THE EFFECTS OF LIMITED STIMULUS INFORMATION ON WORD PROCESSING

THE EFFECTS OF LIMITED STIMULUS INFORMATION ON WORD PROCESSING

by

Stephen Rivers, B.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Arts

McMaster University

April, 1973

MASTER OF ARTS (1973) McMaster University

Hamilton, Ontario

TITLE: The Effects of Limited Stimulus Information on Word Processing

AUTHOR: Stephen Rivers, B.A. (Carleton University)

SUPERVISOR: Dr. Betty Ann Levy

NUMBER OF PAGES: vii, 62

SCOPE AND CONTENTS:

Words and pseudowords were tachistoscopically presented to subjects who had to verify as fast as possible whether the items were members of large or small semantic categories and to decide whether or not the items were in fact words. A dictionary model of word processing was tested, once at the respective thresholds for making wordness and semantic decisions and once at longer presentation times for purposes of comparison to other studies. Response latencies suggest that at short presentation times, when processing is not allowed to advance to the semantic level, a dictionary process seems plausible, whereas at long presentation times retrieval of semantic information may originate from stored categories of words. Final recognition data suggest that items paired with the wordness questions were not processed to as "deep" a level as items paired with the semantic questions.

ACKNOWLEDGEMENTS

This thesis was prepared under the supervision of Assistant Professor, Betty Ann Levy, Ph.D., of the Department of Psychology of McMaster University, Hamilton, Ontario, Canada.

The writer is also indebted to Professors I. Begg and P.L. Newbigging for their advice and encouragement during the revision stages of this paper.

The research reported in this thesis was supported by Grant A7657 from the National Research Council of Canada to Dr. B.A. Levy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

In 1970 David Meyer presented a paper at the Psychonomic Society meeting which was concerned with the cognitive processing of words and pseudo-words. The title of the paper was "Parallel Processes in Word-Recognition" and underscores Meyer's contention that the process involved in deciding whether an item is a word or a meaningless string of letters has two components which operate independently; the operation of one does not necessitate the operation of the other. One component deals with the extraction of structural information and the other with the extraction of semantic information from the item. To identify a stimulus as a word, complete processing of either one of these components is sufficient for a correct decision since structural analysis allows the item to be compared to or "matched" with a stored memorial representation and semantic analysis provides the information that the item has meaning and in most instances meaningful arrays are necessarily words. The data which led to this formulation in Meyer's study are incompatible with the idea that structural and semantic processing are not independent. This notion has been described as a "dictionary process" since the two components of the process operate serially. That is, semantic processing necessitates the completion of structural processing. Just as one looks up an item in a dictionary, definitions cannot be read until the word has been located in the book. Analogously, structural processing leads to locating the item's "address" or position in memory and semantic processing involves "reading" the definition and associated words.

This thesis is concerned with a specific methodological problem related to Meyer's paradigm. In Meyer's experiment, subjects inspected words and pseudowords which were presented non-tachistoscopically and asked to make certain decisions about them. He measured response latency for making these decisions and found that it took just as long for subjects to simply decide that an item was a word as to decide that it was a member of a semantic category. This finding together with supporting data, led Meyer to reject a dictionary model of word processing since this alternative predicts that latencies to wordness decisions (deciding whether the item is a word) are shorter than latencies to semantic decisions (deciding whether it belongs in a semantic category, for example).

The present thesis recognizes the possibility that a serial exhaustive process can also account for Meyer's data. That is, failure to find latency differences between the two types of decision may not have resulted from the operation of independent processes. It may have been that when either decision is made, both components necessarily operate in a serial fashion to completion. The proper but operationally difficult control to have added to Meyer's study was a test for semantic knowledge of the stimulus item at the same time asking the subject a wordness question. This would have determined whether processing had continued to completion in the wordness condition.

The methodological innovation in this paper attempts to circumvent this problem. Instead of presenting stimulus items for long periods of time it was decided to present them tachistoscopically at different presentation times employing the ascending method of limits. As well, a masking stimulus was presented on termination of the test stimulus which was an attempt to arrest processing at the structural stage. If the entire process is serial with the possible feature that output can be initiated after the completion of processing at the structural stage, then correct responses to wordness decisions should be available for output at a faster presentation time than correct responses to semantic decisions. Furthermore, latencies to the former decision process should be faster than to the latter process. These predictions were generally realized in the present study.

The underlying theoretical problem seems to be one of retrieval of stored information for use in decision making. Therefore the structural and functional aspect of long-term semantic memory are important theoretical issues. If, for example, a dictionary process is involved in word processing, then the implication is that retrieval originates from the test item itself. Specifically, subjects might check whether the semantic category asked for is associated with the test item. If they are asked whether a test item such as PEN belongs to both a large and small semantic category on different occasions, then one would expect category size to have little effect on negative responses. Latency to decide that PEN is not an animal should be the same as latency to decide that PEN is not a mammal, since neither category is commonly associated with PEN.

Meyer's model on the other hand, suggests that retrieval can be initiated without first locating the item in storage, since the wordness and meaning components of processing operate in parallel. One alternative retrieval strategy is to search the semantic category in question. Each word stored in such a category would have to be searched until a match is either found in the case of words or not found in the case of pseudowords. The data in the present experiment do not strongly support such a view since there was no effect of category size on response latencies for negative responses.

The thesis which follows is organized into five main sections. First, a brief overview of several theories of long-term semantic memory are reviewed to familiarize the reader with some of the more popular models in the literature. Some of the ideas contained in them are later used to formulate hypotheses about how a particular dictionary retrieval model might operate. Second, selective reviews of two areas of research, Selective Attention and Verbal Learning, are presented as background for the present study. After formal presentation of the experiment itself which includes sections three and four, the results are interpreted in terms of a specific dictionary model and Craik's (1972) "levels of processing" conceptual framework is employed in this discussion.

OVERVIEW OF SOME THEORIES OF LONG TERM SEMANTIC MEMORY

In recent times several theories have emerged, related to how semantic information is stored in and retrieved from the brain. The main point of departure among them is in terms of how an incoming stimulus event is matched with a stored memorial representation of itself. Function and structure seem to imply one another in all of these models since the structure is inferred from the data which themselves reflect the functional aspects of the system. Therefore, when one speaks of how the system functions he is almost always making statements about how the memory store is organized as well.

Four models are briefly presented below. The first is a Dictionary model and includes a description of an attempt to create an analogous system using this model, in a computor. The second is called the Logical Inference model and stresses the interrelationships among memorial elements. The third is the Categorical model and is used by Meyer to explain some of his 1970 data. The fourth, which appears in another one of Meyer's papers (1971), is called the Semantic Distance model.

<u>Dictionary model.</u> Dictionary models which relate to how semantic memory may function must be differentiated from a dictionary process which may be involved in word processing. A dictionary process deals with encoding variables and implies nothing about how semantic memory

is organized. It simply implies that an incoming percept must make contact with a stored memory unit before semantic processing can begin. This is depicted in Figure 1. A dictionary process may be contrasted with notions such as parallel or simultaneous processes between "wordness" information and semantic information. Such a process has been elaborated by Meyer and Ellis (1970). Their "parallel race process" (see Figure 2) suggests that the process involved in deciding that an item is a word is independent of the process involved in making semantic decisions about the item. The inverted bell curves in the figure signify normal distributions of processing times for each process.

A dictionary model on the other hand, may encompass semantic variables as well as the item identification process. For example, Katz and Fodor (1963), Quillian (1967) and Paivio (1971), describe the semantic component of long term-memory (LTM) along the following lines.

The lexical elements of semantic memory are coded along some dimension at the time of perception and then stored in the brain in such a manner as to facilitate retrieval and cognitive associating. To accomplish this, these memory elements or "entries" are organized like a dictionary. According to Paivio (1971, p.421),

> "The lexical information contained in the different classes of markers is expressed in the form of a tree diagram. Figure 3 shows such a diagram for the word bachelor. The representation includes a grammatical marker (noun), semantic markers (human, animal, male) in parentheses, and distinguishers (who has never married, etc.) in brackets."

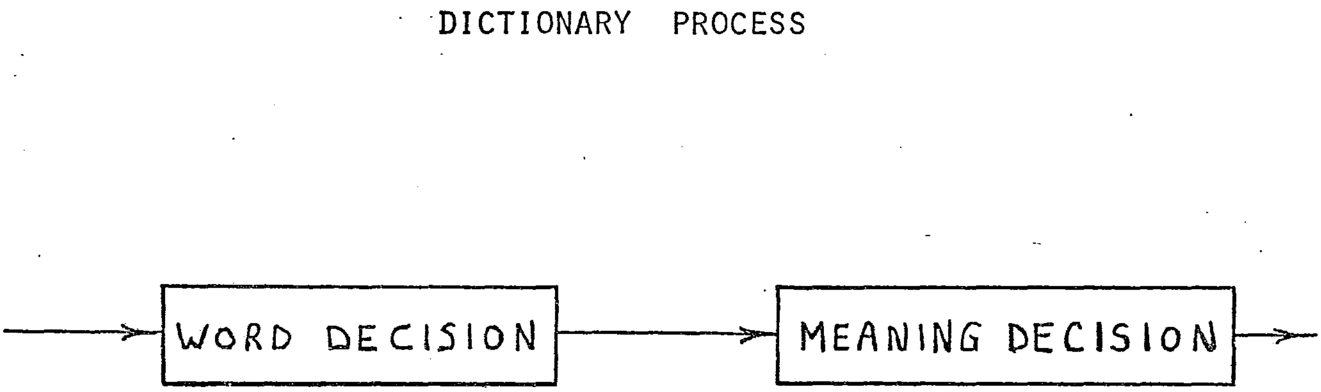

Figure 1 Dictionary process for recognizing whether items are words and retrieving their meanings

PARALLEL "RACE" PROCESSES

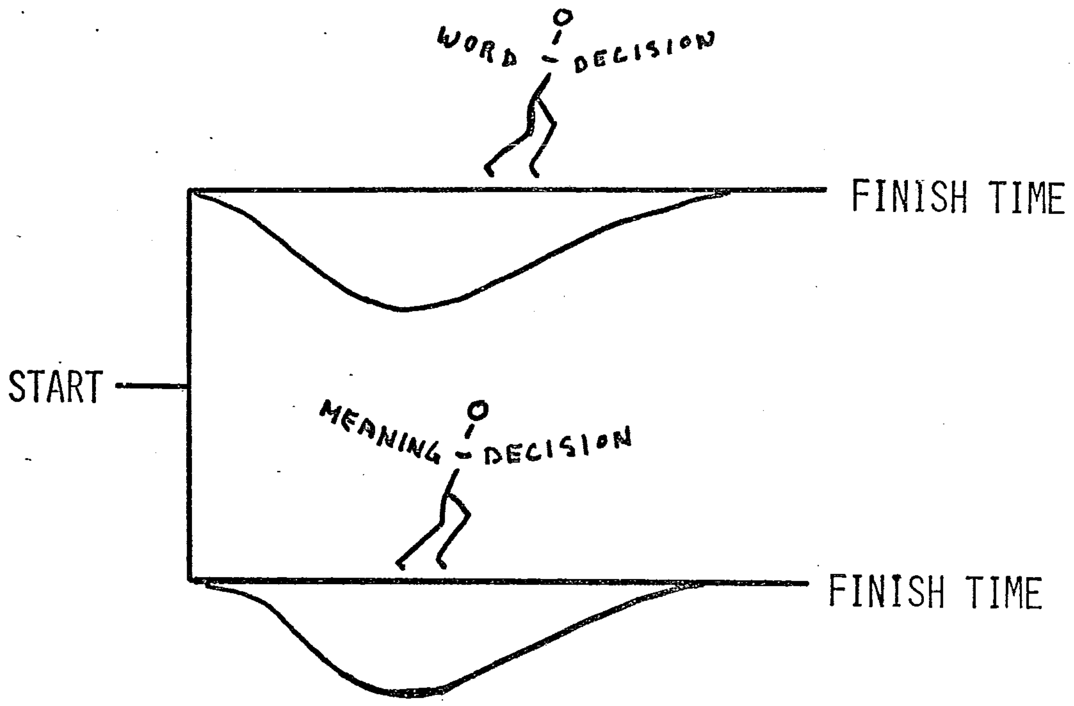

Figure 2 Parallel "race" processes for recognizing whether items are words and retrieving their meanings (From Meyer and Ellis, 1970)

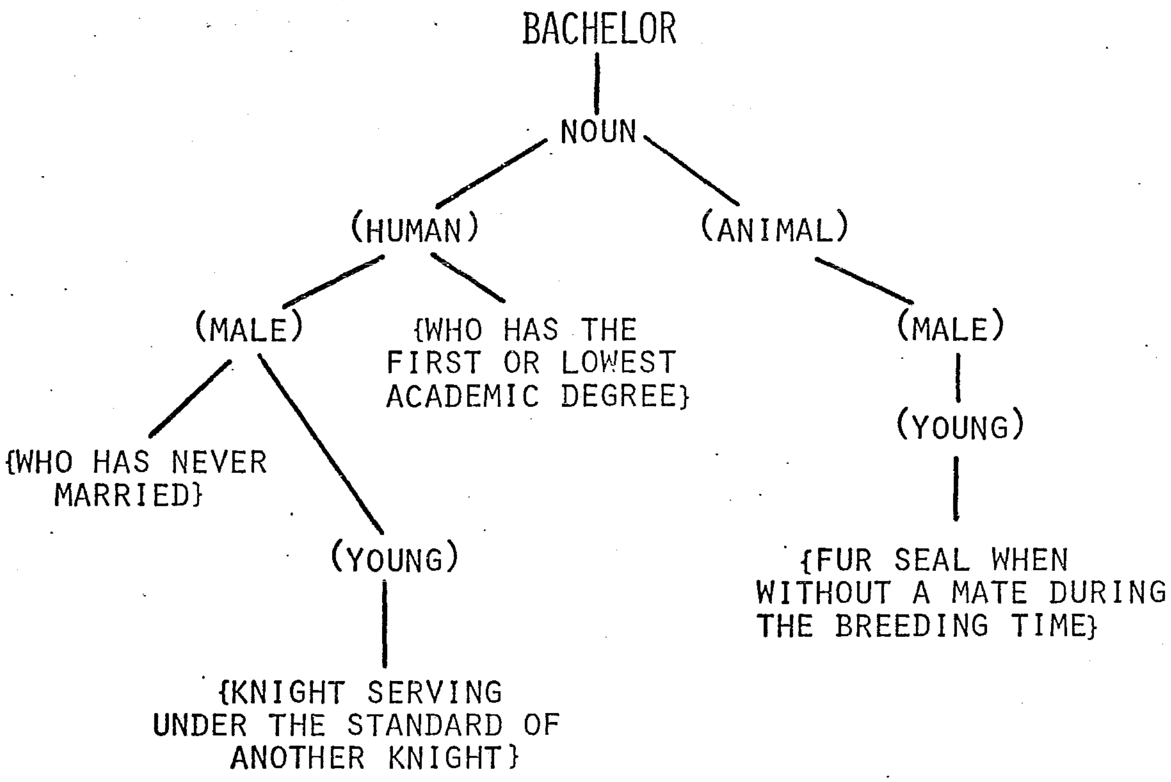

Figure 3 Dictionary entries for the word <u>bachelor</u> in terms of grammatical markers, semantic markers and distinguishers. (From Katz and Fodor, 1963)

In this model then, concrete verbal items are stored as individual coded entries or "markers" together with their conceptual category names and succinct definitions.

Quillian (1967) reviews and discusses work done on building computor programs which use dictionary models of long-term semantic memory. The model memory is made up of nodes, each of which can be thought of as a word label although Quillian prefers "properties" to "words." There are several types of nodes in the model which allow for different ways of relating to the meaning of a name word, either directly by "leading directly into a configuration of other nodes that represents the meaning of the name word" (called type nodes), or indirectly by "having one special kind of associative link that points to that concept's type node" (called token nodes). Configurational meaning of a concept is built on a "plane" of token nodes and this represents its "immediate definition." Thus, according to Quillian, "a word's full concept is defined in the model memory to be all the nodes that can be reached by an exhaustive tracing process, originating at its initial, patriarchical type node, together with the total sum of relationships among these nodes specified by within-plane, token-to-token links. (p.413)."

Ideas such as alphabetical arrangements of the dictionary store and non-random access to various sections of the store have been hypothesized and empirically tested, although most of the evidence is inconclusive. The important point to note is that all dictionary models assume that semantic decisions emanate from the test items themselves. Semantic information cannot be arrived at without activation of the dictionary "unit." As will be seen, other models suggest that semantic information may be retrieved by means other than direct access to the memorial representation.

Logical inference model. One line of theorizing allows for semantic relationships and associations to structurally exist in the memory store. This structure is described in a paper by Collins and Quillian (1970), and takes the form of a "map" of logical inferences (see Figure 4). That is, items in LTM are semantically related by virtue of sharing common properties (e.g., cat, tiger, leopard), or due to continual contrasting (e.g., dog, cat). In this inferential network, the more closely items are related the more confusable they are. The concept of semantic confusability appeared in an earlier paper by these researchers (Collins & Quillian, 1969), wherein they reported that it took longer for subjects (Ss) in a true-false force choice task, to reject false statements involving semantically related concepts than to reject false statements involving semantically unrelated concepts. For example, in one experiment (1969), sentences like A COLLIE IS A CAT (where dog and cat are confused) took longer to reject as false than sentences like AN ELEPHANT HAS A BILL. This they attributed to the fact that when items are confusable, the exhaustive matching process of overlapping attributes consumes more time than when items do not share many common attributes.

Categorical model. The work of Landauer and Freedman (1968) was motivated by the notion that semantic memory may be organized on the basis of semantic categories. According to these authors, one discovers whether or not the word COLLIE is in the memory store, by searching word cate-

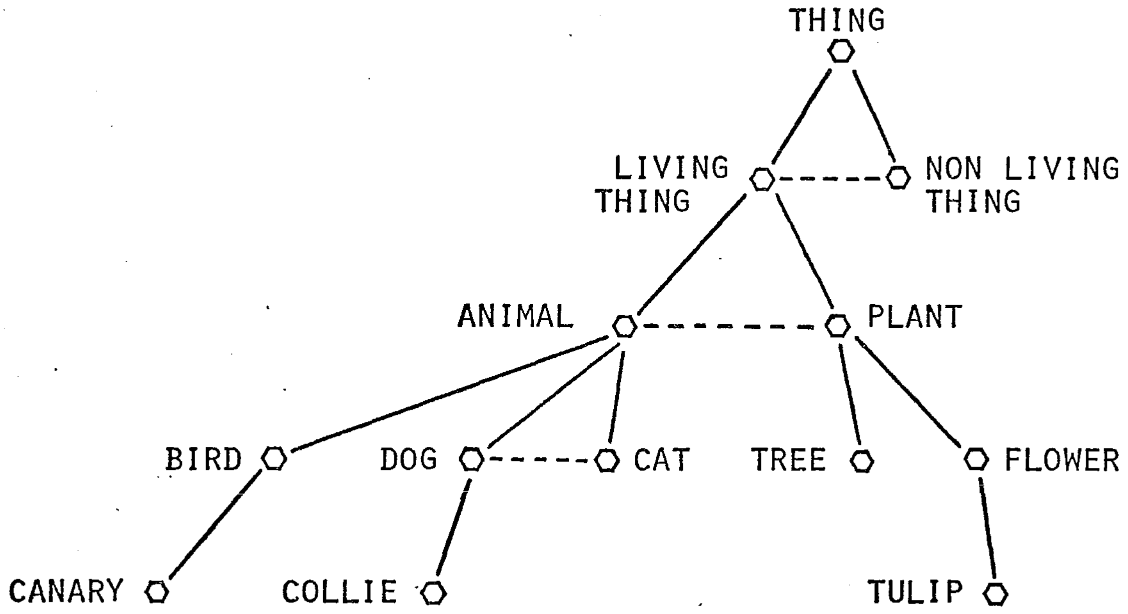

Figure 4 A partial map of long-term semantic memory
(From Collins and Quillian, 1970)

gories such as ANIMAL and DOG. If a "match" is found, one is then potentially able to retrieve and report this item. This is in contrast to Collin's and Quillian's idea that one retrieves items by channelling information through networks of inferential associations.

The technique used to empirically investigate Landauer's and Freedman's claim involves subordinate-superordinate semantic categories. A subordinate-superordinate relationship between categories means that, by definition, all the members of the subordinate category belong to a subset of the superordinate category. Consider the categories MAMMAL and ANIMAL. MAMMAL is a subordinate category of ANIMAL, since all of the members of the category MAMMAL belong to a subset of the category ANIMAL. That is, all mammals are animals, but all animals are not necessarily mammals. Thus, if a member of the subordinate category such as HORSE is selected, this item must belong to both the subordinate and superordinate categories. If two groups of _S_s are then compared, one group being asked whether a HORSE is an ANIMAL, the other being asked whether a HORSE is a MAMMAL, the former group should have longer response latencies than the latter. This is true since it would take more time to search the larger, superordinate category for _exactly the same item_ than the smaller, subordinate category.

Semantic distance model. Another suggested departure from the dictionary model appears in a recent paper by Meyer (1971). Whereas Collin's and Quillian's innovation pertained to a logical structure imposed upon the basic dictionary model, Meyer's seems to be more of a spatial one. He postulated that words which are semantically related, are stored closer together than words which do not share common semantic characteristics. For example, the pair of words, BIRD and INSECT have a smaller "semantic distance" than the pair of words BIRD and CITY. This is why, he explained, it took longer for Ss to decide that a word like BIRD is not a CITY than to decide that a BIRD is not and INSECT (1971).

In a recent article by Rips, Shoben and Smith (1973), the concept of semantic distance is examined and used to explain some of the conflicting data found in the semantic memory literature. Both network models such as Collin's and Quillian's and "set-comparison" or categorical models such as Landauer's and Freedman's, are reinterpreted in terms of the semantic distance variable.

In Collin's and Quillian's network model for example, there may exist more intermediate nodes such as FOWL between the nodes CHICKEN and BIRD than between the nodes ROBIN and BIRD. This would account for their prediction that deciding that a CHICKEN is a BIRD takes longer than deciding that a ROBIN is a BIRD. Rips, et al. concluded that "in general, the semantic distance between an instance (node) and a category reflects the number of intermediate nodes involved, and the latter determines verification time."

Further theoretical consideration. The theoretical issue to which this paper is directed involves the basic dictionary model but an extension of it which incorporates some of the other models as well. Specifically, if the dictionary model of human semantic memory is viable, then there are other questions which must be raised. For example, how does one know that an incoming item is in fact in the dictionary store? Can the dictionary model predict how one determines that an item is a word and not a non-word? If it can, on what basis does one make this kind of decision? Does the item have to travel through inferential maps or semantic categories as Landauer and Freedman would have to predict? Is contact with any stored semantic information necessary to decide whether an item is a word? Or is it possible, as a dictionary process predicts, to dispense with "reading" definitions, or searching categories, etc., when making a "wordness" decision? It might be that a more efficient way of dealing with this kind of decision is to base it on structural information alone (i.e., letter combination probabilities, etc.) and to omit the semantic stage of processing. Just as one looks up a word in a dictionary by starting with the first letter and trying to locate its "address" or position in the book without the benefit of any semantic information, one may similarly retrieve items in memory by first locating its "address" on the basis of structural information obtained and then "reading out" the semantic information located there. The process of obtaining information about an item's wordness may be independent of the process of obtaining semantic information about that item. At the same time, this process would also imply that obtaining semantic information is entirely dependent on the item first being identified as a word. If this is in fact the case, and could be experimentally demonstrated, it would certainly illuminate some of the theoretical issues about how semantic memory is organized. If a dictionary process is valid, then certain models such as the Categorical Search model would have to either be rejected or modified to incorporate the idea of retrieval emanating from the test item and not from categorical searches. It is primarily to such problems that the following thesis is addressed.

SELECTIVE REVIEW OF LITERATURE

It was suggested in the last section that "wordness" processing and "semantic" processing may not be identical. A dictionary process suggests that wordness decisions are independent of semantic decisions but semantic decisions are dependent on the completion of wordness decisions. This possibility derives primarily from two lines of experimental evidence, Selective Attention and Verbal Learning. In this section, these areas of investigation will be briefly summarized with respect to the present study, prior to a statement of the hypotheses.

Human selective attention literature. Work in the area of human selective attention (see reviews by Norman, 1969; Moray, 1969; Swets & Kristofferson, 1970), has brought some evidence to bear on the problem of how visually presented verbal information is processed. Investigations in the auditory modality show that individuals have the capacity to attend to particular stimulus events (occurring in a cluster of events), either by selecting certain events for attention to the exclusion of others or by selectively rejecting ("filtering out") undesired items. One controversy in the selective attention literature involves the locus and function of such a filtering mechanism. Treisman (1967), using Cherry's (1953) shadowing paradigm and results from earlier studies (Moray, 1959; Mowbray, 1964; Treisman, 1960), developed a model which specifies the locus and function of the filter. Briefly, the model assumes that information arriving through parallel channels is analyzed for gross physical characteristics quite early in the processing hierarchy. These features are processed at this point and are necessarily available for further processing. However, if an item is not selected for further processing, no more than its physical features will be available at later stages, and an attenuated version of the original exists somewhere in the system for a short time subsequent to filtering. An item which has been attended to finally reaches a stored "dictionary unit," presumably in LTM. Firing this unit simultaneously causes recognition and conscious perception of the item.

The idea that information which has been filtered can be made available for perception, was challenged by Deutsch and Deutsch (Deutsch, Deutsch & Lindsay, 1967). They contended that Treisman's hypothetical filter mechanism must be very complex indeed, since it must change and adapt to almost every new situation. Furthermore, it must receive "instructions" from a more central location as to what is "desired" or to-be-remembered information. On the grounds that Treisman's formulation necessitates a filter mechanism almost as complex as the central mechanism for which it is reducing input, Deutsch and Deutsch (1963) maintained that an input must go directly to the dictionary unit intact, where each item is fully processed for word identity and meaning. From this point on, selection is achieved on the basis of response salience, i.e., importance to the organism (e.g., hearing one's own name). In this model, a later filter is hypothesized; responses compete for conscious awareness only after items have been fully analyzed structurally and semantically. This is in contrast to Treisman's model, which allows for early selection on the basis of structural and/or acoustic variables.

Despite the Deutschs' well taken point, Treisman's filter mechanism is a necessary component of a word-processing dictionary model. Assuming that an item has already been acquired and is set up in a particular address, a filter which selects relevant information pertaining only to the item's address, seems not only plausible but more direct and functional. If the desired information about an item is simply "Is this a word?," then semantic information becomes superfluous, for there is no reason in this case for exhaustive processing to take place.

Verbal learning literature. A second line of evidence from the area of human verbal learning also relates to the problem of whether or not verbal items must be fully processed before information about them can be used (made functional) by a person.

Until the late 1950's, LTM was the only type of memory systematically researched in the literature and it was the development of the distractor technique in 1959 (Peterson & Peterson) which prompted the notion of a short-term memory system (STM). Although Craik & Lockhart (1972) advanced a new conceptualization of how memory is organized which suggests that memory should not be thought of as separate "boxes" or systems but as a continuum, continued research (see review by Tulving and Madigan, 1970) has demonstrated that these two memory systems may exist and differ quite dramatically in their structure and functions. Short-term memory seems to have a capacity to hold not more than about seven items (Miller, 1956) for only a few seconds (Peterson & Peterson, 1959) without rehearsal. Long-term memory, on the other hand, seems to have no limit to the amount that can be stored, provided rehearsal and some sort of meaningful organization (e.g., chunking) has been imposed upon the items. Interference is postulated as the primary agent of forgetting in LTM, whereas both interference (Keppel & Underwood, 1962; Melton, 1963) and decay (Conrad, 1967) account for information loss in STM. These two systems have also been differentiated on the physiological level. Milner (1959) found that Ss had trouble learning new material once she lesioned the hippocampus. Short-term memory seemed intact however, since Ss could rehearse the items, thereby preventing decay. This distinction can be restated in Hebbian terms: items in STM are maintained by reverberating circuits (rehearsal), whereas more permanent structural changes must take place when an item transfers to LTM (Hebb, 1949).

Variables relevant to STM and LTM have been extensively studied using the probe digit techniques. When Ss are asked to recall an item from a particular position in a list, which is designated by the experimenter (E) after the list is presented, it has been demonstrated that those items most recently presented have a higher probability of being recalled correctly than items presented earlier in the list. This phenomenon is appropriately called the "recency effect" and is also found using the free recall paradigm. Waugh and Norman (1965) interpreted these findings in terms of a dual system of memory and employed William James' (1890) terminology of primary memory (PM) and secondary memory (SM), which correspond functionally to STM and LTM respectively. It is the functioning of PM, wherein recent items are still being rehearsed, which accounts for the recency effect in free recall. (Glanzer & Cunitz, 1966.)

This formulation is also compatible with Craik's "negative recency effect." Craik (1970) found that when Ss were asked to free recall the same list on subsequent occasions, it was the final words in the free recall list which were least well retrieved. This finding demonstrated, according to Craik, that items in PM (i.e., the last items processed) are not as well "registered" as items in SM (i.e., the earlier presented items) and therefore are not retrieved with as high probability at a later time.

This line of evidence led to closer examination of the encoding dimensions which might be specifically related to PM and SM. Ordered recall of a complete set of letters (Conrad, 1962, 1964) or digits (Wickelgren, 1965) is a technique where six-letter or six-digit sequences are presented visually at a rapid rate. The subject is required to write down the letters immediately after the last item is presented. Conrad found that most of the errors which occurred were confusions between letters which sounded alike. For example, when B was the correct response, the subject was much more likely to write down a P or V than a W or F. These results were taken to mean that STM operates as a sort of acoustic "echo-box," since errors made with visually presented material were of an acoustic nature.

Subsequent work by Baddeley and Dale (1966), led to the postulation that since interference in STM is acoustic in nature, meaning must not play an interfering role until the item is transferred to LTM, and therefore all interference in STM is acoustic, while interference in LTM is semantic. This dichotomy received further support from studies by Baddeley (1964), Kintsch and Buschke (1969), Levy and Murdock (1968), and Tulving and Patterson (1968).

This research demonstrates that Ss do emit responses in STM tasks which are relatively free of semantic confusability. This suggests further, that meaningful organization, that is, semantic associations of the types mentioned earlier, have not occurred at the time of response output. If the assumption is made that semantic processing occurs at a deeper level of processing than acoustic-articulatory types of processing, and that semantic processing does not occur in PM, then it follows, as Treisman's model predicts, that input, structural analysis and output are all possible prior to semantic decoding.

Meyer's experiment. In a recent article, Meyer and Ellis (1970) argued against a simple dictionary process which assumes that prior to any semantic decision about an item it must first be located in memory and identified as a word. In Quillian's (1967) computor analogy, this corresponds to the activation of a type node. It is not until after this event occurs that semantic associations are established. Both Treisman's and the Deutschs' models allow for this process to occur, but the Deutschs' would have to maintain that the entire dictionary entry is read prior to the availability of any information.

Meyer employed a technique first used by Landauer and Freedman (1968) to ascertain response latency (processing time) for the semantic categorization of verbal items. Meyer's modification involved presenting Ss with one of three types of question prior to each stimulus exposure. The first set contained questions which asked whether the item to be shown belonged to a large semantic category. For example, IS THIS A KIND OF STRUCTURE? was asked to begin a trial. If the item HOUSE was then shown, the correct answer was "yes." The second set contained questions which asked whether the item to be shown belonged to a smaller semantic category. IS THIS A KIND OF BUILDING? also required a "yes" response when the stimulus HOUSE was presented. The third type of question was, IS THIS A WORD? Naturally, a "yes" response was appropriate in this example. Subjects were presented with words and pseudowords (constructed by changing one of the letters in a word), and response time (RT) measures were obtained from onset of stimulus display to the time S pressed a response button. Essentially, Meyer found that it took Ss just as long to respond to questions pertaining to whether the items were words (wordness decisions), as to respond to questions which required Ss to place items in a large semantic category (meaning decisions).

Because the dictionary model predicts that latencies to IS THIS A WORD? should be faster than to semantically relevant questions, Meyer rejected it. Instead, he postulated parallel decision processes (p.6, l. 12), between "wordness" and "meaningfulness," neither of which was well defined. His alternative "race" model posits that if either decision finishes the processing "race" first, output will be contingent on that independent decision, regardless of the outcome of the "loser."

There are several problems with this formulation. First, one cannot reject the dictionary model on the basis of latency data alone. As was pointed out, both Treisman's and the Deutschs' models allow for a dictionary-like process. However, if the Deutschs are correct, output may simply be impossible before semantic processing is complete. Thus Meyer's latency data may merely reflect output initiated after the completion of a dictionary process. In this case, it would take the same amount of time to report that the stimulus event was a word, as to report information pertaining to its semantic classification.

Second, Meyer referred to "parallel processing" in somewhat the same way Sperling (1967), Estes and Taylor (1966) and Eriksen and Lappin (1965, 1966) did. They postulated that several items in a visual display may be processed as one unit or "chunk," causing the observed minimal variance in latency scores as display size increases. Meyer argued analogously that wordness and meaning are independent variables which have specific processing mechanisms. If so, semantic decisions could be made without information that the item is a word. This leads to what would appear to be a maladaptive system (an inefficient one at least) -- one that searches semantic categories, gaining nothing from the outcome of the word decision. For example, if the outcome of the word decision is negative further semantic processing could be immediately aborted. If these processes are independent, semantic processing would continue even after the item has been identified as a non-word.

In Meyer's paradigm, _S_s had the maximal amount of external stimulus information necessary to complete any processing which might have taken place, whether or not a dependent (dictionary) or an independent (parallel) system exists, since stimulus material was not presented tachistoscopically with a masking stimulus. This allows for an exhaustive dictionary process to be an alternative explanation of Meyer's data. The present issues might be better understood by establishing the minimal amount of information the _S_ requires to complete either a correct wordness decision or a correct meaning decision and at the same time interrupting further processing. By using tachistoscopic presentations and employing the ascending method of limits (see Candland, 1968, Chapter 4), threshold measurements could be obtained for each type of question Meyer used. A "threshold" could be defined as the minimal stimulus exposure duration required for S to complete and respond to a specific decision process. If the presentation duration for wordness questions is less than the presentation duration for semantic categorization questions, then it can be assumed that an early filter mechanism exists, responding to structural variables (e.g., phonetic, lexicographic, etc.),which can initiate response output prior to the completion of semantic processing. This assumption is based on the notion that, if output in the wordness condition can be accomplished to the same degree of accuracy as output in the semantic condition with less stimulus information, then the information required for the latter condition is superfluous to the former, and thus the information used for the wordness decision must somehow be effectively utilized without the necessity of complete processing.

This would help clarify the early vs late filter polemic, but it would still be necessary to study RT measures at each of these duration thresholds, to make statements about the relative rate of processing, which is crucial to any pro-dictionary argument; i.e., a difference in duration threshold for each stage of processing does not necessarily imply that the events occur in the same temporal order as the difference suggests. It merely reflects the minimal amount of stimulus information necessary to make a particular response. If Meyer's experiment is replicated with the stipulation that the data be observed at the respective thresholds for each type of decision process, and it is found, as Meyer found, that RT to wordness and meaning questions still do not vary significantly, then the dictionary model can be rejected with more certainty.

Masking effects. Inherent in the present paradigm is the problem of precise control of the amount of information which undergoes processing. It is not simply a matter of manipulating stimulus duration or intensity tachistoscopically, for it is well documented that stimulus information remains available for up to about one second in a brief sensory store. Sperling (1960; also Averbach & Coriell, 1961; Averbach & Sperling, 1961) went further to demonstrate that visual stimuli do not go directly from retinal images and afterimages to the brain for processing and storage in short- and long-term memory respectively, but intermediate filtering, or "holding" mechanisms might preprocess the information before it reaches the cognitive level of the processing hierarchy.

The advocacy for the existence of a very brief, transient storage-processing system for visual information, coined by Neisser (1967, Chapter 2) the iconic system, stems from the following anecdotal evidence. Despite Miller's (1956) seemingly indisputable claim that a human cannot process or hold more than seven (plus or minus two) discrete items in STM without coding, an individual can, for example, scan reams of names in a telephone book in search of a particular one. Even though he in some way inspects all the names in the list, he does not actually "retain" them in any sense of the word. The iconic storage system was therefore conceptualized to account for an individual's seemingly inexplicable ability to hold information for very brief periods -- long enough to inspect it for critical characteristics, but too short to remember exactly what it was he saw.

In the now classic study by Sperling (1960), _S_s were repeatedly presented with 3 x 3 matrices of random letters for very brief intervals. When required to report as many of the letters as possible, _S_s were able to verbalize up to a maximum of five or six, which corresponds to Miller's (1956) estimated "span of apprehension." However, when each row of the matrix was cued by a particular tone (e.g., a high tone signalled _S_s to report the top row, a medium tone the middle row, etc.) which was sounded immediately or at varying intervals after the actual array was terminated, _S_s could report items from the appropriate row quite accurately; selectively ignoring other uncued items in the array. The _S_s were apparently "reading" from something remaining after the original stimulus had terminated.

By extending the interval between the termination of the stimulus and the onset of the tone, it has been estimated (Sperling, 1967) that the icon fades quite rapidly, rendering almost no information after about a second. Mackworth (1963) found, by varying exposure duration, that the icon requires about 50 msec to reach full strength. It has also been demonstrated (Kahneman, 1968) that when a visual noise field, called a masking stimulus, accompanies or immediately follows a visual presentation, the function of the icon is severely inhibited. This technique can therefore be used to control the amount of information which leaves the iconic system for further processing.

In the literature on the recognition of letters in meaningful and non-meaningful arrays (e.g., Reichter, 1968; Wheeler, 1970), backward masking (see Kahneman, 1968, for a discussion on backward masking) is employed with little discussion on the theoretical basis for its use. There seems to be disagreement as to the exact effects of a mask on an icon. Sperling (1960) maintains that an effective mask immediately terminates an icon, leaving no further information available. He bases this assumption on a particular experiment (1963) wherein he varied the length of exposure time of an array and presented a masking stimulus after each trial. Sperling found that for every 10 msec increase in exposure time, _S_s could report another letter, which was presumably no more than the mask would allow. Neisser (1967) pointed out, however, that Sperling's linear relationship between exposure time and read-in rate may be a function of increasing the strength (through neural summation or some other physiological mechanism) and therefore longevity of the icon up to a maximum level as reported by Mackworth (1963). Neisser maintained that Sperling's read-in rate is too fast: longer presentation times simply induced stronger icons which were not as easily disturbed by a masking stimulus. Therefore, Neisser hypothesized (1967), that the effect of a mask is to degrade the icon, making information more difficult to obtain from it. If the mask in the present study does not destroy the icon, it would be impossible to use the defined threshold to obtain the minimal amount of information required to perform the wordness decision task, because subsequent to the offset of the stimulus, _S_ could still read from the degraded icon and finish processing.

Evidence in favour of Sperling's destruction hypothesis is presented in a study by Liss (1968). He compared Ss' ability to report tachistoscopically presented random letters at different presentation durations under three masking conditions: 1) backward mask (mask comes on immediately after stimulus goes off); 2) concurrent mask (stimulus and mask are on at same time) and, 3) no mask. As well as accuracy measures, Liss asked Ss on a subsequent matching task to judge the items' apparent duration, brightness, contrast, sharpness, and texture in each condition. He reasoned that since it is known that concurrent masking degrades the icon without affecting processing time, backward masking should show qualitatively different effects if Sperling is correct. In fact, this hypothesis was confirmed. There was a discrete exposure duration (around 20 msec) when Ss in the backward condition began making responses with a high probability of being correct. In the backward mask condition the letters appeared very briefly, with higher contrast at 40 msec, than at 9 msec with no mask, and also appeared with higher contrast in the backward condition under a luminance ratio of 1.0 to .37 than in the concurrent condition. These data suggest that once the icon is established, degradation does not occur in the backward masking condition and support Sperling's claim that a backward mask does indeed stop processing at the time of its presentation.

If the assumption is made that a mask stops further processing, then RT at threshold as defined earlier will reflect processing time for a particular type of decision while controlling for the possibility of exhaustive processing in the word condition. This modification to Meyer's paradigm may be a truer test of the dictionary process.

A pilot study was performed to determine the effects of a masking stimulus on tachistoscopically presented words and pseudowords. It was an attempt to determine whether or not Ss could complete the wordness decision prior to full recognition of the item (i.e., verbalization of the item), assuming that verbalization necessitates semantic processing. If they could, then it would imply that full recognition requires more than "wordness" information. The results indicated that the particular mask used had no effect on the icon. Furthermore, only one S yielded data suggesting that discrimination of words from pseudowords is possible prior to full report of the item.

From the data it is impossible to speculate why two of the four Ss responded significantly below the chance level of 50% correct. These two Ss had a strong response bias to say not-word whenever they were guessing, but this factor should not have affected overall accuracy since it would depress the hit rate (saying "word" when it was in fact a word), but correspondingly augment the correct rejection rate (saying "not-word" when it was a pseudoword).

Since icons have been estimated to last up to one second (Sperling, 1967), and presentation times never exceeded this duration, it is probable that Ss were reading information from the icon. Subjects were just as accurate in the M condition as in the NM condition (Pr = .55 vs Pr = .53 of a correct response in each condition respectively). Thresholds were never more than 10 msec apart for each condition.

The main problem illuminated by this study is that examination of prerecognition accuracy was made without identifying at what stage of processing full recognition responses were initiated. It was assumed that to give an accurate whole report on an item, processing (structural and semantic) must be complete. This, however, is not necessarily true. The appropriate condition to have added is semantic categorization questions, such as those used by Meyer and Ellis (1970).

MAIN EXPERIMENT

Purpose. The purpose of the main experiment was four-fold: 1) to test whether a dictionary process does in fact occur when stimulus information is restricted; 2) to provide further tests for a dictionary model of word processing; 3) to provide evidence related to the structural nature of semantic memory; and, 4) to provide evidence to support either Treisman's or the Deutschs' model. A PDP-8L digital computor with an on-line plotter unit was used instead of the three-field tachistoscope used in the pilot study. This change afforded three basic improvements: 1) provided better control of letter shape, size, illumination and duration; 2) provided an effective noise field which had been used as a masking stimulus in previous experiments; and, 3) enabled fast and efficient data collection and reduction.

Seven hypotheses were set up to test both the dictionary process and a particular dictionary model. The model assumes that an incoming item first makes contact with a stored dictionary unit and that access to the dictionary store is random. That is, each unit has an equal probability of being sampled for a match and that once a match is found, the search is terminated. Semantic information associated with the activated unit is then available for further decision making. It is also assumed that negative responses to the question IS THIS A WORD? when pseudowords are used as test stimuli, involves an exhaustive search through the dictionary units. Further, negative responses to semantic types of questions necessitate exhaustive searches through all the dictionary units' definitions for the particular category in question.

<u>Hypothesis 1.</u> If a Treisman-like dictionary model is viable, than at threshold, yes-responses for the word question type (Q-type) will be faster than yes-responses for either semantic Q-type. This is true since semantic questions must wait until a decision has been made concerning the item's wordness.

<u>Hypothesis 2.</u> No-responses to semantic Q-types for words will not differ from each other. This is true since in this model, semantic categories do not have to be searched and therefore the size of the category to which the item belongs is irrelevant.

<u>Hypothesis 3.</u> No-responses to semantic Q-types will be faster for words than for pseudowords. To decide that a word's definition does not include the critical semantic category name, the word must be located and the definition read. To decide that a pseudoword does not have a definition (which is essentially what must be done to make a semantic decision about a pseudoword), all the units must first be checked in order to be sure that the pseudoword does not have a location or "address." On the average, the former process will consume less time since all the dictionary units do not have to be checked.

<u>Hypothesis 4.</u> No-responses to the <u>word Q-type</u> for pseudowords will be just as fast as no-responses to <u>semantic Q-types</u> for pseudowords. This is true since in both these cases all the dictionary units must be checked when the item is a pseudoword. Once it has been decided that the item does not have an address, it has simultaneously been decided that the item has no meaning.

Hypothesis 5. No-responses to the large semantic Q-type for words will be faster than no-responses to the word Q-type for pseudowords. Since stored units are not organized in terms of categories, locating a word and reading its "large" category name should still take less time than searching exhaustively through all the stored units.

Hypothesis 6. There will be more final recognition errors for items paired with the word Q-type than for items paired with semantic Q-types. If, as Craik & Lockhart suggest, semantically processed items are "deeper" processed items, and therefore better "registered" than structurally processed items, the former should be better remembered than the latter. In the present design, the three Q-types were asked an equal number of times. If "registering" does not differentially affect the three Q-type conditions, final recognition errors should be evenly distributed among these three conditions. By chance, this distribution should be 33.3% for each. If Craik is right, there should be fewer errors than chance would predict in the semantic conditions than in the word condition.

Hypothesis 7. If Treisman's early filter model is viable, then the minimum presentation duration for the word Q-type will be less than the minimum duration for either semantic Q-type. This is true since in Treisman's model, structural information is available for output before complete processing has taken place and complete processing requires more stimulus processing time than structural processing.

The comparisons made in the first five hypotheses were also made by Meyer and Ellis (1970), but at a presentation time which afforded unrestricted stimulus information. It would be worthwhile then, to analyze the data at a presentation time approximating the conditions in Meyer's experiment for purposes of replication. Therefore the tenth ranked presentation time following the threshold presentation time was used as a point of observation of the data for these purposes. For example, if the threshold for the word Q-type was 24 msec or the 3rd ranked presentation time (presentation times began at 8 msec and increased in 8 msec units), then mean latency scores were also obtained at the 13th presentation time.

METHOD

Subjects. Ten male and ten female Ss took part in the experiment. Data from one male and one female were omitted from analyses when it was realized that their first language was not English and that this might be an important variable to control for. Of the 18 Ss remaining, six were graduate students at McMaster University, three were part-time laboratory assistants, and the remainder were undergraduates of the University. Each S was paid $2.00 per session which lasted approximately 35-40 minutes. Each S served for four sessions. All Ss had normal or corrected 20/20 vision. Half of the Ss used their preferred hand for the "yes" response button and the other half used their non-preferred hand for the "yes" button.

Apparatus. The apparatus consisted of a PDP-8L Digital computor, which was connected to a Teletype input-output device, a 5 x 5 inch point plotter and a plain 18 x 12 x 2 inch flat box with four response buttons on it, two of which were activated for this experiment. The plotter screen was at eye level when S was seated. The Teletype was turned on when loading the program and printing out data. During stimuli presentation it was turned off since it was quite noisy. The response board rested on the S's lap; a red light remained on in the room throughout the experiment.

Stimulus items. There were 96 test stimulus items as well as 16 pairs of subordinate-superordinate category pairs, the item "word" and a visual noise field. The 96 items were divided equally into 48 stimuli which required "yes" responses and 48 which required "no" responses. The 48 yes items were chosen in the following manner. First, 16 pairs of subordinate-superordinate categories were selected with the specification that three, 3 to 6 letter, high frequency exemplars of each subordinate category were available. These 16 x 3 exemplars constituted the 48 yes-stimuli. The 16 category pairs with their respective exemplars appear in Appendix A. As well, four other category pairs with one exemplar per pair were used as practice stimuli.

The 48 no-stimuli consisted of 16 words and 32 pseudowords. Pseudowords were constructed by changing one letter in a three- to six-letter AA frequency word (see Thorndike & Lorge, 1944), so that its pronouncability was maintained. The no-words were also three- to six-letter AA frequency words. The above items also appear in Appendix A.

On each trial, prior to the appearance of one of the 96 test stimuli, one of the 33 category items appeared on the screen. That is, subordinate-superordinate category labels appeared or the item "word" appeared. It was understood by _S_, that each of these events implied the question, IS THIS A ____________? For instance, if _S_ saw as the first event, the stimulus MAMMAL, he understood that a question was being asked regarding the test stimulus which was to follow. Implicitly, he read this event as "Is this a mammal?" Similarly, if WORD appeared, _S_ understood that he had to discover whether the test item was a word or not.

The letters appeared in approximately the centre of the screen and stood about 1/4 inch high. The mask was generated by programming a random selection of points from a rectangular matrix, the perimeter of which overlapped the area on the screen where the words appeared.

At the termination of the fourth session, Ss were asked to take a final recognition test which included all the test items. Each item was paired with a distractor which was selected from the same semantic category so that memory of the item rather than discrimination of category attributes would be tested. These pairs were typed on sheets of paper and the Ss simply had to circle the member that they recognized as the one used in the experiment.

Design. The design was a within-subject three-way factorial. All of the independent variables were fixed. The three independent variables were: 1) Question type (Q-type) which took on one of three values -- large semantic (LS), small semantic (SS) or word (W); 2) Stimulus Type (S-type) which consisted of stimuli requiring either yes-responses in order to be correct (yes-stimuli) or no-responses in order to be correct (no-stimuli); and, 3) Presentation duration (P-time) which took on 20 discrete values from 8 msec, and were increased by 8 msec intervals, to 160 msec (i.e., 8, 16, 24 ... 160).

The Q-type and S-type variables were combined to yield an equal number of "yes" and "no" trials (48) and an equal number of LS, SS, and W trials (32). The duration variable was overlaid in blocks of 96 trials for each P-value, thus yielding 96 x 20 or 1920 trials per S. See Table 1 for an illustration of the design.

TABLE 1

WITHIN-SUBJECT DESIGN FOR MAIN EXPERIMENT
WITH ASSIGNED Ns FOR EACH CONDITION

Duration (Msec)			Question Type	
		Small semantic	Large semantic	Word
	Stimulus Types			
8 msec	"Yes" words	n= 16	n= 16	n= 16
	"No" words	n= 8	n= 8	
	"No" pseudowords	n= 8	n= 8	n= 16
				N= 96
16 msec				
24 msec				
. . .				
160 msec				

Note.- The n's are the number of trials per subject for four sessions.

Since it was necessary to allow each S to view the test stimuli only once at each P-time, and because it was desirable to pair each Q-type with identical yes-stimuli, the question-stimulus pairings within categories were counter-balanced across Ss. It was possible to do this since all three exemplars of each category pair were yes-words for the three Q-types. For example, HORSE, DOG, and RABBIT all required yes-responses, regardless of whether ANIMAL, MAMMAL or WORD preceded it. Since there are six ways of combining three items with three other items, three groups containing six Ss were essentially formed. Each member of the groups was assigned a different Q-type yes-stimulus combination so that each of the three exemplars was combined with each Q-type an equal number of times. The no-stimuli were paired randomly with the Q-type category names.

The 1920 trials for each S were divided equally into four experimental sessions, each containing four category pairs with a total of 12 yes-stimuli and 12 no-stimuli. (The Ns in Table 1 are for four sessions.) Before the first session, there were 50 practice trials.

Once all of the above conditions were met, the 24 trials (pairs) for each P-time were randomized. In one session, the computor presented 24 question-stimulus pairs at the first P-time (8 msec), and then presented the same 24 pairs in a different random sequence at the next higher P-time (16 msec) and so on, until 20 blocks of 24 trials had been displayed. Thus there were 480 trials per session. The sequence of events for each trial was: 1) the Q-type category name was displayed for 2 secs, and immediately afterwards; 2) the test stimulus was displayed for a constant amount of time (1 sec) although to the S the duration appeared to vary; and, 3) the mask appeared to come on immediately after the test stimulus disappeared, but in fact the mask overlapped the test stimulus in time. Thus the duration of the test stimulus was actually controlled by the interval between test stimulus onset and mask onset.

The two dependent variables were: 1) Reaction time (RT) which was defined as the interval between the test stimulus onset and depression of a button by S; and, 2) response accuracy. These variables were coded and stored in the computor after each trial and were retrieved on paper tapes by E after each session. These tapes were then decoded by another program and summarized print-outs were obtained, showing RTs in msecs and correct-incorrect accuracy measures for each trial.

Procedure. The S was brought into a darkened room and seated in front of the plotter unit. He was given the response box and asked to use either his preferred or non-preferred hand on the "yes" button. The instructions which were then read to him appear in Appendix B. The E then left the room and the program was executed without interruption unless S desired rest periods. All the data were collected and compiled automatically by the computor.

RESULTS AND DISCUSSION

Data. The data were retrieved after each session on paper tapes and then decoded onto Teletype print-outs which were organized in terms of question-stimulus pairs. For each pair, the RT was printed for each P-time in ascending order. The accuracy on each trial was denoted by either a "one," which was a correct response, or a "zero," which was an incorrect response.

Threshold RTs were then recorded for each question-stimulus pair by inspection. A threshold trial was defined as the second of a string of five correct trials in the ascending series of P-times, provided that there was not more than one subsequent, consecutive error and that the total number of errors following the string did not exceed three. Thus, a string of five correct trials did not qualify if: 1) at least two consecutive errors followed, or, 2) at least three errors in total followed.

A program was then written which collapsed all the correct response data without regard to thresholds. Mean RTs with their respective variances and Ns were obtained at each P-time, for each Q-type, on each of the four days. This analysis of variance (ANOVA) will be presented in the next part of this section.

Reaction times which exceeded 2 secs were omitted from all analyses. If a yes-word was disqualified in this manner, it was also dropped from the two other counter-balanced conditions.

The recognition task data were analyzed for yes-words only, since no-words did not appear equally often under each Q-type. It was assumed that since the yes-words did appear equally often under each Q-type, recognition errors made on the words should be equally distributed by chance among the categories and a Chi Square Test should reveal any significant deviation from this expectation.

Analyses. First, a 3-way ANOVA for means (see Edwards, 1968, pp.264-266), was performed on all the correct response data which is illustrated in Figure 5. The summary of this analysis appears in Table 2. The three variables analyzed were days x P-time x Q-type. In Figure 5, the ranked P-times from 1-20 are blocked in fours so that, for example, the first block in each day represents the ranked P-times from 1 to 4 or 8 msecs to 32 msecs. This analysis showed the main effect of days to be significant at the .01 level, $F\ (3,\infty) = 82.5$. Mean RTs dropped steadily from day 1 to day 4, which probably reflects Ss learning and becoming highly practiced at making new types of discriminations. The days variable interacted significantly with the P-time variable, $F\ (57,\infty) = 4.16$ and indicates that on earlier days the slopes of the curves are steeper than on later days. Again, this is probably due to Ss learning how to make appropriate discriminations at earlier P-times since for all of the days, the end-points of each curve lie between 600 and 700 msecs, while the points in the first blocks for each day fall between 550 msecs and 850 msecs. It is thus the strategy used in the first few P-times which changes over days. Days did not interact significantly with Q-type. This demonstrates that practice or learning did not differentially affect the Q-type variable. Therefore, any dif-

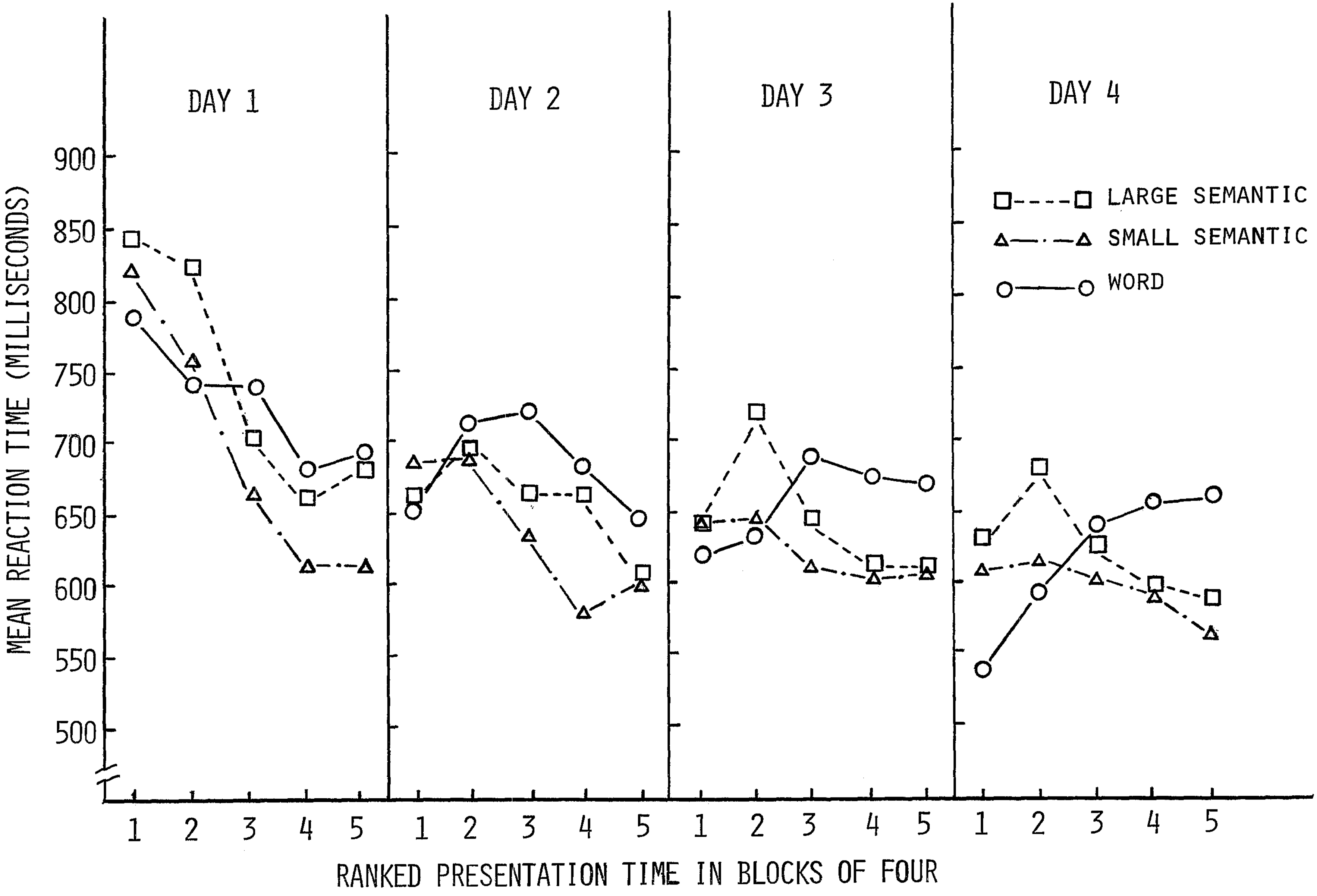

FIGURE 5 DAYS X PRESENTATION TIME X QUESTION TYPE INTERACTION FOR MAIN EXPERIMENT

TABLE 2

ANALYSIS OF VARIANCE SUMMARY TABLE FOR EXPERIMENT

DAYS x PRESENTATION TIME x QUESTION TYPE

Source of Variation	SS	d.f.	MS	*F*
Days	350,009	3	116,669	85.2*
Presentation Time	190,379	19	10,019.9	7.34*
Question Type	56,089	2	28,049	20.56*
Days x P-Time	323,227	57	5,670.6	4.16*
Days x Q-Type	11,229	6	1,871.5	
P-Time x Q-Type	182,618	38	4,805.7	3.5*
Days x P-Time x Q-Type	115,956	114	1,017.2	
Within Treatments (Error)		185	1,364.19	

* Significant beyond the .01 level

ferences among the three levels of the Q-type variable cannot be attributed to differential practice effects. Because of this and the fact that variances were about half on days 3 and 4 than what they were on days 1 and 2, it was decided to analyze data for days 3 and 4 for threshold measurements. The main effect of P-time was also significant, $\underline{F}$ (19,185) = 7.3 and demonstrates that RT is generally faster at longer presentation times than at shorter ones. This is only generally true since there was an interaction between P-time and Q-type, that is, especially on days 3 and 4, the word Q-type increased monotonically as a function of P-time while semantic Q-types first increased and then gradually decreased. The Q-type variable main effect was significant as well, $\underline{F}$ (2,185) = 20.56 and suggests that processing time is differentially affected by this manipulation. The most important verification of this ANOVA was the significant interaction between Q-type and P-time, $\underline{F}$ (38,185) = 3.5. This indicates that at shorter presentation times, the slopes of the three Q-types bear a different relationship to one another than at longer presentation times. The major contributing factor to this interaction was probably the cross-over of the word Q-type condition seen in Figure 5. This suggests that different strategies may be used by Ss when stimulus information is reduced than when they have unrestricted input. Further, it may be that Meyer's model only holds for the later condition and that a different process may operate under the former condition.

Second, an analysis was done to determine whether the three Q-types yielded significantly different threshold values. The mean rank order (from 1 to 20) of the second trial in each string of five correct responses, under each Q-type, was determined. The mean ranked presentation time for the large semantic condition was 5.14, the small semantic condition 5.09, and for the word condition, 4.19. A t-test revealed that the defined threshold in the word condition differed significantly from each of the other two conditions ($t\ (115) = 2.3$, $p < .05$). This finding supports Treisman's early filter model as laid out in hypothesis 7. If making correct wordness decisions requires structural analysis, it appears that this process can operate to initiate output, regardless of whether or not processing continues to the semantic level. Assuming that the mask stopped further processing at threshold, one cannot argue that an exhaustive dictionary process occurred since the wordness and semantic conditions were observed at points equated for accuracy, yet wordness decisions required less stimulus information than semantic decisions. If the mask did not prevent further processing after wordness decisions had been made, then semantic decisions should have been available with the same presentation time as wordness decisions. Since they were not it is reasonable to assume that structural analysis and output are possible prior to further semantic processing.

Third, hypotheses 1-5 were tested by comparing means calculated for threshold trials (T). To try and replicate Meyer's data, means were also calculated at the tenth ranked P-time (T + 10), following the threshold trial. For both of these analyses, data for days 3 and 4 were used since the variances during these days were about half of what they were on day 1, and there was no interaction between days and Q-type.

Tables 3a and 3b show the obtained means for words, pseudowords, yes- and no-responses at T and T + 10. Table 4 shows the predicted differences between means for each of the first five hypotheses and the actual outcomes at threshold and at the tenth P-time following the threshold trial. Since hypotheses 1, 3 and 4 each have two parts there are actually eight comparisons made at T and T + 10. From Table 4 it can be seen that 7 out of 8 of these hypotheses were supported at threshold. The eighth comparison was, however, in the predicted direction but failed to reach statistical significance. Table 4 also shows evidence that Meyer's data wereessentially replicated. One anomaly appears in these data for hypothesis 1: at T + 10 yes-responses to the word Q-type were significantly longer than to both semantic Q-types. This can be interpreted to mean that since yes-response to the word Q-type took longer than to semantic questions, "words" are treated as one large category. Landauer and Freedman (1968) also made this suggestion.

A Chi Square analysis was performed on the final recognition test data and appears in Table 5. A total of only 45 errors were made on the final recognition test and probably reflects learning due to the many repeated exposures of each item. Of these errors, 12 of the items had been paired with a LS question, 10 with a SS question, and 23 with the W question. Chance would predict an equal distribution of errors in each condition. That is, the expected error rate was 15 in each Q-type. A Chi Square test showed that significantly more errors were made for items presented in the word condition, $\underline{X}^2$ (2) = 6.53, $\underline{p} < .05$. This finding supports Craik & Lockhart's (1972) notion of levels of processing.

TABLE 3a

OBTAINED MEANS AND STANDARD DEVIATIONS, DAYS 3 & 4, FOR YES-RESPONSES TO WORDS AT T AND T+10

		Yes Responses	
		At T	At T+10
Large Semantic			
	Mean	678	619
	SD	315	212
Small Semantic	Mean	658	592
	SD	303	186
Word	Mean	627	676
	SD	378	225

TABLE 3b

OBTAINED MEANS AND STANDARD DEVIATIONS, DAYS 3 & 4, FOR NO-RESPONSES TO WORDS AND PSEUDOWORDS AT T AND T+10

		No Responses			
		Words		Pseudowords	
		At T	At T+10	At T	At T+10
Large Semantic					
	Mean	828	884	884	824
	SD	362	221	292	151
Small Semantic					
	Mean	808	775	888	757
	SD	338	130	330	289
Word					
	Mean			885	869
	SD			306	200

Table 4

Predicted and Actual Outcomes of Hypotheses 1-5 at Threshold (T) and by Meyer and Ellis (T+10) with T-test Values

	COMPARISON (a vs b)							
	HYPO 1		HYPO 2	HYPO 3		HYPO 4		HYPO 5
a= b=	yes:Wd yes:LS	yes:Wd yes:SS	No:LS(wds) No:SS(wds)	No:LS(wds) No:LS(pseud)	No:SS(wds) No:SS(pseud)	No:Wd(pseud) No:LS(pseud)	No:Wd(pseud) No:SS(pseud)	No:LS(wds) No:Wd(pseud)
PREDICTION AT T	a < b	a < b	a = b	a < b	a < b	a = b	a = b	a < b
OUTCOME	AP	NAP	AP	AP	AP	AP	AP	AP
t-VALUE	2.51	1.6	1.59	3.76	3.74	.97	1.07	2.25
P <	.02	--	--	.001	.001	--	--	.02
MEYER'S PREDICTION (T+10)	a = b	a = b	a > b	a > b	a > b	a < b	a < b	a = b
OUTCOME	NAP	NAP	AP	AP	NAP	AP	AP	AP
t-VALUE	3.16	4.3	2.17	3.16	1.05	2.6	4.86	1.6
P <	.001	.001	.05	.001	--	.01	.001	--

Key.- yes: = yes-word stimulus
no: = no-word stimulus
Wd: = word Q-type
LS = large semantic Q-type
SS = small semantic Q-type

(wds) = words used as test stimuli
(pseud)= pseudowords used as test stimuli
AP = outcome of experiment was as predicted in hypothesis
NAP = outcome of experiment was not as predicted in hypothesis

TABLE 5

DATA FOR RECOGNITION ERRORS FOR EACH QUESTION TYPE
IN MAIN EXPERIMENT

	Question Type		
	Large Semantic	Small Semantic	Word
Number of errors observed	12	10	23
Number of errors expected	15	15	15

It would be appropriate at this point to discuss Craik's idea of levels or "depth" of processing and show how much of these data can be explained in this frame of reference. As was mentioned earlier with respect to the verbal learning literature, Craik prefers to envisage memory on a continuum of analysis of information rather than on the more traditional "box" models. According to Craik and Lockhart (1972) greater depth "implies a greater degree of semantic or cognitive analysis. After the stimulus has been recognized, it may undergo further processing by enrichment or elaboration." Other researchers such as Treisman (1967) have previously alluded to this conceptionalization. Treisman's early filter model clearly suggests that early levels of processing are concerned with the analysis of physical or sensory features while later stages are more concerned with pattern recognition and the extraction of meaning.

In this experiment it was found that the minimum presentation duration for answering "wordness" questions (which does not require semantic processing) was less than the minimum duration for answering semantic questions. Assuming that the mask governed the level to which processing advanced, it appears that wordness information can be extracted at a lower level of processing than semantic information and that response output can be initiated without necessitating exhaustive processing. This point is supported by the final recognition test data since the items paired with the word questions were not remembered as well as items paired with semantic questions. Thus items in the word condition were not as well registered since they only advanced to a preliminary stage of processing before they were interrupted or made "unprocessable" by the mask. In Meyer and Ellis' (1970) experiment, processing was uninterrupted by a mask and thus both stages of a dictionary process may have occurred, resulting in their failure to find latency differences between the wordness and semantic conditions. Under conditions where the level to which processing could advance was restricted, a Treisman-like process does seem plausible.

At threshold, response latency to the wordness question was shorter than to the semantic question, which suggests that semantic information may be processed at a deeper level than wordness information and a parallel race model based on latency data, as Meyer concluded, is inappropriate. Moreover, negative response data for hypothesis 2 suggest that, assuming random access to the various dictionary units, semantic information derives from activation or firing of dictionary units and not from searching semantic categories, since no-responses to the large semantic Q-type were just as fast as no-responses to the small semantic Q-type. Under this model, hypotheses 3, 4 and 5 which employ pseudowords, yielded evidence which suggest that all the dictionary units must be exhaustively searched when a negative response is required (i.e., the unit is not present).

Data observed at T + 10 were probably subjected to complete processing and are therefore comparable to Meyer's data. Under these conditions, his findings were basically replicated with the exception that it took longer to make a positive decision about an item's wordness than about its meaning. This renders his race model untenable and supports Landauer and Freedman"s (1968) notion that the dictionary units are treated as one large category. The categorical search model can account for these data when processing is allowed to advance since no-responses to the large semantic Q-type were longer than to the small semantic Q-type.

In conclusion, this experiment has demonstrated that under conditions of limited stimulus condition where further processing is precluded, a dictionary process may be involved in word processing. Furthermore, semantic information is probably derived from the test item under these conditions. On the other hand, a categorical search model seems more appropriate for longer processing durations. The implication drawn from this experiment is that more than one mode of operation may be available to the system and that particular experimental conditions may artifactually impose what would appear to be variations in structure. The polemic of dictionary vs category model may not be theory specific but rather, paradigm specific. Further research should be directed toward discovering the range and diversity of the processes available to the system.

REFERENCES

Averbach, E. & Coriell, A.S. Short-term memory in vision. Bell System Technical Journal, 1961, 40, 309-328.

Baddeley, A.D. Semantic and acoustic similarity in short-term memory. Nature, 1964, 204, 1116-1117.

Baddeley, A.D. & Dale, H.C.A. The effect of semantic similarity on retroactive interference in long- and short-term memory. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 419-420.

Candland, D.C. Psychology: The Experimental Approach. New York: McGraw-Hill, 1968.

Cherry, E.C. Some experiments on the recognition of speech with one and two ears. Journal of the Acoustic Society of America, 1953, 25, 975-979.

Collins, A.M. & Quillian, M.R. Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 240-247.

Collins, A.M. & Quillian, M.R. Does category size affect categorization time? Journal of Verbal Learning and Verbal Behavior, 1970, 9, 432-438.

Conrad, R. An association between memory errors and errors due to acoustic masking of speech. Nature, 1962, 196, 1314-1315.

Conrad, R. Acoustic confusions in immediate memory. British Journal of Psychology, 1964, 55, 75-84.

Conrad, R. Interference or decay over short retention intervals? Journal of Verbal Learning and Verbal Behavior, 1967, 6, 49-54.

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

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

Deutsch, J.A. & Deutsch, D. Attention: Some theoretical considerations. Psychological Review, 1963, 70, 80-90.

Deutsch, J.A., Deutsch, D. & Lindsay, P.H. Comments on "Selective attention: Perception or response?" Quarterly Journal of Experimental Psychology, 1967, 19, 362-364.

Edwards, A.L. Experimental Design in Psychological Research. (Third Edition). Holt, Rinehart and Winston, Inc., 1968.

Eriksen, C.W. & Lappin, J.S. Internal perceptual system noise and redundancy in simultaneous inputs in form perception. Psychonomic Science, 1965, 2, 351-352.

Eriksen, C.W. & Lappin, J.S. Independence in the perception of simultaneously presented forms at brief presentations. Journal of Experimental Psychology, 1966, 72, 20-25.

Estes, W.K. & Taylor, H.A. Visual detection in relation to display size and redundancy of critical elements. Perception and Psychophysics, 1966, 1, 9-16.

Glanzer, M. & Cunitz, A.R. Two storage mechanisms in free recall. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 351-360.

Hebb, D.O. The Organization of Behavior. New York: Wiley, 1949.

James, W. Attention. In Principles of Psychology. New York: Holt, 1890, 402-458.

Kahneman, D. Method, findings, and theory in studies of visual masking. Psychological Bulletin, 1968, 70, 404-425.

Katz, J.J. & Fodor, J.A. The structure of a semantic theory. Language, 1963, 39, 170-210.

Keppel, S. & Underwood, B.J. Proactive inhibition in short-term retention of single items. Journal of Verbal Learning and Verbal Behavior, 1962, 1, 153-161.

Kintsch, W. & Buschke, H. Homophones and synonyms in short-term memory. Journal of Experimental Psychology, 1969, 80, 403-407.

Landauer, T.K. & Freedman, J.L. Information retrieval from long-term memory: Category size and recognition time. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 291-295.

Levy, B.A. & Murdock, B.B., Jr. The effects of delayed auditory feedback and intralist similarity in short-term memory. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 887-894.

Liss, P. Does backward masking by visual noise stop stimulus processing? Perception and Psychophysics, 1968, 4, 328-330.

Mackworth, J.F. The duration of the visual image. Canadian Journal of Psychology, 1963, 17, 62-81.

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

Meyer, D.E. Dual memory-search of related and unrelated semantic categories. Paper presented at the 42nd Annual Meeting of the Eastern Psychological Association, New York, N.Y., April 15-17, 1971.

Meyer, D.E. & Ellis, G.B. Parallel processes in word-recognition. Paper presented at the 10th Annual Meeting of the Psychonomic Society, San Antonio, Texas, November 5-7, 1970.

Miller, G.A. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Milner, B. The memory defect in bilateral hippocampal lesions. Psychiatric Research Reports, 1959, 11, 43-52.

Moray, N. Attention in dichotic listening: affective cues and the influence of instructions. Quarterly Journal of Experimental Psychology, 1959, 11, 56-60.

Moray, N. Attention. Selective Processes in Vision and Hearing. Hutchinson Educational Ltd., 1969.

Mowbray, G.H. Perception and retention of verbal information presented during auditory shadowing. Journal of the Acoustic Society of America, 1964, 36, 1459-1464.

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

Norman, D.A. Memory and Attention. New York: Wiley, 1969.

Paivio, A. Imagery and Verbal Processes. Holt, Rinehart and Winston, Inc., 1971.

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

Quillian, M.R. Word concepts: a theory and simulation of some basic semantic capabilities. Behavioral Science, 1967, 12, 410-430.

Reichter, G.M. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Rips, L.J., Shoben, E.S. & Smith, E.E. Semantic distance and the verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 1-20.

Sperling, G. The information available in brief visual presentations. Psychological Monographs, 1960, 74, No.11.

Sperling, G. A model for visual memory tasks. Human Factors, 1963, 5, 19-31.

Sperling, G. Successive approximations to a model for short-term memory. Acta Psychologica, 1967, 27, 285-292.

Swets & Kristofferson, A.B. Attention. Annual Review of Psychology, 1970, 21, 339-366.

Thorndike, E.L & Lorge, I. The Teacher's Word Book of 30,000 Words. New York: Bureau of Publications, Teachers College, 1944.

Treisman, A.M. Contextual cues in selective listening. Quarterly Journal of Experimental Psychology, 1960, 12, 242-248.

Treisman, A.M. Reply to "Comments on 'Selective attention': Perception or response?", Quarterly Journal of Experimental Psychology, 1967, 19, 361-367.

Tulving, E. & Madigan, S.A. Memory and verbal learning. Annual Review of Psychology, 1970, 21, 437-484.

Tulving, E. & Patterson, R.D. Functional units and retrieval processes in free recall. Journal of Experimental Psychology, 1968, 77, 239-248.

Waugh, N. & Norman, D.A. Primary Memory. Psychological Review, 1965, 72, 89-104.

Wheeler, D.D. Processes in word recognition. Cognitive Psychology, 1970, 1, 59-85.

Wickelgren, W.A. Acoustic similarity and intrusion errors in short-term memory. Journal of Experimental Psychology, 1966, 71, 396-404.

APPENDIX A

STIMULUS ITEMS FOR MAIN EXPERIMENT

A. "Yes" Words

Superordinate	Subordinate	Exemplars (3)
ANIMAL	MAMMAL	HORSE DOG RABBIT
MACHINE	VEHICLE	CAR TRAIN JET
APPAREL	FOOTWEAR	SHOE SOCK BOOT
STRUCTURE	BUILDING	HOUSE HUT STORE
WEAPON	FIREARM	GUN PISTOL RIFLE
FOOD	MEAT	BACON ROAST STEAK
JOB	PROFESSION	DOCTOR LAWYER TEACHER
PLANT	TREE	ELM OAK SPRUCE
SMOKED	CIGARETTE	PLAYERS CAMEO EXPORT
UTENSIL	CUTLERY	FORK SPOON KNIFE
AUTHOR	POET	SHELLEY KEATS BURNS
MONEY	COIN	DIME NICKEL PENNY
SUBSTANCE	ROCK	GRANITE COAL MARBLE
LIQUID	DRINKABLE	WATER MILK BEER
PLACE	PLANET	MARS SATURN PLUTO
PERSON	SOLDIER	MAJOR COLONEL PRIVATE

B. "No" Stimuli

Words	Pseudowords	Pseudowords
ANGER	NITLE	KNOP
BLOCK	SROWN	CIOL
CAPITAL	RASULT	RAUND
CURTAIN	ASGUE	GRUND
DOOR	COURPE	MAFER
FOLDER	VOPAL	WEITH
GUILT	ANZUAL	WIKING
HAND	STUDK	HUBE
LETTER	SUG	BOTTXE
LIGHT	LYMP	TILLOW
MOULD	CEOLING	SNY
NAME	FACILY	TABE
PAINT	WAEL	QUIER
SOUND	SOPG	OCE
TARGET	CHADE	HORER
TIMER	LEBER	THBEE

APPENDIX B

INSTRUCTIONS USED IN MAIN EXPERIMENT

This is an experiment to determine how quickly and accurately you can answer certain types of questions. On each trial there will first appear a single word on the screen before you. This word is asking a question about a stimulus which will immediately follow it. For example, if the word COUNTRY appears, it is asking whether or not the next stimulus belongs to that class of nouns. If the word that you then see is CANADA, push the yes-button with your index finger; if TULIP appears, you would push the no-button because a tulip is not a country. As well as questions about classifying nouns, you may also be asked the question WORD? In this case a word or a pseudoword like BILK may appear. Pseudowords may also appear after the first type of question. For example, if BILK appears after a question like COUNTRY, the obvious answer is "no." The question stimuli will remain on the screen long enough for you to read them without any difficulty. However, some of the other stimuli, especially early in the sessions, may last an insufficient amount of time for you to be sure of your answers, or even see the stimuli for that matter. Nevertheless, you must give a response on every trial, so please guess if you are not sure. Remember, we want you to respond as fast as possible, but also to be as accurate as you can. The question stimuli will serve as a warning for the test stimuli which follow and the trials will follow one another continuously. Work as steadily as possible and whenever you feel that you want to rest, hold a button down after any trial and until you release it, the program will remain stopped. If you notice anything else coming into your field of vision, just ignore it and concentrate on your task. Are there any questions? Alright, here are some practice trials . . .