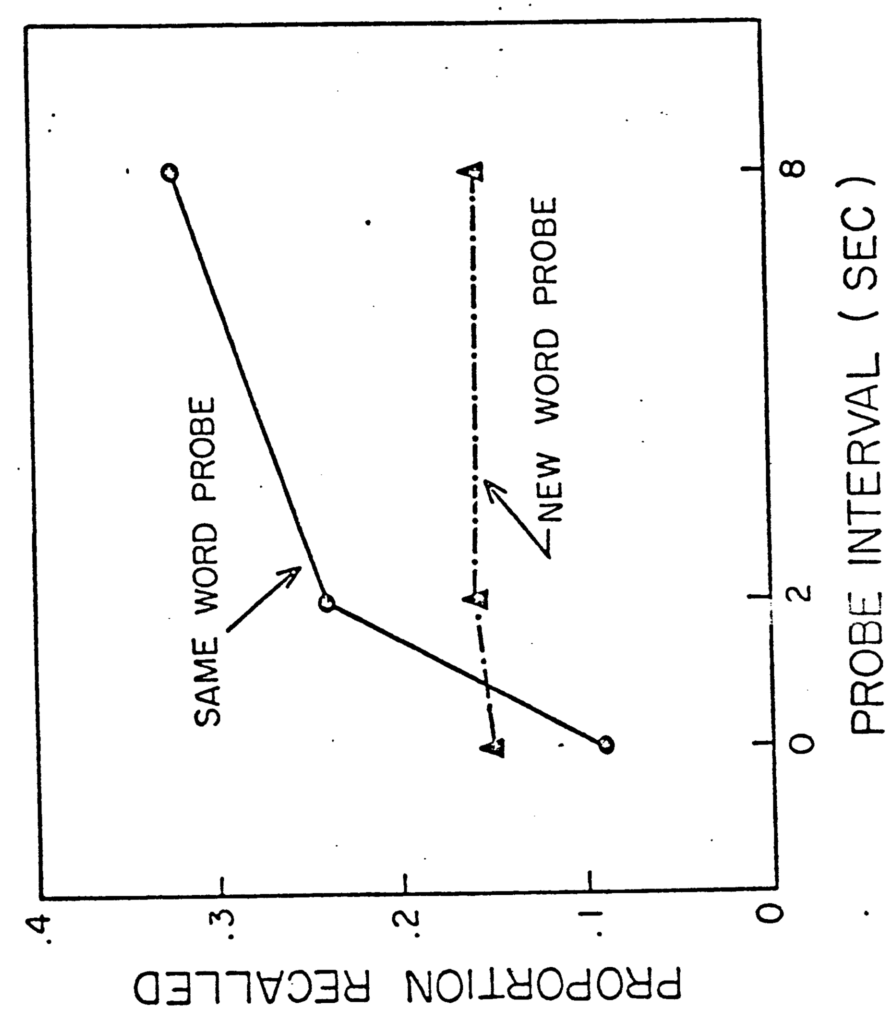
Consider an experiment in which subjects are given a series of single item Brown-Peterson recognition trials. A word is presented, the subject counts backwards for 0, 2 or 8 seconds and then is presented with a probe word to which he must respond "same" or "different" as rapidly as possible. Under these conditions virtually no errors occur, but after 45 such trials an unexpected final free recall is given. The question of interest is whether final recall is a function of the length of the retention interval between the initial presentation of the item and its probe.

The results from such an experiment are shown in Figure 2 and the answer is clearly affirmative; when the probe is the same word, final free recall increases with the duration of the filled interval. No such increase occurs if the probe is a new word, a result which rules out an explanation in terms of covert rehearsal during the retention interval.

While the detailed explanation of these results may take many forms, their general interpretation seems quite clear; as the initial retention

---

Insert Figure 2 about here



interval increases (and with it the degree of forgetting of the item) the depth of processing of the probe necessary to arrive at a "same" decision also increases and thus, as a by-product, subsequent retrieval is facilitated.

Several points concerning these results can be made. Firstly, they suggest that even over the restricted range of retention intervals considered (0-8 seconds) there are substantial differences in the qualitative aspects of the retrieval processing involved in making the recognition response. Accounts of recognition memory which ignore the analysis and processing of the probe itself are missing what is perhaps the major aspect. It is important to remember that it is some encoded version of the probe and not the probe itself that is utilised to contact the stored trace.

Secondly, the results add weight to the conviction that examining the consequences of recognition for subsequent retrieval may be a fruitful approach to analysing recognition itself. It should be possible to establish systematic relations between the conditions under which recognition occurs (probe context, lure similarity, etc) and subsequent retrieval (free or cued-recall, recognition, etc.). A recent study by Bartlett and Tulving (1974) further exemplifies this approach.

<u>The word frequency effect</u>

It is well established that whereas common words (like TABLE, DOG) are better recalled than rare words (like GIMLET, ATOLL), the superiority is reversed in recognition -- that is, rare words are recognized better than common words (Kintsch, 1970; Gregg, this volume). From the standpoint of the present system, two factors are invoked to account for the

recognition result: First, common words are relatively easily encoded and thus the resultant episodic trace is not particularly rich, rare words demand more analysis and result in a richer trace. Second, the attributes comprising the trace of a rare word form a more distinctive and unique event in episodic memory; thus when the word is re-presented for recognition, the encoded probe word evokes the episode more easily. It seems likely, also, that rare words give rise to more nearly identical encodings on successive occasions -- ATOLL or GIMLET are thought of in the same way on each exposure while TABLE or DOG may lead to somewhat different encodings from one occasion to the next. Another way of expressing this difference is to say that the meaning of low frequency words is less sensitive to changes in context. By this view, the encoded attributes of a low frequency (rare) probe are more likely to "overlap" the attributes of the encoded first presentation. It is suggested, then, that rare words are better recognized since their traces are richer, more unique and since successive presentations yield highly similar encodings.

Furthermore, since the meaning of a low frequency word is less obvious and since the word itself thus demands more processing before its gist is extracted, more attention will be paid to the word's surface features -- its constituent letters, phonemes and syllables. Thus the encoded trace of a low frequency word may contain a substantial "physical" or surface features component. One experimental result which fits this view was reported by Schulman (1967); he found that polysyllabic words (compared to shorter words) boosted recognition performance for low frequency words but not for high frequency words -- the polysyllabic words may have resulted in a richer physical trace for the rare words.

In the present system, recall is conceptualized as the guided reconstruction of the original encoding from the basic information provided by the retrieval information. On the other hand, recognition relies less on the reconstructive efforts of the system since more retrieval information is provided by the stimulus; in this case the uniqueness and distinctiveness of a low frequency word are the factors which give rise to its superior recognition.

We would like to get away from Kitsch's (1970) proposal that recall consists of search and decision phases while recognition involves only the latter phase. As Tulving and Thomson (1973) point out, this model rules out the possibility that recall could be superior to recognition but in fact, this outcome is demonstrated by these authors. However, rather than talk about recognition and recall being "on a continuum," we prefer to think of the processes reflecting different aspects of the same system. Recall and recognition have some operations in common; the similarity between the procedures will depend on the particular situation. Thus, since recall and recognition reflect different questions being asked of the system, it should not be too surprising that some variables such as word frequency, affect the processes in different ways.

Some Theoretical Issues Reconsidered

The effects of expectancy on encoding

It has been argued that the episodic trace is a by-product of the perceptual operations carried out on the stimulus input. Since these operations themselves may be modified by various factors, it follows that the episodic trace will be rich or impoverished depending on the amount of perceptual processing necessary to carry out the original perceptual

task. One factor which influences the ease of perception is expectancy. If a particular word is expected (due to prior experience, immediate context or instructions, for example) then it will be recognised easily, with a minimum of perceptual analysis; if the word is not expected, more processing will be necessary to identify the word. Thus the end product — word identification — is the same in both cases, but since fewer perceptual operations are required in the first case, the resultant episodic trace will be less rich than the second trace.

Expectancy serves much the same role for encoding as set does in problem solving. Set may facilitate or retard problem solving, and similar effects are expected with the encoding of verbal material. Facilitating effects usually stem from increasing efficiency as intermediate steps are deleted in a manner similar to practice at a skill. We would like to carry this parallel one step further by examining its consequences for retention. Increasing the efficiency of encoding by the dropping out of steps will have the consequence of excluding these steps from the episodic trace. Thus, paradoxically, when expectancy facilitates initial encoding, it reduces the richness, and later effectiveness, of the resulting memory trace.

Before going on to deal with some negative effects of expectancy on encoding, one positive example will be cited. One function of expectancy is to focus attention on those aspects of the stimulus that are important for confirmation or denial of the expectancy. This focussing of attention has the effect of reducing the number of aspects of the stimulus that are processed and affects memory in that only those aspects of the stimulus that have been given attention will be remembered. One example of this principle comes from an experiment by Quartermain and Mangan (1956). Subjects were instructed to find a key that had been hidden in the room. In a later test of their retention for the contents of the room, subjects were best able to remember those things that could have contained a key. Instructions to find a key led subjects to deal with only particular aspects of the total stimulus situation and it was only those aspects that had been dealt with that were well remembered.

The negative effects of expectancy are really further examples of the same general principle. When context makes a word highly expected, it is necessary to deal with only a few of the word's physical features in order to identify it. Thus if the word is embedded in a context of continuous prose, physical features such as misspellings, type font, and so on, can go unnoticed and will form little or no part of the episodic trace.

In addition to limiting the physical aspects of a stimulus that are dealt with during encoding, expectancy or prior experience in a situation may also reduce the number of cognitive operations required to arrive at an encoding. This point is best illustrated with an example. Suppose that a subject has been given the problem of adding 35 and 17. After he has produced 52 as a correct answer, he is immediately given the same problem. On this second occasion, the subject is likely to give the sum from memory rather than constructing a solution by adding the two numbers. The external stimuli, 35 and 17, have served only to bring the sum to mind so that the relationships intermediate to the final solution have been completely bypassed. As a result, the intermediate relationships will not be represented in the memory trace of the second problem occurrence

and cannot be used later to aid in reconstruction of either the problem or the solution. Again, the same principle is likely to apply to the encoding of verbal material. Retention over the long term requires more than simple pronunciation of an item. If an item is to be remembered, it must be placed in relationship with other events so that a structure involving the item is constructed. As in the arithmetic example, prior experience can limit the number of operations required in the construction. A case in point is the transition of a metaphor to a cliché. As a result of frequent use, the components of a metaphor are no longer compared in order to construct a meaning. Rather, the metaphor leads directly to a meaning in the same way that the two numbers in the arithmetic example led to their sum.

Examples given above have illustrated two ways in which prior experience or expectancy can influence the encoding of a stimulus. First, stimulus priming may occur so that few physical aspects of the stimulus are attended to in order to identify it. In addition, the operations required to give the stimulus a place in a structure may be abbreviated or totally by-passed due to the reproduction of a prior encoding from memory. In contrast to an interference theory of forgetting (e.g., Postman & Underwood, 1973), the present position emphasises the encoding of a stimulus input as well as the process occurring during the retention interval. The failure to recognize or recall particular details of a stimulus may be due to failure to deal extensively with those aspects during encoding as well as an effect of interference during the retention interval or at the time of test. The next few paragraphs will illustrate how our position can be applied in the interpretation of several memory phenomena.

The first data areas to be examined are those of learning to learn and proactive inhibition. Increased efficiency in performing a task after practice on similar tasks is a common observation in both verbal and non-verbal situations. For example, when several unrelated lists of paired-associates are learned, fewer trials are required to reach a performance criterion on lists presented later in the series. One interpretation of this improvement (Warr, 1964) is that subjects are performing the same operation for all lists but quickening the performance of those operations with practice. That is, subjects are learning the same thing about words in later lists as those in earlier ones but doing it faster. This appears to be the interpretation of learning to learn that underlies attempts to control degree of original learning in studies of proactive inhibition. The notion is that degree of original learning must be controlled by equating performance levels on an immediate test before differences in rate of forgetting can be detected (Underwood, 1964). When multiple lists are learned, later lists are presented for fewer trials since a performance criterion is reached more rapidly due to the effects of learning to learn. Proactive inhibition is then shown by the rapid forgetting of lists learned later in the series and explained in terms of interference between lists at the time of test. According to our view, attempts to equate degree of original learning have been unsuccessful and proactive inhibition is at least partially due to subjects learning less about later lists in the series.

Rather than a quickening of operations, learning to learn can be seen as the result of a reduction in the number of operations performed. Words in earlier lists are dealt with more extensively in order to meet

the same performance criterion as words in later lists. Some operations performed on earlier lists are unnecessary to satisfy the performance criterion imposed on the immediate test so they are eliminated from the learning of subsequent lists. However, some of those operations that have been eliminated are important in aiding the reproduction of list items in the long term. If this view is correct, the amount learned is not equated by continued training on successive lists until a common criterion is gained or by the more elaborate techniques proposed by Underwood (1964). A performance criterion can be gained in a number of ways so an equivalence in immediate recall performance does not imply equal learning in the sense of equally rich encodings. An alternative approach would be to equate lists on the number of exposures or study time rather than on a performance criterion. When this is done, proactive inhibition is largely eliminated (Warr, 1964). However, equating study time is not a totally satisfactory solution either. There are still likely to be differences in the particular operations performed as a function of learning to learn, although study time is equated. A subject that has gained experience with a task in unlikely to act as if he were naive simply because he is required to spend the same amount of time on the task as he did originally. If we are to understand proactive inhibition, we must first understand the operation of learning to learn. Proactive inhibition is likely to be a consequence of changes in operations performed on the study material as well as interference between lists at the time of test.

The notions described earlier can also be applied to the effects of proactive inhibition in the Brown-Peterson paradigm. Our analysis would hasize the effect of prior lists on the encoding of items in later sts. When instances of the same category are presented on successive als, subjects may come to expect instances of that category so that operations necessary to identify a presented category member are previated. In this regard, the situation is analogous to the example presenting the same addition problem twice in succession. For items esented later in the series, dimensions or attributes that are common ross lists can be given directly from memory, and require little new nstruction.

The last data area to be examined in this section is that of the acing of repetitions. Both recall and recognition are generally found be enhanced when repetitions of an item during study are spaced ther than continuous. Our interpretation of the spacing effect is that e continuous repetition of an item leaves the analysing processes so ghly activated or "primed" that further encodings of the item involve ry few operations. This view is very similar to that presented by ckhart (1973). That is, the final product can be repeated or rehearsed thout constructing that encoding anew when repetitions are massed. ain, the situation is quite similar to the example of solving the same ithmetic problem twice in a row. The sum of the numbers or a final coding can be repeated without repeating all of the operations which re originally necessary to construct the sum or final encoding.

Data supporting our interpretation of the spacing effect come from veral sources. First, there is now a large body of data showing that action time to a stimulus input is decreased if that input is a repetition of an immediately prior event (e.g., Remington, 1969). Thus, massed

repetitions either quicken the operations or, as we would prefer to argue, reduce the number of operations that are required, prior to responding. In the retention literature, the effect of spacing of retentions has been attenuated when each repetition is accompanied by a different modifier (Jacoby, 1972; Gartman & Johnson, 1972; Madigan, 1969). This effect has been interpreted as showing that retention is enhanced by increasing the number of access routes to an item in memory. However, it is just as reasonable to conclude that a change in modifiers is important because it requires the subject to go through constructive acts prior to arriving at a final encoding of a repetition. That is, for massed repetitions without a new modifier, the subject may go directly to an earlier final encoding rather than constructing it. With spaced repetitions, the encodings of a repeated item might be quite similar, but due to forgetting during the interval between presentations, the subject has constructed the encodings for each of the repetitions. One very nice piece of evidence for this position comes from an investigation by Hintzman, Block and Summers (1973). In that investigation, each presentation of an item was presented in a different modality, auditory or visual. A later test for modality of input then allowed retention to be traced to each presentation of an item. Application of this technique revealed that it was retention of the second presentation of a repeated item that suffered under conditions of massed repetition. This result is precisely what would be expected if encoding of a massed repetition was highly primed so that a subject paid little attention to the physical characteristics of the item in order to identify it.

The STM/LTM distinction

Does the present system retain the distinction between short-term and long-term memory? In their previous model Craik and Lockhart (1972) proposed that memory is essentially one system, but that subjects can also retain information by maintaining activity in some part of the analysing structures. This "maintenance processing" at some specific level of analysis was seen as equivalent to keeping the item in consciousness, paying attention to the item, rehearsing the item or retaining the item in primary memory. The characteristics of short-term memory defined in this way would thus not be absolute but would depend on the level of analysis at which processing was maintained -- at deeper levels, encoding would be more semantic in nature, "capacity" would be greater in that more use could be made of learned rules, and the rate of forgetting would be slower since deeper processing yields a more durable code.

In the present model we want to keep the notion that retention in primary (or short-term) memory is equivalent to continued activation of some part of the analysing structures, but also incorporate Jacoby's (1974) suggestion that many short-term memory phenomena are attributable to the retrieval strategy used -- that is, recent items are scanned, and the target selected on the basis of some characteristic specified by the retrieval question. The main point we wish to make is that the notion of one short-term memory mechanism with a range of specific characteristics (limited capacity, acoustic coding, etc.) may be an oversimplification -- rather, we should think of the phenomena of recent memory and ascribe these phenomena to the different questions asked, and to the different strategies employed by the system.

First, some short-term phenomena are ascribed to the continued activation of the analysing structures with the subjective concomitant that the items are still "in mind"; memory span tasks and the Sternberg reaction-time task (when small set sizes are used) may be described in these terms. Other phenomena of recent memory, involving matching or probe techniques may be tapping the scanning or "looking back" strategy (Jacoby, 1974) -- Shulman's (1970) results and Waugh and Norman's (1965) probe-digit technique are examples of this mode of operation.

As a further suggestion, the reconstruction strategy may also be more effective for recently presented items -- since the specific processing operations have recently been performed and possibly because recent traces provide more effective feedback and guidance to the reconstructive processes. Recency effects with this retrieval strategy should be relatively "long-term" in nature (that is, not transient fragile effects) and such recency phenomena have been reported by Bjork and Whitten (1972) and by Tzeng (1973). In summary, we believe that it still makes sense to distinguish "short-term" from "long-term" memory, but the characteristics of short-term retention will depend not only on material and task (which in turn will influence the depth of encoding), but also on the retrieval strategy utilised by the subject.

Recognition and Recall — concluding comments

We are now in a position to make some concluding comparisons between recognition and recall. In the first place it seems clear that there is little to be gained by attempting to answer such global questions as whether or not recognition involves a search or retrieval phase and whether or not it can be distinguished from recall on this basis. Such questions presuppose a view of memory which is altogether too simple. The only sensible approach is to examine the form of processing necessary to execute the recall or recognition and then to document their similarities and differences.

We have argued that two basic modes of retrieval exist for recognition — reconstruction and scanning. The same two modes exist for recall and in this sense it is possible to argue that recognition and recall are basically the same process. However, within each mode (whether it be recall or recognition) the processing may differ with task conditions. In the case of the reconstruction mode, the amount and depth of processing necessary to construct an encoding which will "resonate" with the episodic trace can vary considerably, and if the term "retrieval" is to be assigned a meaning which is more specific than the operational definition given in the introduction, it should be in reference to this process of reconstruction. In the case of recall, the reconstruction mode can be seen to operate in a similar fashion, the only difference being in the nature of the experimenter-provided information. In the case of recognition, the reconstruction process necessary to achieve resonance may be no more than the normal, and highly practiced analysis of meaning that typically occurs when a word is read or heard. In such cases recognition or familiarity would seem,

in subjective experience, to be direct and involuntary — without the conscious awareness of a search process. In fact, if such analysis of the probe yields little or no resonance, the item will typically be rejected as new. Thus if the context of the probe is such as to bias a very different meaning (as in Tulving & Thomson's 1971 study) a miss will frequently result. If resonance is present but weak, additional processing of the probe may be executed in an attempt to resolve the ambiguous status. Mandler's post recognition retrieval check can be viewed as one form that this additional processing might take. As the degree of such processing extends beyond the initial "reading" of the item, processing may be under more conscious control and a subjective sense of search memory may result.

In the case of recall, reconstruction of the episodic trace will typically demand more extended processing simply because the experimental situation provides less retrieval information. We have argued that this extended processing is not one of "generate and recognize" but rather a continuous interplay between the reconstruction process and the episodic trace. Of course, a compelling feature of the "generate and recognize" model is that recalled items must, in some sense, have been acknowledged as "old" prior to overt recall; after all, subjects do not typically emit items they judge not to have been in the list. According to the views presented here, however, such final editing would usually be trivial, if indeed it occurred explicitly at all. Items which are generated as candidates for recall are not fortuitously old or new in a way analogous to a yes/no recognition test. The retrieval process itself ensures a high probability that such items will be list members. The important editing occurs, not following an item's retrieval, but as an integral part of the processing (which we have labelled "guided reconstruction") involved in the items production in the first place.

Recognition and recall, then, are not related by the former being a sub-component of the latter; rather, one is typically a more extended — and in that sense more complex -- version of the other. Moreover, the terms represent rather loose labels for what in fact is a variety of retrieval conditions and the development of a comprehensive theory of human memory would be better served , not so much by direct comparisons of recognition and recall, but by a more detailed analysis of the processing involved in each.

REFERENCES

Bartlett, J. C. & Tulving, E. (1974) Effects of temporal and semantic encoding in immediate recall upon subsequent retrieval. Journal of Verbal Learning and Verbal Behavior, 13, 297-309.

Bjork, R. A., & Whitten, W. B. (1972) Recency - sensitive retrieval processes in long-term free recall. Paper presented at the meetings of the Psychonomic Society, St. Louis, Missouri.

Bower, G. H., Clark, M. C., Lesgold, A. M. & Winzenz, D. (1969) Hierarchical retrieval schemes in recall of categorical word lists. Journal of Verbal Learning and Verbal Behavior, 8, 323-343.

Bregman, A. S. (1968) Forgetting curves with semantic, phonetic, graphic and contiguity cues. Journal of Experimental Psychology, 78, 539-546.

Carey, S. T. & Lockhart, R. S. (1973) Encoding differences in recognition and recall. Memory and Cognition, 1, 297-300.

Craik, F. I. M. (1970) The fate of primary memory items in free recall. Journal of Verbal Learning and Verbal Behavior, 9, 143-148.

Craik, F. I. M. (1973) A "Levels of Analysis" view of memory. In P. Pliner, L. Krames, & T. M. Alloway (Eds.), Communication and affect: Language and thought. New York: Academic Press.

Craik, F. I. M. & Jacoby, L. L. (1975) A process view of short-term retention. In F. Restle (Ed.), Cognitive theory: Volume one. Potomac, Md.: Erlbaum Associates.

Craik, F. I. M. & Lockhart, R. S. (1972) Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Craik, F. I. M. & Watkins, M. J. (1973) The role of rehearsal in short-term memory. Journal of Verbal Learning and Verbal Behavior, 12, 599-607.

D'Agostino, P. R. (1969) The blocked-random effect in recall and recognition. Journal of Verbal Learning and Verbal Behavior, 8, 815-820.

Davis, J. C., Lockhart, R. S. & Thomson, D. M. (1972) Repetition and context effects in recognition memory. Journal of Experimental Psychology, 92, 96-102.

Evans, J. D. & Jacoby, L. L. (1973) Long-term storage: Rehearsal or encoding demands. Paper presented at the Forty-Fifth Annual meeting of the Midwestern Psychological Association, Chicago.

Frost, N. (1972) Encoding and retrieval in visual tasks. Journal of Experimental Psychology, 95, 317-326.

Gartman, L. M. & Johnson, N. F. (1972) Massed versus distributed repetition of homographs: A test of the differential encoding hypothesis. Journal of Verbal Learning and Verbal Behavior, 11, 801-808.

Götz, A. & Jacoby, L. L. (1974) Encoding and retrieval processes in long-term retention. Journal of Experimental Psychology, 102, 291-297.

Gregg, V. (1975) Word frequency, recognition and recall. In J. Brown (Ed.), Recognition and recall, London: Wiley.

Hardyck, C. D. & Petrinovich, L. F. (1970) Subvocal speech and comprehension level as a function of the difficulty level of reading material. Journal of Verbal Learning and Verbal Behavior, 9, 647-652.

Hintzman, D. L., Block, R. A. & Summers, J. J. (1973) Modality tags and memory for repetitions: Locus of the spacing effect. Journal of Verbal Learning and Verbal Behavior, 12, 229-238.

Horowitz, L. M. & Manelis, L. (1973) Recognition and cued recall of idioms and phrases. Journal of Experimental Psychology, 100, 291-296.

Hyde, T. S. & Jenkins, J. J. (1969) The differential effects of incidental tasks on the organization of recall of a list of highly associated words. Journal of Experimental Psychology, 82, 472-481.

Jacoby, L. L. (1972) Effects of organization on recognition memory. Journal of Experimental Psychology, 92, 325-331, a.

Jacoby, L. L. (1972) Context effects of frequency judgments of words and sentences. Journal of Experimental Psychology, 94, 255-260, b.

Jacoby, L. L. (1973) Test appropriate strategies in retention of Categorized lists. Journal of Verbal Learning and Verbal Behavior, 12, 675-682.

Jacoby, L. L. (1974) The role of mental contiguity in memory: Registration and retrieval effects. Journal of Verbal Learning and Verbal Behavior, 13, 00-00.

Jacoby, L. L. & Bartz, W. H. (1972) Encoding processes and the negative recency effect. Journal of Verbal Learning and Verbal Behavior, 11, 561-565.

Jacoby, L. L. & Hendricks, R. L. (1973) Recognition effects of study organization and test context. Journal of Experimental Psychology, 100, 73-82.

Kintsch, W. (1970) Models for free recall and recognition in D. A. Norman (Ed.), Models of human memory, New York: Academic Press.

Kohler, W. (1947) Gestalt Psychology. New York: Liveright.

Kolers, P. A. (1973) Remembering operations. Memory and Cognition, 1, 347-355.

Kroll, N. E. A., Parks, T. E., Parkinson, S. R., Bieber, S. L., & Johnson, A. L. (1970) Short-term memory while shadowing: Recall of visually and of aurally presented letters. Journal of Experimental Psychology, 85, 220-224.

Light, L. L. & Carter-Sobell, L. (1970) Effects of changed semantic context on recognition memory. Journal of Verbal Learning and Verbal Behavior, 9, 1-11.

Light, L. L. & Schurr, S. C. (1973) Context effects in recognition memory: Item order and unitization. Journal of Experimental Psychology, 100, 135-140.

Lockhart, R. S. (1973) The spacing effect in free recall. Paper presented at the meeting of the Psychonomic Society, St. Louis, November.

Madigan, S. A. (1969) Intraserial repetition and coding processes in free recall. Journal of Verbal Learning and Verbal Behavior, 8, 828-835.

Mandler, G. (1972) Organization and recognition. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press, pp. 139-166.

Mazuryk, G. F. & Lockhart, R. S. (1974) Negative recency and levels of processing in free recall. Canadian Journal of Psychology, 28, 114-123.

McCormack, P. D. (1972) Recognition memory: How complex a retrieval system? Canadian Journal of Psychology, 26, 19-41.

McNeill, D. & Lindig, K. (1973) The perceptual reality of phonemes, syllables, words, and sentences. Journal of Verbal Learning and Verbal Behavior, 12, 419-430.

Melton, A. W. (1963) Implications of short-term memory for a general theory of memory. Journal of Verbal Learning and Verbal Behavior, 2, 1-21.

Miller, G. A. (1963) Comments on Professor Postman's Paper. In C. N. Cofer & B. S. Musgrave (Eds.), Verbal behavior and learning: problems and processes, New York: McGraw Hill.

Mistler-Lachman, J. L. (1974) Queer sentences, ambiguity and levels of processing. Memory and Cognition, 2, 000-000.

Murdock, B. B., Jr. (1968) Modality effects in short-term memory: Storage or retrieval? Journal of Experimental Psychology, 77, 79-86.

Murdock, B. B., Jr., & Dufty, P. O. (1972) Strength theory and recognition memory. Journal of Experimental Psychology, 94, 284-290.

Neisser, U. (1964) Visual search. Scientific American, 210, 94-102.

Norman, D. A. (1969) Memory and attention. New York: John Wiley and Sons, Inc.

Okada, R. (1971) Decision latencies in short-term recognition memory. Journal of Experimental Psychology, 90, 27-32.

Posner, M. I., & Warren, R. E. (1972) Traces, concepts, and conscious constructions. In A. W. Melton and E. Martin (Eds.) Coding processes in human memory. Washington, D. C.: V. H. Winston & Sons.

Quartermain, D. & Scott, T. H. (1960) Incidental learning in a simple task, Canadian Journal of Psychology, 14, 175-182.

Remington, R. J. (1969) Analysis of sequential effects in choice reaction time. Journal of Experimental Psychology, 82, 250-257.

Savin, H. B. & Bever, T. G. (1970) The nonperceptual reality of the phoneme. Journal of Verbal Learning and Verbal Behavior, 9, 295-302.

Schulman, A. I. (1967) Word length and rarity in recognition memory. Psychonomic Science, 9, 211-212.

Schulman, A. I. (1971) Recognition memory for targets from scanned word lists. British Journal of Psychology, 62, 335-346.

Schulman, H. G. (1970) Encoding and retention of semantic and phonemic information in short-term memory. Journal of Verbal Learning and Verbal Behavior, 9, 499-508.

Sokolov, E. N. (1963) Perception and the conditioned reflex. New York: Macmillan.

Sutherland, N. S. (1972) Object recognition. In E. C. Cartarette & M. P. Friedman (Eds.) Handbook of Perception, New York: Academic Press.

Treisman, A. (1964) Monitoring and storage of irrelevant messages in selective attention. Journal of Verbal Learning and Verbal Behavior, 3, 449-459.

Tulving, E. (1968) Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), Verbal learning and general behavior theory. Englewood Cliffs: Prentice Hall, (a).

Tulving, E. (1968) When is recall higher than recognition? Psychonomic Science, (b), 10, 53-54.

Tulving, E. (1972) Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Press.

Tulving, E. & Pearlstone, A. (1966) Availability versus accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behavior, 5, 381-391.

Tulving, E. & Thomson, D. W. (1971) Retrieval processes in recognition memory: Effects of associative context. Journal of Experimental Psychology, 87, 116-125.

Tulving, E. & Thomson, D. M. (1973) Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373.

Tversky, B. (1973) Encoding processes in recognition and recall. Cognitive Psychology, 5, 275-287.

Tzeng, O. J. L. (1973) Positive recency effect in a delayed free recall. Journal of Verbal Learning and Verbal Behavior, 12, 436-439.

Underwood, B. J. (1964) Degree of learning and the measurement of forgetting. Journal of Verbal Learning and Verbal Behavior, 3, 112-129.

Von Frisch, K. (1953) The dancing bees. New York: Harcourt, Brace.

Warr, P. B. (1964) The relative importance of proactive inhibition and degree of elarning in retention of paired associate items. British Journal of Psychology, 55, 19-30.

Warrington, E. K. & Sanders, H. I. (1971) The fate of old memories. Quarterly Journal of Experimental Psychology, 23, 432-442.

Waugh, N. C. & Norman, D. A. (1965) Primary memory, Psychological Review, 72, 89-104.

Wickens, D. D. (1970) Encoding categories of words: An empirical approach to meaning. Psychological Review, 77, 1-15.