THE UNIVERSITY OF MANITOBA

DEPTH OF PROCESSING: IS THERE A SEMANTIC CONTINUUM?

by

Donald Wilson McMurray

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

WINNIPEG, MANITOBA

December, 1978

DEPTH OF PROCESSING: IS THERE A SEMANTIC CONTINUUM?

BY

DONALD WILSON McMURRAY

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

The levels of processing approach to memory proposed by Craik and Lockhart (1972) has recently come under attach because of the circularity that has been inherent in relating depth to the strength of the memory trace (Nelson, 1977). Seamon and Virostek (1978), however, derived separate theoretical and empirical orderings of a range of classification tasks from across the various processing domains. A significant rank correlation between the ordered scale and stimulus word free recall strengthened the claim of earlier studies (Craik & Tulving, 1975; Hyde & Jenkins, 1973) that memory is a by-product of perceptual/cognitive analyses performed on the stimulus at input. This research sought to extend Seamon and Virostek's findings by operationalizing the term 'depth' and testing its predictive capabilities within a single domain in a series of incidental learning studies.

The framework for the six experiments reported here was the recent revision of the depth of processing idea advanced by Lockhart, Craik and Jacoby (1976), i.e., the notion of hierarchically organized domains of processing. Specifically, it was hypothesized that within the semantic domain more elaborate semantic analyses, defined in terms of the qualitative nature of the mental operations required to perform the task, would yield superior recall.

Experiment I involved subjects rating a variety of semantic or meaning producing activities against criteria that were postulated as underlying depth of processing. Subjects assessed each task in terms of the amount of conscious effort or attention required to carry out the task. A depth-ordered scale was derived from paired-comparison judgments collected on nine semantic tasks. The tasks ranged from judging whether a word was living or non-living to providing a short definition

to determining how two unrelated words were similar.

In two follow up experiments (Experiment II & III) independent groups of subjects were unexpectedly tested for recall of lists of 24 unrelated words after engaging in one of the scale-ordered tasks that required them to generate their own encodings. Predictions based on the scale were empirically validated for most tasks when stimuli were low (Experiment II) or medium (Experiment III) in meaningfulness.

Additional experiments explored the generalizability of the findings of the first two studies by equating for study time in a within subject design (Experiment IV) and introducing a self-paced procedure in Experiment V. Experimental demands were also varied so that subjects rated words in all cases. Under these circumstances the relationship between the depth-ordered scale and the retention functions associated with each task broke down.

Experiment VI employed the original experimental procedures (i.e., subject generated encodings) but retained highly meaningful stimuli as in Experiments IV and V. Although order of recall correlated well with the scale the effect was not as strong as in Experiments II and III and differences across conditions were statistically indistinguishable.

It was concluded that the effectiveness of depth of processing within the semantic domain is a function of level of meaningfulness of the material under study vis à vis the stored contents of existing cognitive structures, the qualitative nature of the encoding operations and the extent to which the learner is actively engaged in those operations. Results were interpreted in the light of Craik and Jacoby's (1975) two process model of short-term retention.

ACKNOWLEDGEMENTS

I extend a deep feeling of gratitude to John McIntyre who supervised the development of this research project in his capacity as chairman of the dissertation committee. In addition to offering advice and continued encouragement at every stage, John gave freely of that which I have come to know and value over the years, his self. This provided me with the necessary resolve and direction to complete this enterprise.

For their valuable criticisms and suggestions I acknowledge a debt to Dr. Murray Singer and to Dr. John Whiteley, both of whom contributed substantially in setting the study in perspective in terms of its overall scope and objectives. My thanks also go to Dr. Joanne Rogan for her comments relating to the analysis and presentation of the data.

I am extremely grateful to Joanne Stothers for the innumerable hours she devoted to typing, proofreading and correcting the various drafts of the manuscript.

Finally, my deepest thanks are for Anne who stood at my side, and whose constant emotional support and solace were a source of strength and inspiration to me through all those times when the way ahead seemed unclear.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

DEPTHS OF PROCESSING: IS THERE A SEMANTIC CONTINUUM?

The levels of processing view propounded by Craik and Lockhart (1972) has exerted a considerable influence on the direction of recent research in human memory. Although originally intended as an alternative conceptual framework to multistore models of memory it has assumed the status of a theory especially as it relates to control processes in short-term retention. In the original paper, memory for an event was conceived of as a by-product of perceptual analyses with higher levels of recall reflecting deeper levels of processing at input. Preliminary shallow analyses were thought to be concerned with physical or structural features of the stimulus. Progressively deeper analyses were more concerned with extracting meaning attributes and/or the developing of associative relationships with the stored contents of existing cognitive structures. The depth to which a stimulus was analyzed was, therefore, thought to be a function of the type of analysis or level of encoding employed as well as the attention required of the central processor for the analysis. This too, varied according to the familiarity or level of meaningfulness of the to-be-remembered (TBR) item. It was initially thought that highly meaningful material was processed much faster than low meaningful material.

The memory trace was the end-product with deeper initial encodings leading to richer, more durable traces. This notion supplanted the need to postulate separate memory stores along with a variety of mechanisms for transferring material from one store to another. Memory was simply characterized as the by-product of various information processing activities or stimulus analyzing operations.

The depth of processing approach, although partially formulated on the basis of a review of some of the earlier studies of Jenkins and his associates (Hyde & Jenkins, 1969; Johnston & Jenkins, 1971), generated an enormous amount of research following the Craik and Lockhart paper. Most of the findings provide experimental support for the notion that deeper encodings result in better retention for both free recall (Craik & Tulving, 1975; Epstein, Phillips & Johnson, 1975; Hyde & Jenkins, 1973) and recognition (Craik & Tulving, 1975; Elias & Perfetti, 1973; Seamon & Murray, 1976). Most of these studies involved the use of an incidental learning paradigm where an attempt was made to gain control over the encoding strategies adopted by the subject. Usually the subject was required to perform an orienting task by judging selected characteristics of the stimulus materials. It is assumed that engaging a subject in an orienting task provides a fairly strong means of controlling that subject's encoding operations (Postman, 1976). Typically, in a levels of processing experiment, the material is held constant and the task demands are varied for each of several groups of subjects. For example, Goldman and Pellegrino (1977) presented three different orienting tasks to subjects. If the stimulus to be processed was COPPER then the three tasks thought to represent progressively deeper levels of analysis were as follows: (a) "Does the word contain the letter p?" (b) "Does the word rhyme with 'hopper'?" and (c) Gold is a metal. Is ________?" Results indicated that the deeper processing suggested by task (c) produced better memory than did the more superficial, less meaningful, processing entailed in tasks (a) and (b).

This new look in memory research represents a partial shift away from traditional variables (e.g., stimulus meaningfulness, study time and associative strength between items) to a focus on the qualitative nature of the mental operations at input. Deeper levels of processing presumably yield richer, more durable memory traces. The approach has not been without its detractors, however, and as we shall see there are still problems, especially as they relate to determining what criteria can be found that will predict, a priori, that a particular level of encoding will induce a better memory trace than another.

Recently minor modifications have been made to the concept of depth of processing. Factors such as stimulus elaboration (Craik & Tulving, 1975), distinctiveness of encoding (Moscovitch & Craik, 1976; Lockhart, Craik & Jacoby, 1976), precision of encoding (Stein, 1978) attempt to focus more closely on the mechanisms mediated by the various orienting tasks operating at input.

The Craik and Tulving (1975) series of experiments, as Postman (1976) points out, represents the most detailed attempt to date to systematically explore the relationship between levels of processing of TBR words and their subsequent retention. These authors concluded that the notion of "depth" was too simplistic as an explanatory concept to account for the complete range of perceptual/cognitive analyses that might be performed on a stimulus during encoding. Instead they opted for the concept of "domains" of processing (Lockhart, Craik & Jacoby, 1976), a term suggested by Sutherland (1972).

Additonally, they speculated that even within a domain depth ought to be replaced by "elaboration" or "spread" of encoding since these terms were more descriptive of the nature of the analyses carried out at input. Nevertheless, the stress still remained very much on the qualitative nature of the encoding operations. In fact, Craik and Tulving (1975) adduced strong evidence to show that a fairly rapid and "minimal semantic analysis is more beneficial than an elaborate structural analysis (Experiment 5)" (p.289).

In adopting the idea of domains, Lockhart, Craik and Jacoby (1976) suggested that various characteristics of words (e.g., orthographic, phonemic or semantic) can be best thought of as existing at separate levels. Hence the concept of depth was retained in the sense that the domains can be hierarchically arranged. The important question became one of whether stimulus processing necessarily proceeds through each of these domains until the level of analysis demanded by the encoding task is reached. The authors themselves are somewhat vague on this point. At one stage (p.78) it is claimed that with practice a complete domain can be circumvented as when an accomplished reader bypasses the phonemic stage completely (something the present author agrees with entirely), yet later in the same paper it is stated (p.80) that input from the phonemic domain is a necessary prerequisite for word perception. This is difficult to reconcile with evidence (Cattell, 1885) that subjects are capable of perceiving whole words faster than they can a single letter. Obviously word perception does not depend totally on either letter-by-letter identification or some grapheme-phoneme recoding mechanism.

More recently Graesser and Mandler (1975) conducted a study that also indicated that domains may be bypassed with the learner proceeding straight to the target domain. These authors showed that subjects do not retain memory for the surface structure of sentences when forced to concentrate on semantic analyses. Mistler-Lachman (1975) also demonstrated that subjects can resolve deep structure or semantic ambiguity before detecting and resolving ambiguities at the surface structure level of sentences. Subsequently, one of the co-authors in the paper under review conceded that deeper semantic analyses can occur without any prior structural analyses (Moscovitch & Craik, 1976).

In conclusion, the status of depth as a concept descriptive of the types of analyses that exist either within or across domains remains unclear. The present study explores the plausibility of utilizing depth as an explanatory concept within a domain (e.g., the semantic domain) and to assess whether qualitatively deeper stimulus analyses will yield a richer memory trace.

Certainly there is some experimental support for Craik and Tulving's (1975) suggestion that within a particular domain depth may be better characterized by the "spread" or number of encoded features rather than some ordered levels of analysis. Within the semantic domain, for example, Klein and Saltz (1976) found that subjects encoding along two uncorrelated dimensions evidenced better recall than subjects who elaborated the stimuli in terms of a single encoding dimension. In attempting to examine what kinds of existing cognitive structures might be involved in semantic deep-level processing these authors initially estimated correlations between the dimensions "happy-sad," "big-little," "fast-slow," and "pleasant-unpleasant," based on subjects' ratings in a pilot study. Subsequently, under incidental learning conditions, subjects evidenced better recall after rating words on two dimensions than on a single dimension, especially if the dimensions were uncorrelated, e.g., "pleasant-unpleasant" and "big-little" (where the correlation between the ratings was .05).

However, the phrase "greater depth" implies qualitatively deeper levels of cognitive analysis, especially within the semantic domain. Moreover, it is intuitively appealing to think in terms of a semantic continuum since the concept of semantic processing is so diverse and hitherto generally not well understood. Added to this is the idea that it is highly improbable that all semantic processing is equivalent. Given a gradation of memory performance contingent on type of depth of processing, the continuum might eventually be used to examine more closely the psychological nature of semantic processing itself (Baddeley, 1978).

There is some empirical support for a depth-ordered semantic continuum. For example, Epstein, Phillips and Johnson (1975) demonstrated that asking subjects to find similarities between pairs of words with different meanings or differences between words that were similar in meaning led to better recall than indicating similarities for similar words or differences for different word pairs. The implication is that encoding deeper semantic features produced a more elaborate memory trace. These results are, of course, entirely compatible with Lockhart, Craik and Jacoby's (1976) contention that "a reasonably familiar pattern or stimulus - response sequence [i.e., judging similarities for similar words] will be analyzed and encoded by a moderate number of analytic operations and will result in a moderately rich memory trace . . . Conversely, if the stimulus is novel, or difficult to process. . . more analyses are carried out and a richer memory trace results" (p. 79). Notice too, that the different semantic tasks in the Epstein, Phillips and Johnson study could be characterized as requiring differing amounts of conscious effort or attention in order that the task be successfully carried out. This is important because memory then becomes not only a function of the number of encoded features but it also is dependent on the qualitative nature of the cognitive analyses, i.e., deeper analyses require more conscious attention and hence produce richer and longer lasting memory traces.

Defining Depth of Processing

There has been a good deal of criticism aimed at the levels position in terms of the circularity that exists with respect to an adequate definition of depth of processing (Baddeley, 1978; Goldman & Pellegrino, 1977; Nelson, 1977). Generally, depth has been defined in terms of degree of stimulus elaboration and is reflected in increased levels of recall or recognition. Similarly, it is argued that well-remembered stimuli must have been more deeply processed. Processing tasks are chosen because researchers intuitively feel task A requires more elaborate processing than task B. Most of the processing activities fall into semantic and non-semantic categories, but as some authors suggest (Postman, 1975; Postman & Kruesi, 1977) no independent criteria for distinguishing semantic from non-semantic tasks have been developed. Consequently, until recently, the only validation for any "levels" notion has been the memorial consequences of the various orienting tasks themselves.

In an attempt to induce depth of processing and to break the circularity surrounding the concept, Craik and Tulving (1975, Experiment 2) explored the possibility that response latency might be associated with depth of encoding. Subjects were asked one of three types of questions: 1) "Is the word in capital letters?" 2) "Does the word rhyme with PALACE?" 3) "Does the word fit the sentence 'He lived with his mother and father in a ____________'?" Results confirmed that as decision latency increased for Tasks 1), 2), and 3) (thought to reflect ever deeper levels of processing) so did recognition performance. It could be argued that these findings are predictable on the basis of total study time. In re-analysis of the data from Experiment 2 and a subsequent study (Experiment 5) the same authors found a negative correlation between study time and recognition performance and hence concluded that retention is not solely a function of study time. (For a fuller discussion of the Craik and Tulving, 1975, series of experiments see Appendix A.) That processing time is not a major determinant of memory performance has also been suggested by other studies (Carpenter, 1974; Gardiner, 1974; Goldman & Pellegrino, 1977; Seamon & Murray, 1976).

In response to Nelson's (1977) argument that depth as a construct was scientifically meaningless due to the lack of suitable definition Seamon and Virostek (1978) derived an ordering of classification tasks based on subject-defined depth of processing. Subjects were instructed to order thirteen tasks (suggested by Nelson as being representative of the various processing domains) in terms of the depth of processing required by each task. Depth was defined as the degree of difficulty associated with each task. The subject-defined processing depth was tested for its predictive capabilities in an incidental learning paradigm that employed the same questions with different subjects. A significant correlation between the median ranks of the classification questions and their associated free-recall scores strengthened the claim that memory performance is a positive function of increasing depth of stimulus processing. Additionally the requirement that depth of processing be defined independently of memory performance was met at least for a range of tasks drawn from across the orthographic, phonemic and semantic domains. It remains true that no one has produced an independent depth-ordering of processing tasks within a single domain.

Levels within Levels

Baddeley (1978) lamented the fact that there has been no real attempt to delineate levels within any one domain be it structural, phonemic or semantic. Within the semantic domain, for example, there is a need to further differentiate levels of analysis and to assess the relative effects of various kinds of semantic processing tasks on memory (Craik & Tulving, 1975; Klein & Saltz, 1976; Schulman, 1974; Seamon & Murray, 1976). As stated earlier, it is the intention of the present study to explore the proposition that within a single domain different encoding processes will produce different memory traces.

Despite Baddeley's (1978) claim that he is unaware of any evidence suggesting that deeper levels of processing within the semantic domain lead to better retention both the Epstein, Phillips and Johnson (1975) and the Klein and Saltz (1976) studies described above tend to support the notion. In the latter case the authors anticipated a richer memory for the two-dimension rating group because it was reasoned that more semantic attributes of the stimulus were actively encoded. This additive notion of encoding of stimulus attributes has received experimental support elsewhere, at least for free recall (Goldman & Pellegrino, 1977). Others too, have found that variable encoding of the same items tends to facilitate retention (Martin, 1968; Nelson & Hill, 1974). Within the phonemic domain, however, Nelson (1977) found that subjects required to make two different decisions for each of two repetitions retained these words no better than subjects making the same decision on each of two repetitions. Also, Hyde (1973) found that subjects who processed words on two semantic dimensions showed no better retention than either of the two groups who processed TBR words on a single semantic dimension only.

Despite the kinds of objections registered by Nelson (1977) and Nelson and Vining (1978) depth of processing continues to enjoy a great deal of attention and the amount of research the concept is generating, shows no sign of abating. While it is generally conceded that the reliability of depth of processing as a phenomenon, especially across domains, has been adequately demonstrated it is equally certain that the validity of the concept has not been fully documented. Because of this, the concept lacks both explanatory and predictive powers. For example, how do we determine, a priori, whether generating a free associate in response to a stimulus or examining it to see if it fits meaningfully into a given sentence frame will lead to better memory for that stimulus? If depth of processing within the semantic domain is to have any construct validity then two things must occur:

1. The term "depth of processing" must be operationalized by clearly specifying the operations used to measure it. If depth implies ordinality then there must exist an independently devised depth-ordering of processing tasks within the semantic domain.

2. The resulting taxonomy of processing tasks must be tested empirically to see if the rankings can predict previously obtained retention differences.

It may be said that apart from establishing some predictive capability for the ordering of depth of processing tasks, this type of exercise will reveal little of the nature of depth of processing as a psychological concept. Yet it remains true that the theoretical principle that retention increases as function of depth of stimulus processing has not been fully explored. It is the contention here that deeper semantic encodings (to be independently and operationally defined) will lead to richer, more persistent memory traces.

Therefore, as a preliminary exercise, the present study began by scaling a variety of semantic or meaning-producing activities by having subjects rate all tasks in a series of paired-comparisons using criteria thought to encompass the concept of depth of processing. Following this the scale was empirically validated in a series of incidental learning experiments. If tasks rated as requiring deeper semantic analyses in fact lead to stronger memory traces then an appealing notion would be confirmed.

EXPERIMENT I

The specific purpose of the initial experiment was to establish an ordinal scale for nine tasks requiring processing English words and which have been designated as involving semantic operations by 'levels' researchers in recent years (Bellezza, Cheeseman & Reddy, 1977; Craik & Jacoby, 1975; Craik & Tulving, 1975; Hyde & Jenkins, 1969, 1973; Mistler-Lachman, 1975; Nelson, Wheeler, Borden & Brooks, 1974; Shulman, 1971).

The nine tasks selected were not exhaustive of the kinds of semantic tasks employed but were chosen so as to be representative of the range. Intuitively they were thought to vary with respect to the complexity of the cognitive operations involved.

Method

Subjects. Twelve undergraduate students from a second year experimental psychology course were recruited to perform the rather demanding rating task which involved Fechner's (1860) method of paired comparisons (see Torgerson, 1958). It was reasoned that as psychology majors the subjects would be more committed to the task demands of the experiment than any random sample of psychology undergraduates. Subjects received partial course credit for their participation.

Rating tasks. The nine tasks that were selected for rating were as follows:

1. Write a few words saying how the two words are different. (DIFF)
2. Living-Nonliving. Subjects were asked to circle each word as 'Living' if the word described something that was in their opinion living, otherwise they were to circle the word 'Nonliving.' Similar

judgments were required for tasks 3 and 7. (L-NL)

3. Strong-Weak (S-W) - see Task 2.

4. Write a single meaningful sentence using both words. (2-WS)

5. Write two free associates for each word, i.e., which are the first two words that come to mind when you think of the word. (2-FA).

6. Write a meaningful sentence incorporating each word. (1-WS)

7. Pleasant-Unpleasant (P-UP) - see Task 2.

8. Write a few words saying how the two words are similar (SIM)

9. Write an appropriate definition for each word. (DEF)

Procedure. Each subject was presented with a four-page booklet containing the 36 paired comparisons he/she was to make. Comparisons were made with respect to a specific pair of English nouns which was different for each subject. The 12 word pairs were chosen from Paivio, Yuille and Madigan's (1968) norms. The nature of these word pairs is more fully described in Experiment II. Each subject was randomly assigned a word pair and the order of pages containing the ratings was randomized across subjects with the additional requirement that half of the subjects would rate from the bottom to the top of each page and the other half from the top to the bottom. The ordering of tasks on successive pages was randomly determined with the proviso that no task was assigned the same serial position on more than one page.

As part of the general instructions the subjects were given practice with each task by collectively working through the list of nine tasks, keeping the pair 'pliers-crab' in mind. The Experimenter demonstrated the procedure associated with each task. Following this the subjects were told that they would be required to compare each task to every other task in succession and to decide for each particular pair which task involved more conscious effort or attention.

At this juncture the Experimenter elaborated on the nature of depth of processing and its relationship to terms such as "depth" or "spread of encoding" as portrayed in the literature (Craik & Jacoby, 1975; Craik & Tulving, 1975; Lockhart, Craik & Jacoby, 1976). Illustrations using tasks from the non-semantic domains were used to sharpen the concept in the minds of subjects prior to the rating task. Care was taken to emphasize that depth was not necessarily to be equated with time but may likely have more to do with the complexity and qualitative nature of the encoding operation or the amount of conscious effort or attention required to carry out the task. Examples, again drawn from the non-semantic domains, were used to bolster this point. Subjects were then given a word pair to practice on for themselves and encouraged to think about the requirements of each task. The order in which tasks were practised was randomly determined for each subject. Following practice with each type of task (see Appendix B for Practice Sheet) subjects were given a single word pair (which was subsequently to be used in a later experiment) and proceeded to indicate which task of all possible pairs required more conscious effort or attention (see Appendix C for an example of the paired-comparisons task). Although encouraged to be consistent in their ratings the subjects were expressly instructed not to refer back to earlier ratings but rather to re-focus on the criteria for rating if uncertain about a particular pair (see Appendix G).

Results and Discussion

Throughout this paper an effect is considered significant using a significance level $\alpha = .05$. The response patterns for the paired comparisons between the nine semantic tasks are contained in Appendix D. Where all raters were consistent with respect to their choices the comparisons were readily reducible to rankings. Two raters (1 and 4) displayed some inconsistencies in their choices but Kendall's coefficient of consistency was high for both (.96). The overall ranks along with the word pairs upon which the ratings were based are shown in Table 1.

It was important to establish the extent to which the raters were in agreement with their ordering of the tasks. Accordingly, Kendall's coefficient of concordance (W) was calculated to be .74 and the chi-square of 71.13 (with 8 degrees of freedom) associated with this was highly significant (see Appendix E). There was thus a high inter-rater consistency in the ratings.

Since W is a measure of agreement among the raters one is tempted to conclude that the rankings are correct and were ordered in terms of the perceived amount of conscious attention or depth of processing. Unfortunately, the high degree of consensus may also mean that the judges were making the comparisons in terms of some other criteria (e.g., time or difficulty). Due to the relatively small number of circularities in the overall comparisons, however, it would appear that the judges were rating along the dimension on which they were trained, otherwise it would be reasonable to expect more inconsistencies in the comparisons and less consensus among the judges. It is concluded, therefore, that the high degree of agreement in fact reflects an ability of the subjects to discriminate among the various tasks along a particular dimension in a consistent fashion.

Table 1

Derived Ranks on Depth of Processing for Nine Semantic Tasks

Rater #	SIM	DIFF	DEF	2-WS	1-WS	2-FA	S-W	P-UP	L-NL	Word Pair Rated
1	2	3	1	4	5	8	6	8	8	ensemble - abduction
2	5	1	2	4	5	7	3	8	9	residue - nephew
3	1	2	4	5	6	3	7	8	9	buffoon - amazement
4	1	3	5	2	5	5	7	8	9	extermination - wench
5	1	3	5	2	4	6	8	9	7	pacifism - henchman
6	1	2	4	5	3	6	7	9	8	charlatan - impotency
7	5	4	2	3	1	6	9	8	7	maker - opinion
8	2	1	3	5	4	6	8	9	7	wholesaler - deceit
9	1	2	4	3	5	9	7	6	8	originator - malice
10	1	4	2	3	5	9	7	6	8	folly - armadillo
11	1	5	3	2	4	9	6	7	8	fallacy - arbiter
12	1	2	3	5	6	4	7	8	9	bard - charter
Σ	22	32	39	42	54	78	82	94	97	

Table 2 contains the rank-ordering of the nine semantic tasks. From an examination of the rankings in Tables 1 and 2, apart from some natural breaks, there does appear to be evidence for a real continuum. It is interesting to note that the three tasks requiring some type of relational encoding (SIM, DIFF and 2-WS) were consistently rated near the top of the scale. Tasks calling for the listing of definitive, accompanying or associative attributes (DEF, 1-WS and 2-FA) were ranked higher than tasks calling for a judgment confirming the presence or absence of a particular semantic attribute, i.e., S-W, P-UP, and L-NL.

Table 2

Ordinal Scale of Nine Semantic Processing Tasks

Derived from Paired Comparisons Data

1.	Finding a Similarity between Two Words	(SIM)
2.	Finding a Difference between Two Words	(DIFF)
3.	Writing a Definition for Each Word	(DEF)
4.	Writing a Single Sentence for Both Words	(2-WS)
5.	Writing a Single Sentence for Each Word	(1-WS)
6.	Writing 2 Free Associates for Each Word	(2-FA)
7.	Deciding if each Word is Strong or Weak	(S-W)
8.	Deciding if each Word is Pleasant or Unpleasant	(P-UP)
9	Deciding if each Word is Living or Nonliving	(L-NL)

EXPERIMENT II

The derived scale of depth-ordered semantic tasks depicted in Table 2 was used in Experiment II to assess the question of whether there exists a continuum of processing within the semantic domain. If depth of processing has been operationally defined as a result of the scaling in Experiment I, then the ranking of the tasks should be highly correlated with, and predictive of memory performance after subjects engage in each type of processing under incidental learning conditions. It was predicted that better retention would be associated with tasks rated at the top of the scale than those ranked toward the bottom of the scale.

Method

Subjects. 135 undergraduate students, both males and females, from introductory psychology sections at the University of Manitoba served as subjects in this experiment. Because the nature of the semantic tasks required that subjects be familiar with the words used in the experiment, only native speakers of English participated in the study. Subjects were run in groups of 7-12 persons and all subjects received nominal course credit for their voluntary participation.

Design. Independent groups of 15 randomly assigned subjects were allotted one of the nine processing tasks in an incidental learning paradigm.

Stimuli. The stimuli were 24 unrelated nouns selected from Paivio, Yuille and Madigan's (1968) norms. They were chosen according to the following criteria:

a) no word was an obvious associate of any other list word

b) each word was low in both related 'm' and 'I'

Throughout the present series of studies it was decided to minimize the possibility of subjects encoding via non-verbal images by keeping rated 'I' for all stimulus words as low as possible. It has been established that imagery value is a potent positive factor in the acquisition of verbal information (Smythe & Paivio, 1968; Paivio, Smythe & Yuille, 1968). It is recognized, however, that a problem exists in attempting to nominally manipulate certain item properties (i.e., level of rated 'm') while keeping others constant (e.g., rated 'I') because of the complex pattern of intercorrelations among various scaled attributes (Paivio, Yuille & Madigan, 1968).

c) none of the words was categorically related

The stimuli (see Appendix F) had a mean 'm' of 4.43 (SD = .57) and a mean 'I' of 3.87 (SD = .80).

For tasks requiring subjects to relate to two words at a time (SIM, DIFF, and 2-WS), pairs were formed so as to match as closely as possible on rated 'm' and 'I', so long as the two words did not begin with the same letter. An additional restriction was that no two living things were paired (e.g., charlatan - wench).

Two sets of stimuli were prepared. For the six single-word conditions the 24 stimuli were typed in upper case letters on acetate. They were mounted on 2 x 2 inch (5.08 x 5.08 cm) slides for projection by a Kodak Carousel Projector. The second set of 12 slides contained two words typed one above the other. For each set the order of presentation was the same for all nine tasks.

<u>Procedure</u>. Fifteen subjects were randomly assigned to each of nine experimental conditions according to the order of their appearance in the laboratory with the restriction that within a particular session there were *n* subjects in each condition before there were *n*+1 subjects in any condition. This only occurred where various tasks shared the same rate of presentation of the stimuli. An attempt was made to keep the *n* the same across experimental conditions throughout the study, although subjects who missed participating because of illness or lateness made this requirement difficult to meet with any real stringency.

Subjects were informed that the purpose of the experiment was to collect information on how people perceived and reacted to certain aspects of English words. They were told that a series of nouns were to be presented visually and they would have to complete a specific task with respect to that word (or two words) (see Appendix G for general instructions). Response booklets containing specific instructions relative to each task (see Appendix G) were issued to each subject and following a practice slide stimuli were presented either singly or in pairs according to the particular experimental condition. For tasks L-NL, P-UP, and S-W the presentation time was 10 seconds. For the 2-FA task subjects saw the word for 15 seconds, while for the DEF and 1-WS tasks the words were exposed for 20 seconds. The remaining three tasks (2-WS, DIFF, and SIM) had a longer presentation rate of 30 seconds to allow time for completion of the task. Slide onset was under the control of an electronic timer.

After list presentation the experimenter collected the booklets and issued subjects with written instructions for an unexpected test of free recall (see Appendix G). There was a three-minute period between the conclusion of the processing and the signal to begin recalling the words. Subjects were allowed a maximum of three minutes for free recall after which they were questioned as to whether they had expected a memory test. No one reported having that expectation.

<u>Results and Discussion</u>

The raw data for Experiment II are contained in Appendix H while the summaries of the various analyses are shown in Appendix I. It should be noted that comparisons among the various tasks were based on the number of words correctly recalled. The possibility that subjects in the relational encoding tasks (2-WS, SIM & DIFF) were organizing both words into a higher order memory unit (Tulving, 1968) is acknowledged. In fact, an examination of subjects' free recall sheets confirmed that subjects tended to recall words presented as pairs in adjacent positions. The mechanism involved is also seen as being similar to redintegrative memory as described by Horowitz and Prytulak (1969); recall of one item of the pair leads almost directly to the other. In the present study, however, the relationship of prime interest is between the depth-ordered scale and memory performance. The number of responses of each of the nine groups is shown in Table 3. An analysis of variance revealed an overall significant difference among the nine means, $\underline{F}$ (8,126) = 6.65. As Table 3 reveals there is some partial support for the prediction that recall would be a function of scaled depth of processing. For example with the exception of the last three tasks the order of recall followed nicely from predictions based on the independently ordered scale. Clearly, however, there are some anomalies in terms of what the scale predicts. To begin with the SIM task, rated as the task requiring greatest depth of processing, ranked eighth in overall recall. Similarly, the DEF and DIFF groups produced relatively poor recall in terms of their rankings.

Careful inspection of the recall protocols for the DEF group revealed that of the 360 definitions that should have appeared in the response booklets, on 108 occasions the definition was not known. In the few instances where something was written it was usually inappropriate, e.g., "A buffoon is a kind of monkey;" "Armadillo is a type of flower" or "A henchman is one who henches." Except for "bard" and "fallacy,"

Table 3

Mean Recall for Each Group in Experiment II

	Type of Processing and Presentation Time								
	L-NL	P-UP	S-W	2-FA	1-WS	2-WS	DEF	DIFF	SIM
Seconds	10	10	10	15	20	30	20	30	30
$\bar{X}$	5.20	7.30	7.20	7.46	10.20	11.53	8.73	9.40	6.80

only 11 of those 108 words appeared in the subjects' recall sheets. In the case of "bard" and "fallacy" they were the third and second last words presented, respectively. Despite the three minutes between the end of list presentation and the beginning of recall a recency effect may account for the few times these two words were recalled at all. Bjork and Whitten (1974) and Tzeng (1973) have shown that recency effects are not altogether dissipated by interpolated activity. It is concluded, therefore, that the DEF, SIM, and DIFF tasks could not be expected to show good recall when such a large proportion of the words were unfamiliar to some subjects. In fact, it seems clear that these tasks which required either listing, comparing or contrasting definitive attributes of the words or their referents, were beyond the capabilities of subjects when the stimuli were so unfamiliar that essential features could not be extracted from semantic memory. This line of reasoning is supported by Jacoby, Bartz and Evans (1978) who claim that if the task demands require the subject to "deal with meaning (this) should have little effect if the material is so impoverished as to be unable to support a meaningful analysis" (p. 332).

It is likely that subjects were attempting to retrieve items by utilizing a type of backwards scanning mechanism (Lockhart, Craik & Jacoby, 1976) to examine the contents of episodic memory. Because of the time lag between acquisition and recall due to the relatively long presentation rates associated with these tasks, it is not surprising that the items were relatively inaccessible.

Despite the low Spearman rank order correlation ($\rho = .38$) between the scaled tasks and the recall scores, primarily due to the problems described above, Table 3 indicates that generally recall increased as a function of rated depth of processing. Since there was some general support for the depth-ordered scale it was decided to repeat Experiment II using stimuli more compatible with the working vocabularies of the average college or university student, thus enabling a fairer assessment of the higher order tasks and their effects on memory performance.

EXPERIMENT III

For the most part the research design, procedures and tasks used in Experiment II remained the same for the present study. The only obvious change was to increase the meaningfulness of the stimuli while attempting to keep the 'I' values at a minimum. As explained earlier, this was done to minimize the possible additive effects of imagery during acquisition across the various incidental orienting tasks (Paivio, 1976).

Method

Subjects and Design. The subjects were 135 University of Manitoba undergraduates, both males and females, participating to fulfill an introductory psychology course requirement. Again, participants were restricted to those whose first language was English. The subjects were randomly assigned in order of their appearance to one of the nine between-subjects conditions.

Stimuli. The stimuli were 24 unrelated nouns selected from the same norms as Experiment II (Paivio, Yuille & Madigan, 1968). Again, obvious associates and categorically related words were not chosen. The stimuli were still relatively low meaningful words with a mean 'm' of 5.77 (s.d. = .78) and mean 'I' of 4.61 (s.d. = .81). The average 'm' for these words was significantly higher than those used in Experiment II, $t(46) = 7.05$, and yet still below the mean 'm' for Paivio, Yuille and Madigan's (1968) norms. The rationale for increasing the 'm' value for the stimuli as mentioned above was based on the assumption that subjects would be more familiar with their meaning attributes and hence better able to perform the more complex tasks (i.e., SIM, DIFF, and DEF). The complete list is contained in Appendix F.

Procedure. The experimental procedures detailed on pages 19 and 20 and instructions to subjects as described in Appendix G for Experiment II were retained for Experiment III. In order to restrict presentation rate to three levels, however, the 2-FA task which occupied the subjects for 15 seconds per word in Experiment II was altered so that subjects would be required to produce a single free associate for each stimulus word in 10 seconds (1-FA), thus equating this task for time with all other single words tasks, except DEF and 1-WS which were maintained at a presentation rate of 20 seconds per word.

Results and Discussion

The raw data for Experiment III are contained in Appendix H while the summary of the one-way analysis of variance is shown in Appendix I. As in Experiment II, results indicated an overall significant difference among the mean number of words recalled for each of the nine experimental conditions ($F(8,126 = 7.31)$. These means are presented in Table 4 and with a single exception the data conform with the predictions based on the scale developed in Experiment I. In fact, the Spearman rank order correlation coefficient was highly significant ($\rho = .98$).

Table 4

Mean Recall for Each Group in Experiment III

	L-NL	P-UP	S-W	1-FA	1-WS	2-WS	DEF	DIFF	SIM
Seconds	10	10	10	10	20	30	20	30	30
$\bar{X}$	6.40	7.53	7.93	8.06	8.26	10.53	11.60	10.86	11.80

Probing the omnibus F-test, again with the use of Tukey's (HSD) method for pairwise contrasts (see Kirk, 1968) revealed several significant pairwise comparisons. For example, judging how two words might be similar in meaning (SIM) yielded higher recall than all other conditions except DEF, DIFF, and 2-WS. Similarly DEF outperformed all groups except 2-WS, DIFF, and SIM. Other significant contrasts were DIFF was superior to both P-UP and L-NL. Additionally 2-WS recalled significantly more than L-NL.

Of all the significant differences, one of particular importance is the difference between DEF and 1-WS since (a) they are both single word tasks and (b) they both received the same rate of presentation (i.e., 20 seconds). DEF was rated as involving greater depth of processing than 1-WS and the difference in level of recall empirically validates their difference in rank at least for this level of meaningfulness.

The suspicion that the unexpected low recall of groups associated with the SIM, DIFF, and DEF groups in Experiment II was due to the impoverished nature of the stimuli vis-à-vis these particular tasks was confirmed by the results of the present study. Using a series of *t*-tests to compare differences in level of recall for each condition from Experiment II to Experiment III it was found that both SIM and DEF groups increased in their level of recall significantly ($t(28) = 3.72$ and $t(28) = 2.67$, respectively). The DIFF group also increased its level of recall ($t(28) = 1.70$) although this difference was not of the same magnitude. In all other conditions there were no statistically significant improvements in recall.

Interestingly, the level of recall for 1-WS and 2-WS was less in Experiment III than Experiment II. One possible interpretation is that in Experiment II these items, being essentially disimilar, as well as low in meaningfulness, were given a more unique encoding which subsequently proved to be more resistant to inter-item interference at recall (Wickelgren, 1975). To bolster this point further Stein (1978) and Craik and Tulving (1975) both agree that, in addition to depth of processing, retention also rests on the subject's ability to encode inputs in a unique way thus enabling the differentiation of one encoding from subsequent encodings. To reiterate, we suggest that the tendency for 1-WS and 2-WS groups to show better recall in Experiment II with lower 'm' words may be explained by positing that the sentential contexts for these unrelated words were so dissimilar as to provide unique encodings which subsequently facilitated recall.

While the overall results of Experiments II and III appear to validate the scale and provide general support for the hypothesis that deeper or more elaborate encoding within the semantic continuum leads to better memory performance, at least for free recall, there may be some alternative explanations that can account for the present data.

It would appear that recall might also be an increasing function of presentation time despite considerable evidence to the contrary cited earlier. Certainly there are some large differences in study time among the single word tasks. Before this explanation can be entertained seriously, however, one has to account for the significant pairwise comparison (DEF > 1-WS) already mentioned since both tasks enjoyed the same presentation rate. Also, DIFF and SIM recalled more words than DEF although the total study time was six and eight minutes for the DIFF, SIM and DEF respectively.

The results of the present experiment support the concept of a continuum of depth of processing within the semantic domain. Taken together with the data of Seamon and Virostek (1978) there would seem to be good evidence for the levels of processing hypothesis both within and across domains.

EXPERIMENT IV

Experiment IV attempted to resolve some of these concerns about the possibility of relational encoding in some tasks and total study time differences as well as test the generalizability of the effects obtained in Experiments II and III by:

1) varying further the level of meaningfulness of the stimulus materials.

2) restricting the orienting tasks to five, all of which involved single word tasks. A new semantic task, i.e., judging whether a word is an exemplar of a given category was included in the range of tasks.

3) altering the demands placed on the subject so as to require ratings on a bipolar dimension for each task. This change in experimental procedure allows for equality of presentation time across each of the semantic processing conditions.

Method

Subjects. Fifty University of Manitoba undergraduates were tested in groups of 7-12 and received nominal course credit in return for their participation. Once again the language restriction operated and only native English speakers were admitted to the study.

Design. All subjects in the present experiment received the same set of stimulus words and employed each of the five processing conditions in a completely within-subject design.

Stimuli. The stimuli were 34 unrelated nouns taken from Riegel's (1965) Michigan Restricted Association norms and represented an overall increase in level of meaningfulness compared to the words used in Experiment II and III (see Appendix F).

Tasks. The tasks performed by the subjects were as follows:

1. Sentence (SENT). This task required the subject to rate a sentence on a "good-poor" dimension, e.g., "The anger of the crowd grew in intensity over the referee's decisions."

2. Definition (DEF). The subject was required to rate a definition on a "good-poor" dimension, e.g., "Anger is a strong emotion, like hate, where often the adrenalin runs high."

3. Free Associate (FA). e.g., "the first word that comes to mind when you look at the word (ANGER) could be HATE." Subjects rated on a "likely-unlikely" scale.

4. Category (CAT). "The word above (ANGER) is a type of emotion. Subjects rated along a "good example - poor example" dimension.

5. Living-Nonliving (L-NL). Subjects were asked to judge if each word (e.g., ANGER) was a living or a nonliving thing.

For each processing condition the Experimenter constructed the rating tasks so that subjects would be expected to always rate towards the positive end of each continuum (except of course in the L-NL condition). That is to say, each definition was written so as to include at least three of the definitive attributes listed in Riegel's norms. Similarly, the SENT sentences written for each word were approximately the same length as the DEF sentences but did not necessarily contain a definitive attribute. In the FA condition the strongest free associate given for each word was taken from that section of the norms. For the CAT condition the most often cited superordinate was chosen, against which the exemplar was to be rated as a good example or a poor example of that category. Appendix G gives an example of the definition, sentence, free associate, category judgments and living-nonliving tasks.

Booklets were made up according to the following procedures:

1) The 30 TBR words were randomly broken up into groups of five.

2) Within each group of five words tasks were counter-balanced across subjects and this was continued until 50 booklets were made up.

3) The order of words within the booklets was shuffled thoroughly so that the order of words and tasks was completely randomized across subjects.

4) The only restrictions to 3) was that no two successive words started with the same letter and no more than two adjacent items required the same processing task.

The remaining four words were used as two primacy and two recency buffers and were counter-balanced and randomized in the same way as already mentioned before being added to the front and back of each booklet.

Procedure. Each subject received a booklet face down on the desk and was told, as in earlier studies, that the experiment was concerned with how people perceive and relate to certain aspects of common English words. Briefly, the nature of each task was described and demonstrated on an overhead projector. The time allotment for each task was fixed at 10 seconds per word because of the within-subject nature of the design. Subjects were instructed to turn the page and complete the next task at the sound of a tone but not before.

Following completion of the rating tasks subjects were given an unexpected test of memory under free recall conditions. At the conclusion of free-recall subjects rated all possible pairs of tasks in a paired-comparison procedure with respect to the five semantic processing tasks used in this experiment (see Appendix C for an example of the 10 comparisons). The criterion for rating each pair was the same as in Experiment I, i.e., which task required more conscious effort or attention to carry out?

Results and Discussion

Ratings. The response patterns for each of the 50 subjects are depicted in Appendix D. Note that seven raters (Ss 9, 11, 16, 33, 34, 39, and 45) were somewhat inconsistent in their rankings and Kendall's coefficient of consistency in each case was .8. This coefficient was subsequently utilized to determine that each rater's preferences contained a single circular triad (e.g., A→B→C→A). The probability of d (the number of circular triads) for this few stimuli is quite high at .766 (see Kendall & Babington-Smith, 1939) and accordingly the null hypothesis that the choices were allotted at random cannot be rejected for these subjects. However, there was a high degree of consistency among the other 43 raters. An analysis to determine the degree of overall agreement included all 50 judges. Kendall's coefficient of concordance (W) was found to be .58. The X^2 associated with this value of 116.5 is highly significant (see Appendix E. The probability under the null hypothesis (that the judges' ratings are unrelated) is therefore rejected. The rank order of the five tasks based on the paired comparison data is shown in Table 5.

Recall. The mean recall of 30 words for each type of processing is shown in Table 6. Although the booklets contained 34 words, in order to minimize the primacy/recency effects the first two and last two items in each booklet were struck from the recall sheets of each subject. An inspection of Table 6 indicates an almost complete reversal in terms of scale order of the tasks in Table 5.

Table 5

Ordinal Scale of Five Semantic Processing Tasks

Derived from Comparisons Data in Experiment IV

1. Rating a Sentence on "Good-Poor" dimension (SENT)
2. Rating a Definition on "Good-Poor" dimension (DEF)
3. Rating a Given Free Associate as "Likely-Unlikely" (FA)
4. Rating a Word as "Good-Poor" example of a given category (CAT)
5. Rating words as Living-Nonliving (L-NL)

Table 6

Mean Recall of Words as a Function of

Type of Semantic Processing in Experiment IV

	L-NL	CAT	FA	DEF	SENT
$\bar{X}$	2.74	1.84	2.48	2.00	1.20

The obtained F-ratio was significant, $F(4,196) = 15.667$, (see Appendix I) and Tukey's HSD test (Kirk, 1968) revealed the L-NL yielded significantly higher recall than all other conditions except FA. Additionally, both DEF and FA tasks produced better recall than SENT.

The immediate and perplexing question that springs to mind is why did the SENT task, ranked as the task requiring greatest depth of processing in terms of conscious effort, lead to the poorest recall?

Moreover, how does one account for the superior retention of the L-NL group against all other groups except FA in spite of its being ranked at the bottom of the scale by 41 out of the 50 judges?

It is difficult to account for the present findings in any straightforward manner especially the marked increase in recall associated with L-NL task. One possibility is that subjects, when carrying out the L-NL task, did so quickly and subsequently engaged in further encoding strategies for the remainder of the 10 second period allotted for each task. It was obvious, on the basis of casual observation, that subjects were completing this task in less than half of the allotted time. Why they might engage in additional processing remains unclear but subjects may well have been sensitized by the nature of the other processing tasks in this within-subject design to engage in additional processing. One could also speculate that the L-NL task called for an objective judgment (i.e., one that is potentially right or wrong) hence requiring more conscious attention, but then so did the CAT task and words associated with this task were relatively poorly recalled. The poor recall of CAT group is interpretable in the light of a study done by Buschke and Lazar (1973) who found that category cues supplied only at input (as in the present study) tended to create a type of cue overload since they recalled fewer words than a group who received the TBR words in the absence of any cues at input.

Postman (1975) suggested that a particular orienting task e.g., L-NL may not only influence the initial encoding of the TBR item but if time permits secondary activities such as displaced rehearsal may take place. Jacoby (1974) called this "looking back" through the list and the effect of this may have led to some functional organization of the list that enhanced subsequent levels of recall. What is being suggested is that the within-subject design potentially enabled the subject to engage in additional processing (of the kind required by some of the 'deeper' tasks) when time permitted. This could not have occurred for the SENT and DEF tasks since subjects were occupied for most of the 10 second period for rating.

One still has to account for the present pattern of results which are discrepant not only with the scale but also the data obtained in Experiments II and III. As Postman (1975) correctly pointed out "the item property nominally manipulated may not be the functionally effective one" (p.45), and this may be due to the complex intercorrelation of item attributes as mentioned earlier. In the present study there was an overall increase in the 'm' of the stimulus words. As rated 'm' of word stimuli increases it is more difficult to hold "I" at a low or constant value due to the high intercorrelation between the two attributes (Paivio, Yuille & Madigan, 1968). Perhaps in the context of making a L-NL judgment the subjects were encouraged to summon up an image of the word's referent thus adding a highly effective imagery dimension to the subject's encoding. Were this the case, the recall results would not be surprising since the advantage of imagery over verbal encoding strategies are well documented in free recall (Paivio, 1971). In addition, Paivio (1972 , 1976) posited that highly concrete items may be dually encoded as a result of both imaginal and verbal processing. Moreover, he postulated that these two codes may combine additively to facilitate memory performance. Levy and Craik (1975) advanced further evidence to the effect that separate codes may be additive in their effects on memory.

The relatively poor recall associated with the two tasks where the TBR word was embedded in a sentential context also bears close scrutiny. In terms of imagery it has been demonstrated that sentences are less likely to invoke an image than concrete words or their referents (Paivio, 1971). A more plausible explanation might involve the nature of orienting tasks themselves which were changed from subject-generated (and perhaps idiosyncratic) encodings to a rating task where the encodings for SENT, DEF, FA, CAT tasks were supplied by the experimenter. It is probably legitimate to ask whether presented encodings can potentiate a conflict or type of interference with the subject's own encoding. A study by Duffy and Montague (1971) provides partial support for this interpretation. They showed that providing subjects with sentence mnemonics hindered learning performance. These authors concluded that experimenter-supplied mnemonics may constitute a source of interference and effectively prevent subjects from generating more distinctive encodings for themselves. Certainly there is ample evidence to demonstrate that subject-generated mnemonics produce better memory performance than equivalent ones supplied by the experimenter (Bobrow & Bower, 1969; Bower & Winzenz, 1970). Bobrow and Bower argued that subject-generated encodings actively involve the subject, thereby fostering a deeper understanding of verbal sequences, and hence facilitate recall.

That the change in the experimental demands made on the subjects may have affected the type of encoding merits further consideration. For instance, the subjects were no longer required to write each stimulus word and it may be argued that in the SENT and DEF tasks, for example, the subjects were rating the sentences <u>per se</u> as "good" or "poor" without being disposed to selectively attend to the TBR word embedded in that sentence (Postman & Kruesi 1977). In fact, both tasks may have been construed by subjects as judging "how well" the TBR word fitted into the given sentence. This task parallels tasks used by Craik and Tulving (1975) and Hyde and Jenkins (1973). While the former study found providing sentential contexts for TBR words facilitated recall, Hyde and Jenkins' (1973) results were similar to the present findings, i.e., experimenter-provided sentential contexts led to relatively poor recall. Subjects may not have been selectively attending to and encoding the TBR word as consistently for the SENT task where often the TBR word occurred at the end of the sentence as part of the predicate (e.g., the detective in charge of homicide expected an early arrest). Perhaps it would have been better attended to if it had been the subject of the sentence, as was the case in the DEF task where the TBR word always occurred as the subject of the sentence, e.g., "Homicide, like murder, is the intentional slaying of a human being."

Furthermore, it could be argued that although the sentence for the DEF and SENT tasks were experimenter supplied, the sentential contexts were more distinctive for the DEF task in the sense that definitive features or attributes of the TBR word were contained in each sentence. This notion of "distinctiveness" parallels the concept developed by Moscovitch and Craik (1976) and empahsized by Reed Hunt, and Mitchell (1978) where distinctiveness refers to the extent that features encoded as a result of some incidental task are specific to the TBR word. If this were the case, however, recall for the DEF group should have produced rich encodings and, therefore, better recall. Although the definitive attributes were clearly specified in each sentence (see Appendix F) as

Klein and Saltz (1976) noted, perhaps the attributes have to be both specified and activated during encoding for recall to be maximal.

EXPERIMENT V

The overriding concern with the results of the previous experiment was the possibility that there may have been severe carry-over effects, especially to those tasks where the rating was completed well within the prescribed time period of 10 seconds, thus affording subjects the opportunity for further encoding. It is felt that this may have been especially true for the L-NL, CAT, and FA tasks. Accordingly, it was decided to repeat Experiment IV but change the rate of presentation from a fixed interval (10 seconds) to a self-paced procedure for each subject to minimize the possibility of subjects engaging in processing additional to the task requirements. Secondly, it was decided to revert to a between-subject design to eliminate potential carry-over effects across experimental conditions.

Method

Subjects. Seventy-five male and female subjects were tested in groups of 5-7. In return for their voluntary participation they were given a nominal course credit. The subjects were selected from the same population, with the same language restriction as in all previous experiments.

Design. The experiment was a between-subject design utilizing the same five semantic orienting tasks from Experiment IV. Groups of 15 subjects were randomly assigned to each type of processing task according to the order of their appearance in the laboratory.

Stimuli and materials. Twenty-four words were randomly selected from the list of 30 items used in Experiment IV for presentation in booklet form (see Appendix F). The order of word presentation was the same across all conditions. Once again buffer items were employed to counter possible primacy and recency effects.

Procedure. Following the same general introduction as before, subjects in each condition were instructed as to the nature of their particular rating task. Examples were demonstrated on an overhead projector and the experimenter answered individual queries regarding the experimental procedure. Subjects were permitted to proceed through the booklets at their own pace signalling to the experimenter at the conclusion of the task. The time taken for the task was noted for each subject following which he/she was given a signal to turn over the sheet containing written instructions for free recall (see Appendix G). Following free recall subjects were debriefed and questioned as to whether they had expected a memory test. None admitted having that expectation.

Results and Discussion

The reader is referred to Appendix H for an overview of the raw data for this study and to Appendix I for the relevant statistical analyses.

Recall. Table 7 shows the mean recall for each of the five experimental conditions. An analysis of variance revealed a significant *F* ratio ($F(4,70) = 10.18$), with all pairwise comparisons between SENT and the other four conditions being significant according to Tukey's (HSD) criterion (Kirk, 1968). None of the other pairwise comparisons was significant.

It is disconcerting to note that once again the order of recall bears little resemblance to the ordinal scale as determined in Experiment IV. Apart from the low recall of the SENT condition there are no significant differences between the means of the remaining four tasks. This is interpreted to mean that how facilitative a particular orienting task is as an encoding mechanism is dependent upon both the nature of

Table 7

Mean Recall and Average Processing Time Per Item

for Each Group in Experiment V

Type of Processing					
	L-NL	CAT	FA	DEF	SENT
Mean Recall	9.80	11.53	8.73	11.2	5.60
Mean Processing	4.33	6.57	7.53	8.92	9.16

the TBR material and the degree of conscious or cognitive effort involved in encoding (i.e., experimenter supplied vs. subject-generated encodings). The question again arises as to whether rating an externally provided encoding requires as much conscious attention or effort as an encoding generated by the subject.

Processing time. As can be seen in Table 7 there are substantial differences in the amount of time spent rating the words in terms of each particular type of processing, ($F(4,70) = 16.02$). Several pairwise comparisons as determined by Tukey's (1953) HSD test were also statistically significant. The L-NL group's average encoding time of 4.33 seconds per item was faster than all other conditions. In spite of this, mean recall of L-NL was superior to that of the SENT group, for example, where the average time required to complete the orienting task for each item was more than twice as long. This finding adds further weight to studies (already cited) reporting a lack of correlation between processing time (once thought to be an independent index of depth of processing) and retention.

The present study would appear to cast doubt on some of the reasons offered for the disparate results of Experiments II and III versus IV. To begin with, it can no longer be claimed that the relative advantage of the L-NL conditions vis-à-vis the other semantic processing tasks was totally a function of carry-over effects or displaced rehearsal. By purposely running subjects in small groups, it was obvious to the author that subjects were reasonably task-oriented, i.e., they seemed not to spend unequal amounts of time performing the task for successive words. On the contrary, they appeared to work through the booklets at a steady, even pace. Thus, the performance of the L-NL task in Experiment IV cannot unequivocally be attributed to any additional encoding that may have taken place during that portion of time left over after the rating of each word which it is now assumed must have taken less than five seconds. Note, however, that the recall of L-NL in the present study ranks third as against first in Experiment IV. Perhaps, reverting to a between-subject design, thus eliminating sequencing effects from task to task, partially accounts for this relative decrement in recall.

What still needs to be explained is: (a) the poor recall of the SENT group despite taking the longest amount of time to complete the task for each word. (b) The apparent absence of a continuum of processing within the semantic domain, the SENT group notwithstanding.

To restate an earlier line of reasoning, it may be argued that in the SENT condition the features of the task were irrelevant or even antagonistic for optimal encoding of the TBR word in the context of rating the sentence along a "good-poor" dimension. Despite being instructed to look at the word first and then rate the sentence in terms of the word at the top of the page some subjects reported during debriefing that they often took no particular notice of the word, and instead simply rated the sentence itself. There is no way of knowing how often this happened, especially since the subjects were not required to copy the TBR word down, as they were in Experiments II and III.

The second source of concern is twofold. To begin with there is no solid evidence for a continuum of semantic processing in terms of significant differences in recall between the various levels of encoding. Secondly, there is no relationship between the scale, derived from the paired comparison data, and the retention functions in Experiments IV and V. It should be borne in mind, however, that both the scaling and recall data were obtained from the same subjects. This was not the case for the originally derived scale. Moreover, in Experiment IV there was no concerted attempt to sensitize the raters to the concept of depth of processing or its associated connotations beyond the printed instructions (see Appendix G) as was done with the original scaling exercise. It may well be that although there was high agreement among the raters, they could have been allotting their preferences in terms of some other criterion, e.g., time required to complete the task. There is some indirect evidence to support this notion. The scale ordered tasks depicted in Table 5 correlate perfectly with the amount of time required to carry out each of the tasks (see Table 7). Since the rankings were done by all subjects immediately after recall, it is quite possible that 'time' became the dominant dimension in their minds. If this, in fact, happened and since it is clear from the results of the present study that time *per se* is not an index of depth, or a predictor of retention, it is not surprising that the previously obtained relationship between scaled semantic tasks and their retention measures dissipated completely in Experiments IV and V.

EXPERIMENT VI

In order to clarify some of the issues involved, and in an attempt to more carefully delineate and control the variables related to depth of processing, it was resolved to carry out a further study where the original design was reinstated, i.e., subjects were required to generate their own encodings for each of the semantic tasks. It was also decided to maintain the rated 'm' of the stimuli at a high level so that the question of whether the breakdown in the continuum of processing was a result of the change in the task demands on the subjects or was due to the increase in the meaningfulness of the stimuli could be answered. It may be that with highly familiar stimuli a reasonably strong memory trace can be formed irrespective of the type of semantic encoding employed.

Experiment VI was a replication of Experiments II and III in terms of general design and procedures. Changes relating to the choice of tasks and the selection of stimulus materials are detailed below.

Method

Subjects and Design. The subjects were 72 undergraduates, both males and females, drawn from the same population described in earlier studies. Independent groups of 18 randomly assigned subjects were allotted one of four semantic processing tasks in an incidental learning paradigm.

Tasks It was decided to use only four of the tasks from the original scale, namely, SIM, DEF, 1-WS and L-NL. In this way both ends of the scale were represented (SIM and N-NL) and additionally, the only significant post hoc comparison obtained in Experiment III between tasks requiring the same processing time (DEF and 1-WS) could again be made.

Similarly, recall for SIM was greater than for 1-WS and this comparison could again be drawn.

Stimuli. The stimuli were 24 unrelated nouns chosen from Paivio, Yuille and Madigan's (1968) norms with the same restrictions already mentioned for Experiments II and III. The mean 'm' was 7.6 (SD = .60). The mean 'I' was also high at 6.06 (SD = .62). See Appendix F for the complete list of words. The average 'm' was significantly higher than those words used in Experiment III, ($t(46) = 9.08$). Unfortunately, it proved difficult to keep the 'I' values at a minimum because of the high intercorrelation between these two variables at this level of meaningfulness.

Procedure. The experimental procedures, general and specific instructions described in detail on pages 19 and 20 and laid out in Appendix G were retained for this experiment.

Results and Discussion

The raw data for each of the 18 subjects in the various experimental conditions are contained in Appendix H while the summary of analysis of variance appears in Appendix I. Table 8 shows the mean recall following incidental learning in each of the four semantic processing tasks.

An examination of Table 8 reveals that, while the F-ratio is not significant ($F(3,68) = 2.18$), there is a return to an order of recall in terms of what the original scale predicted. This garners further support for the present thesis of a functional continuity of depth of processing within the semantic domain.

When the results of the present five recall studies are considered together, it seems apparent that retention is not solely a

Table 8

Mean Recall for Each Group in Experiment VI

	Type of Processing and Presentation Time			
	L-NL	1-WS	DEF	SIM
Seconds	10	20	20	30
$\overline{X}$	12.33	13.16	13.72	14.83

function of depth of processing within a specified domain. Rather the formation and subsequent durability of the memory trace is most likely a multiplicative relationship that includes at least some of the following variables.

1. The depth of processing of the stimulus materials in terms of the amount of conscious effort directed towards the extraction of deeper semantic features of the TBR items.

2. The compatibility between the inherent meaningfulness of the stimulus materials and the learner's existing cognitive store of related concepts and knowledge.

3. The degree to which the learner is actively engaged in the encoding process.

Taking these variables one at a time, the findings of Experiments II, III, and VI are reasonably predictable from the independently derived index of depth of processing described in Experiment I. The discrepancies that were manifested in the recall data of Experiments IV and V might be accounted for when considering the second and third variable. For one thing, there was a noticeable increase in level of meaningfulness of the TBR material in Experiments IV and V and it has been suggested that retrieval from the episodic memory trace at recall may be effected by scanning the recently encoded events or by a reconstruction of the stimuli from the stored contents of semantic memory (Lockhart, Craik & Jacoby, 1976). Where there already exists a large number of relevant features in semantic memory it is postulated the TBR item can be accessed by a variety of attribute-specific encoding strategies, thus enhancing the probability of correct recall. If the TBR item has few existing features in semantic memory then tasks requiring the extraction of the deepest features will result in the formation of a richer, more durable trace. In the case of low 'm' stimuli (Experiment II) there was such a dearth of features upon which to draw that those tasks requiring the analysis of deeper features (i.e., SIM, DIFF, and DEF) led to poor recall because the subjects were for the most part unable to make the required responses with respect to those selected attributes of the TBR units. Also Kintsch(1974) stated "a person's memory structure, his knowledge of the use and meaning of a word determines whether and how a memory episode will be encoded" (p.78). Craik and Tulving (1975, Experiment 7) contended that the results of their study indicated that elaborate semantic processing may not facilitate later recall if comprehension of the stimulus is not achieved. Furthermore, as mentioned earlier, the strategies of these subjects was probably restricted to scanning back through the temporally tagged items in episodic memory. With the large time lag between presentation and recall for these tasks (SIM, DIFF, DEF) it seems likely these items became increasingly inaccessible (Craik & Jacoby, 1975). Note, however, an overall improvement in performance for these tasks relative to the other semantic processing tasks in Experiment III, where now the relevant features could be extracted from semantic memory and used as inputs into the episodic memory system. Additionally, in Experiment III, the order of recall for all nine tasks mirrored almost perfectly recall as predicted by the scale.

With a further increase in meaningfulness (Experiment VI), however, although the continuum remained intact, the boundaries, as reflected by the recall scores, narrowed considerably to the point where it was not possible to demonstrate statistically that a 'deeper' task requiring finding a similarity between two unrelated words (SIM) produced significantly better recall than any of the other tasks selected from the continuum. It is reasoned that being able to extract relevant features from semantic memory provides an additional boost to the retrieval of the memory trace from episodic memory at the time of recall. Since the TBR units now are well known and logically have many relevant features stored, it is conceivable that almost any type of semantic analysis might activate the complete network of attributes associated with the TBR item. Note that the original Craik and Lockhart (1972) model accommodates nicely the increased level of recall across all tasks in Experiment VI when those authors stated "At deeper levels the subject can make greater use of learned rules and past knowledge; thus material can be more efficiently handled and more can be retained" (p.676).

This may have occurred in Experiment VI in addition to the type of scanning operation postulated by Lockhart, Craik and Jacoby (1976). In fact, where both strategies can be employed it is arguable that the probability of recall is enhanced considerably. The data of Experiments II, III, and VI seem consonant with this line of reasoning.

What is most perplexing, however, is the effect of the third factor listed above and its relationship to the overall pattern of results obtained in the present series of experiments. Why, for example, did changing the task demands in Experiment IV and V lead to such radical recall differences compared to the rather neat patterns that emerged in the other studies? Arguments relating the efficacy of subject-generated encoding versus experimenter-supplied encodings have already been presented. In a related vein, one has to view with some uncertainty the assumption that subjects faithfully comply with the experimental instructions. Further, it is questionable whether with highly meaningful material the subject can control the type or extent of encoding that takes place. Lockhart, Craik and Jacoby (1976) cite the well known Stroop effect as an example of some subjects' relative inability to control their level of encoding by complying with experimental instructions. This problem is endemic to the nature of incidental learning studies and complicates enormously the effects, analysis, and interpretation of various kinds of instructions given to subjects in incidental learning paradigms.

The problem alluded to here was encountered by Paivio and his associates who were exploring the comparative effects of imaginal and verbal encoding strategies on the retention of verbal materials that varied in level of concreteness. Their hypothesis was that a set to image highly concrete noun pairs should lead to better retention than verbal encodings whereas there should be no differences between the two types of processing for abstract pairs. A series of studies (Paivio &

Yuille, 1967; Paivio & Yuille, 1969; Yuille & Paivio, 1968) failed to demonstrate the predicted interaction between level of concreteness and type of encoding. Finally, Paivio and Foth (1970) obtained the interaction by requiring the subjects to overtly generate their own encodings, i.e., they were required to pictorially represent their images or write down the verbal encoding they had generated. It is entirely feasible that in the present set of studies we can only claim to have had a fairly strong measure of control over the nature of the subjects' elaborative strategies in Experiments II, III, and VI where overt subject-generated encodings were recorded in the context of each semantic processing task. It was in these studies that a reliable relationship was found between the depth-ordered scale and the various retention functions.

Conclusion

The purpose of the present series of experiments was twofold. Firstly although much evidence pointed strongly to the notion that semantic processing consistently ensured a more durable memory trace than processing material using coding operations from non-semantic domains (Seamon & Virostek, 1978) no one had seriously entertained the thesis that within a single domain deeper levels of processing could produce stronger memory traces. Craik and Tulving (1975) suggested that at a given level more elaborative or richer encodings would increase the probability of recall. Lockhart, Craik and Jacoby (1976, p. 81) stated that further elaborations within a domain and hence the richness of the memory trace was a function of the number and nature of the features analyzed. To the present author the term "depth of encoding" as originally proposed (Craik & Lockhart, 1972) might be used to predict memory performance within a domain, specifically, the semantic domain.

This naturally leads us to the second purpose of the present study. As some researchers (Baddeley, 1976; 1978; Nelson, 1977) had nicely pointed out, the concept of depth was potentially meaningless because of its inherent circularity. Seamon and Virostek (1978) dispelled this criticism by effectively demonstrating that recall increased as a positive function of subject defined depth of processing. In order to investigate the possibility of a continuum within a single domain it was decided to independently define depth by deriving an ordinal scale of selected semantic orienting tasks. The criteria employed to obtain the scale were based on the latest descriptions of the concept as expressed by Lockhart, Craik and Jacoby (1976). They stated "to the extent that the task requires further processing of the stimulus input conscious effort is involved" (p. 81). Of course, the idea of attention as the central processor in primary memory was not new. It was discussed fully by Craik and Lockhart (1972) and its history is well documented by others (Murdock, 1972) with its origins being attributed to William James (1890).

The findings of the present series of experiments tends to confirm the hypothesis that within a single domain greater depth of processing, as defined by tasks involving progressively deeper cognitive analyses, is associated with higher levels of retention. After obtaining a depth ordered scale (Experiment I) where subjects appeared to be able to qualitatively differentiate among the semantic tasks, subsequent studies (Experiments II, III, and VI) provided a broad base of empirical support for the scale. The anomalous findings in Experiments IV and V have been discussed elsewhere. It should be stressed, however, that it is probably no longer tenable to attempt to account for memory performance solely in terms of a single variable such as depth of processing. Jacoby, Bartz and Evans (1978) argue convincingly for the need to identify classes of situations wherein particular functional relationships may be observed. More specifically, they imply that future researchers should "treat meaning (a variable that has traditionally been related to the rate of learning and level of retention) as being dependent on both the meaning-producing activity of the learner and the potential meaningfulness of the material" (p. 342).

The overall findings of the present studies can be interpreted in the light of Craik and Jacoby's (1975) theoretical account of the processes thought to be operating in short-term retention. The present author agrees with the contention that memory performance is determined in part by both the depth of stimulus analysis and the compatibility of the TBR item with the analyzing structures. The latter is taken to refer to existing cognitive structures, whose contents have been referred to as constituting semantic memory (Craik & Jacoby, 1975). It is stressed, however, that the analyzing mechanism necessary for perceptual encoding is activated by the process of conscious attention. Presumably, tasks requiring more conscious attention produce stronger traces which become inputs in the episodic memory system. These inputs can be formed both for the results of perceptual analyses and analyses requiring the extraction of stored features from semantic memory.

Retrieval from episodic memory can occur in one of two ways. Jacoby (1974) proposed that recently encoded events can be accessed by a backwards search or scanning operation with the likelihood of successful retrieval diminishing as the time between input and recall increases. Alternatively, after some delay the subject may take recourse to a type of reconstructive strategy where the contents of semantic memory (that may have been involved in original encoding) are actively deployed . This type of feedback is postulated to be particularly effective for the reconstruction of deeper semantic encodings, thus producing better recall.

In the present studies (especially Experiment II) it is suggested that in some instances, subjects were using only the scanning mechanism as a retrieval strategy, especially where the combination of the nature of the TBR material and the task demands effectively precluded the use of the reconstructive strategy. The latter was probably ineffectual in those cases where little or no information was available from semantic memory as an input during the original formation of the episodic memory trace. Presumably, this happened with DEF, SIM and DIFF tasks in Experiment II. Also, there was a much longer delay between presentation and test for items associated with these tasks, thus limiting the effectiveness of the one remaining retrieval strategy, scanning back through the temporally ordered episodic memory traces. If one wonders why the impoverished nature of the stimuli did not interfere with subjects carrying out the shallower tasks (e.g., Pleasant-Unpleasant) that still, nevertheless, required them to orient to meaning attributes of the TBR material, it could be argued that subjects tended to make better use of the scanning strategy because of the relatively shorter time lag between presentation and test.

With highly meaningful stimuli subjects were possibly making use of both retrieval modes. Moreover, the compatibility of the TBR items existing cognitive stores was extremely high. Therefore, relatively little conscious effort was required to carry out each of the semantic tasks. If conscious effort or attention is related to depth, the relative superiority of the higher ordered tasks (found to some extent in Experiment II and more strongly in Experiment III) would be expected to dissipate as the meaningfulness of the material increases. This is precisely what happened in Experiment VI.

References

Adams, J.A. Human Memory. New York: McGraw-Hill, 1967.

Arbuckle, T.Y. & Katz, W.A. Structure of memory traces following semantic and non-semantic orientation tasks in incidental learning. Journal of Experiment Psychology: Human Learning and Memory, 1976, 2, 362-369.

Atkinson, R.C. & Shiffrin, R.M. Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), The Psychology of Learning and Motivation, Vol. II. New York: Academic Press, 1968.

Auble, P.M. & Franks, J.J. The effects of effort toward comprehension on recall. Memory and Cognition, 1978, 6, 20-25.

Baddeley, A.D. The psychology of memory. New York: Basic Books, 1976.

Baddeley, A.D. The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. Psychological Review, 1978, 85, 139-152.

Baddeley, A.D. & Hitch, G. Working memory. In G.H. Bower (Ed.), The psychology of learning and motivation. Vol. 81, New York: Academic Press, 1974.

Barclay, J., Bransford, J., Franks, J., McCarrell, N. & Nitsch, K. Comprehension and semantic flexibility. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 471-481.

Bellezza, F.S., Cheesman, F.L. & Reddy, B.G. Organization and semantic elaboration in free recall. Journal of Experimental Psychology: Human Learning and Memory, 1977, 5, 539-550.

Bower, G.H. & Clark, M.C. Narrative stories as mediators for serial learning. Psychonomic Science, 1969, 14, 181-182.

Bjork, R.A. & Whitten, W.B. Recency-sensitive retrieval processes in long-term free recall. Cognitive Psychology, 1974, 6, 173-189.

Bobrow, S.A. & Bower, G.H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1969, 80, 455-461.

Bousfield, W.A. The occurrence of clustering in the recall of randomly arranged associates. Journal of General Psychology, 1953, 49, 229-240.

Bower, G.H. Mental imagery and associative learning. In L.W. Gregg (Ed.), Cognition in learning and memory. New York: Wiley, 1972.

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

Bower, G.H. & Winzenz, D. Comparison of associative learning strategies. Psychonomic Science, 1970, 20, 119-120.

Broadbent, D.E. Perception and communication. London: Pergamon, 1958.

Buschke, H. & Lazar, G. Cue encoding and recognition in facilitation of recall. Journal of Experimental Psychology, 1973, 97, 75-78.

Carpenter, P.A. On the comprehension, storage and retrieval of comparative sentences. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 401-411.

Cattell, J. McK. Veber die Zeit de Erkennung und Bennenung von Schriftzeichen, Bildern und Farben. Philosophice Studien, 1885, 2, 635-650. Translated and reprinted in James McKeen Cattel Man of Science, 1860-1944. Lancaster, Penn: Science Press, 1947, 1.

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

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

Craik, F.I.M. 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, 1973.

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

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.

Craik, F.I.M. and Tulving, E. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 1975, 104, 268-294.

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

Delin, P.S. Success in recall as a function of success in implementation of mnemonic instructions. Psychonomic Science, 1968, 12, 153-154.

Duffy, T.M. & Montague, W.E. Sentence mnemonics and noun-pair learning. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 157-162.

Eagle, M., & Leiter, E. Recall and recognition in intentional and incidental learning. Journal of Experimental Psychology, 1964, 68, 58-63.

Epstein, M.L., Phillips, W.D. & Johnson, S.J. Recall of related and unrelated word pairs as a function of processing level. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 149-151.

Elias, C.S. & Perfetti, C.A. Encoding task and recognition memory: The importance of semantic encoding. Journal of Experimental Psychology, 1973, 99, 151-156.

Fechner, G.T. Elemente der Psychophysik. Leipzig: Breitkop and Härtel, 1860.

Fischer, R.P. & Craik, F.I.M. Interaction between encoding and retrieval operations in cued recall. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 701-711.

Frase, L.T. & Kamman, R. Effects of search criterion upon unanticipated free recall of categorically related words. Memory and Cognition, 1974, 2, 181-184.

Gardiner, J.M. Levels of processing in word recognition and subsequent free recall. Journal of Experimental Psychology, 1974, 102, 101-105.

Glanzer, M. & Koppenaal, L. The effect of encoding tasks on free recall: stages and levels. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 21-28.

Goldman, S.R. & Pellegrino, J.W. Processing domain, encoding elaboration, and memory trace strength. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 29-43.

Graesser, A. & Mandler, G. Recognition memory for the meaning and surface structure of sentences. Journal of Experimental Psychology: Human Learning and Memory, 1975, 3, 238-248.

Herriot, P. Attributes of memory. London: Methuen, 1974.

Horowitz, L.M. and Prytulak, L.S. Redintegrative memory. Psychological Review, 1969, 76, 519-531.

Hunt, R.R. & Mitchell, D.B. Specificity in nonsemantic orienting tasks and distinctive memory traces. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 121-135.

Hyde, T.S. Differential effects of effort and type of orienting task on recall and organization of highly associated words. Journal of Experimental Psychology, 1973, 97, 111-113.

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

Hyde, T.S. & Jenkins, J.J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 471-480.

Jacoby, L.L. Encoding processes, rehearsal, and recall requirements. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 302-310.

Jacoby, L.L. The role of mental contiguity in memory: Registration and retrieval effects. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 483-496.

Jacoby, L.L., Bartz, W.H. & Evans, J.D. A functional approach to levels of processing. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 331-346.

Jacoby, L.L. & Goolkasian, P. Semantic versus acoustic coding: retention and conditions of organization. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 324-333.

James, W. Principles of psychology. New York: Holt, 1890.

Jenkins, J. Remember that old theory of memory? Well, forget it! American Psychologist, 1974, 29, 785-795.

Johnston, C.D. & Jenkins, J.J. Two more incidental tasks that differentially affect associative clustering in recall. Journal of Experimental Psychology, 1971, 89, 92-95.

Kendall, M.G. & Babington-Smith, B. On the method of paired comparisons. Biometrika, 1939, 31, 324-345.

Kintsch, W. The representation of meaning in memory. Hillsdale, N.J.: Erlbaum Associates, 1974.

Kirk, R.E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole Publishing Co., 1968.

Klein, K. & Saltz, E. Specifying the mechanisms in a levels-of-processing approach to memory. Journal of Experimental Psychology: Human Learning and Memory, 1976, 6, 671-679.

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

Levy, B.A. & Craik, F.I.M. The co-ordination of codes in short-term retention. Quarterly Journal of Experimental Psychology, 1975, 27, 33-45.

Lockhart, R.S., Craik, F.I.M. & Jacoby, L.L. Depth of processing in recognition and recall: Some aspects of a general memory system. In J. Brown (ed.), Recognition and recall. London: Wiley, 1976.

Loftus, G.R. & Loftus, E.F. Human memory: The processing of information. Hillsdale, NJ: Erlbaum Associates, 1976.

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

Mandler, G. Organization and memory. In K.W. Spence and J.T. Spence (Eds.), The psychology of learning and motivation, Vol. 1. New York: Academic Press, 1967.

Martin, E. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. Psychological Review, 1968, 75, 421-441.

McDaniel, M.A. & Masson, M.E. Long-term retention: When incidental semantic processing fails. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 270-281.

McMurray, D.W. & Duffy, T.M. Meaningfulness and pronounceability as chunking units in short-term memory. Journal of Experimental Psychology, 1972, 96, 291-296.

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.

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.

Mistler-Lachman, J. Depth of comprehension and sentence memory. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 98-106.

Mistler-Lachman, J. Queer sentences, ambiguity and levels of processing. Memory and Cognition, 1975, 3, 395-400.

Mondani, M.S., Pellegrino, J.W. & Battig, W.F. Free and cued recall as a function of different levels of word processing. Journal of Experimental Psychology, 1973, 2, 324-330.

Morris, C.D., Bransford, J.D. & Franks, J.J. Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 519-533.

Moscovitch, M. & Craik, F.I.M. Depth of processing retrieval cues and uniqueness of encoding as factors in recall. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 447-458.

Murdock, B.B., Jr. Recent developments in short-term memory. British Journal of Psychology, 1967, 58, 421-433.

Murdock, B.B., Jr. Short-term memory. In G.H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory, New York: Academic Press, 1972.

Murdock, B.B., Jr. Human memory: Theory and data. Potomac, Md.: Lawrence Erlbaum Associates, 1974.

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

Nelson, D.L., Wheeler, J.W., Jr., Borden, R.C. & Brooks, D.H. Levels of processing and cuing: Sensory versus meaning features. Journal of Experimental Psychology, 1974, 103, 971-977.

Nelson, T.O. Repetition and depth of processing. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 151-171.

Nelson, T.O. & Hill, C.C. Multiple retrieval paths and long-term retention. Journal of Experimental Psychology, 1974, 103, 185-187.

Nelson, T.O. & Vining, S.K. Effect of semantic versus structural processing on long-term retention. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 198-209.

Norman, D.A. Models of human memory. New York: Academic Press, 1970.

Norman, D.A. & Bobrow, D.G. Description: A basis for memory acquisition and retrieval. Technical Reports, No. 74, Center for Human Information Processing, University of California, 1977.

Paivio, A. Imagery and verbal processes. New York: Holt, Rinehart and Winston, 1971.

Paivio, A. A theoretical analysis of the significance of imagery in learning and memory. In P. Sheehan (Ed.), The Function and Nature of Imagery. New York: Academic Press, 1972.

Paivio, A. Imagery in recall and recognition. In J. Brown (Ed.), Recall and Recognition. London: Wiley, 1976.

Paivio, A. & Foth, D. Imaginal and verbal mediators and noun concreteness in paired-associate learning: The elusive interaction. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 384-390.

Paivio, A., Yuille, J.C. & Madigan, S.A. Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology Monograph, 1968, 76, (1, Pt. 2).

Paivio, A. & Yuille, J.C. Mediation instructions and word attributes in paired-associate learning. Psychonomic Science, 1967, 8, 65-77.

Paivio, A. & Yuille, J.C. Changes in associative strategies and paired-associate learning over trials as a function of word imagery and type of learning set. Journal of Experimental Psychology, 1969, 79, 458-463.

Parkinson, S.R., Parks, T.E. & Kroll, N.E.A. Visual and auditory short-term memory: Effects of phonemically similar auditory shadow material during the retention interval. Journal of Experimental Psychology, 1971, 87, 274-280.

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

Postman, L. Verbal learning and memory. Annual Review of Psychology, 1975, 26, 291-335.

Postman, L. Methodology of human learning. In W.K. Estes (Ed.), Handbook of Learning and Cognitive Processes. Vol. 3, Hillsdale, N.J.: Erlbaum Associates, 1976.

Postman, L. & Kruesi, E. The influence of orienting tasks on the encoding and recall of words. Journal of Verbal Learning and Verbal Behavior, 1977, 353-369.

Reed Hunt, R. & Mitchell, D.B. Specificity in non-semantic orienting tasks and distinctive memory traces. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 121-135.

Riegel, K.F. The Michigan Restricted Associate Norms (Rep. #3) Ann Arbor, University of Michigan, Department of Psychology, 1965.

Roenker, D.L., Thompson, C.P. & Brown, S.C. Comparison of measures for the estimation of clustering in free recall. Psychological Bulletin, 1971, 76, 45-48.

Rosch, E. Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 1975, 104, 192-223.

Rosenberg, S. & Schiller, W. Semantic coding and incidental sentence recall. Journal of Experimental Psychology, 1971, 90, 345-346.

Saltz, E. The cognitive bases of human learning. Homewood, Illinois: Dorsey, 1971.

Schallert, D. Improving memory for prose: The relationships between depth of processing and context. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 621-632.

Schnur, P. Testing the encoding elaboration hypothesis: The effects of exemplar ranking on recognition and recall. Memory and Cognition, 1977, 5, 666-672.

Schulman, A.I. Memory for words recently classified. Memory and Cognition, 1974, 2, 47-52.

Seamon, J.G. & Murray, P. Depth of processing in recall and recognition memory: Differential effects of stimulus meaningfulness and serial position. _Journal of Experimental Psychology: Human Learning and Memory_, 1976, _2_, 680-687.

Seamon, J.G. & Virostek, S. Memory performance and subject-defined depth of processing. _Memory and Cognition_, 1978, _6_, 283-287.

Shriffin, R.M. & Atkinson, R.C. Storage and retrieval processes in long-term memory. _Psychological Review_, 1967, _76_, 179-193.

Shulman, H.G. Encoding and retention of semantic and phonemic information in short-term memory. _Journal of Verbal Learning and Verbal Behavior_, 1970, _9_, 499-508.

Shulman, H.G. Recognition memory for targets from a scanned word list. _British Journal of Psychology_, 1971, _62_, 335-346.

Siegel, S. _Nonparametric statistics for the behavioral sciences_. New York: McGraw-Hill, 1956.

Stein, B.S. The effects of cue-target uniqueness on cued recall performance. _Memory and Cognition_, 1977, _5_, 319-322.

Stein, B.S. Depth of processing reexamined: The effects of the precision of encoding and test appropriateness. _Journal of Verbal Learning and Verbal Behavior_, 1978, _17_, 165-174.

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

Till, R.E., Cormak, D.R., & Prince, P.L. Effects of orienting tasks on sentence comprehension and cued recall. _Memory and Cognition_, 1977, _5_, 59-66.

Till, R.E., Diehl, R.L. & Jenkins, J.J. Effects of semantic and non-semantic cued orienting tasks on associative clustering in free recall. Memory and Cognition, 1975, 3, 19-23.

Till, R.E. & Jenkins, J.J. The effects of cued orienting tasks on the free recall of words. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 489-498.

Torgerson, W.S. Theory and methods of scaling. New York: Wiley, 1958.

Treisman, A.M. Strategies and models of selective attention. Psychological Review, 1969, 76, 282-299.

Treisman, A. & Tuxworth, J. Immediate and delayed recall of sentences after perceptual processing at different levels. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 38-44.

Tulving, E. The effect of alphabetical subjective organization on memorizing unrelated words. Canadian Journal of Psychology, 1962a, 16, 185-191.

Tulving, E. Subjective organization in free recall of "unrelated words. Psychological Review, 1962b, 69, 344-354.

Tulving, E. Intratrial and intertrial retention: Notes toward a theory of free recall verbal learning. Psychological Review, 1964, 71, 219-237.

Tulving, E. Theoretical issues in free recall. In T.R. Dixon and D.L. Horton (Eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

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

Tulving, E. Cue-dependent forgetting. American Scientist, 1974, 62, 74-82.

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

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

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

Tzeng, O.J.L. Positive recency effect in a delayed free recall. Journal of Verbal Learning and Verbal Behavior, 1975, 12, 436-439.

Underwood, B.J. Attributes of memory. Psychological Review, 1969, 76, 559-573.

Walsh, D.A. & Jenkins, J.J. Effects of orienting tasks on free recall of incidental learning: "Difficulty," "effort," and "process" explanations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 481-488.

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

Wicklegren, W.A. More on the long and short of memory. In D. Deutsch and J.A. Deutsch (Eds.), Short-term memory. New York: Academic Press, 1975.

Wood, G. Organizational processes and free recall. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Yuille, J.C. & Paivio, A. Imagery and verbal mediation instructions in paired-associate learning. Journal of Experimental Psychology, 1968, 78, 436-441.

APPENDICES

APPENDIX A:

Historical Review

Depth of Processing - An Historical Review

Antecedents of Levels of Processing

During the 1960's a general consensus evolved about the nature of human memory that took the form of computer based models. Most of the proposed models were described in terms of a series of stores each with its own information processing capabilities (Atkinson & Shiffrin, 1968; Broadbent, 1958; Murdoch, 1967; Neisser, 1967; Waugh & Norman, 1965). It was largely agreed that at least three types of storage systems existed: a type of sensory or iconic memory, a short term store (STS) and a long term store (LTS). It was proposed that new information enters directly into memory through sensory receptors and is first registered in a modality specific storage system where certain perceptual analyses are carried out on the sensory features of the stimulus item. Information could only be held here for very brief periods of time and was highly susceptible to decay unless specifically maintained in conscious awareness. Although the capacity of this system was thought to be large, its sole function was to register and retain the incoming information long enough for various types of coding processes to act upon the information. These processes might involve transforming and transferring selected features to the next store called the short-term store (STS) whose capacity was sharply limited (Broadbent, 1958; Miller, 1956). In fact, Murdock (1967) in his modal model felt that short term memory, if required to maintain perfect information, was perhaps limited to four items.

Information could be maintained in STS by continued attention and/or rehearsal and was thought to be predominantly acoustic (Baddeley, 1966; Conrad, 1964) or phonemic in nature (Shulman, 1971), although probably visual as well (Kroll, Parks, Parkinson, Bieber & Johnson, 1970;

Parkinson, Parks & Kroll, 1971), and possibly even semantic (Shulman, 1970, 1972). In any event, the duration of the newly formed trace was thought to be relatively short - up to thirty seconds (Shiffrin & Atkinson, 1969), and information could be lost rapidly through displacement (Waugh & Norman, 1965) or interference (Adams, 1967).

Information residing in STS, if rehearsed or reorganized in terms of existing cognitive stores, could be transferred to a more permanent form of storage with a certain probability. This store was referred to as long term memory (LTM) and its capacity was thought to be limitless.

It is not the purpose of the review section of this paper to debate the various models of human memory that flourished in the 1960's but merely to set the stage for Craik and Lockhart's (1972) highly influential paper. The interested reader is referred to Norman (1970) or Loftus and Loftus (1976) for excellent overviews of the various models and their component storage, transfer and retention characteristics.

Craik and Lockhart (1972) vigorously attacked the notion of multistore models of memory, especially the necessity for information to be transferred from one memory store to another. Instead, they postulated that much research could be better understood within the framework of their "levels of processing" view. After highlighting the limitations and inconsistencies of the "box model" approach in terms of capacity, coding, and retention characteristics they suggested that memory trace was a by-product of perceptual analysis and the retention was a positive function of the depth of stimulus elaboration. Essentially, memory was viewed as a "continuum, from the transient products of sensory analysis, to the highly durable products of semantic-associative operations" (p.676).

It was their contention that during stimulus encoding the TBR material is subjected to different levels of processing with different kinds of information being extracted at each stage. Sensory features, or surface structure details, are extracted first. At deeper levels, as a word is recognized, it may trigger associations, images or stories, based on the subject's prior experiences with that word. To the extent the analysis made contact with (and use of) the contents of existing cognitive structures the probability of recall for that word was greatly enhanced.

Whenever a new conceptual framework is erected to account, in very general terms, for a huge body of existing data related to short-term retention, the explanatory and predictive capabilities of that concept are open to scrutiny. The authors (Craik & Lockhart, 1972) of the proposition that memory was a function of depth of stimulus processing could hardly have envisaged the tremendous amount of research and literature the levels approach would generate. In the ensuing years evidence both pro and con accumulated, oftentimes necessitating revisions or modifications to the original concept as its explanatory powers were tested and retested.

Before embarking on an examination of the major highlights relating to the evolution of this new process view of memory, it may well be useful to discuss in detail some of the major papers that were precursors to Craik and Lockhart's new conception of memory. Many of the studies and most subsequent studies involved the use of the incidental learning paradigm where the TBR material remains constant and the experimenter manipulates various task variables. Accordingly, to marshall support for the importance of the type of processing that goes on at input, this review begins by looking at the early incidental learning studies and the revelation that retention varied as a function of the nature of the orienting task and in some instances was equal to intentional learning. From there the paper attempts to trace the various elaborations of the original Craik and Lockhart (1972) claim that short-term retention is best understood in terms of the type of processing carried out on incoming information with deeper, semantic processing being associated with richer, longer lasting traces.

Incidental Learning Studies

To lend support to their thesis, Craik and Lockhart reviewed a series of studies involving the incidental learning paradigm wherein the experimenter presumably has control over the type of processing the subject applies to the material, a control that does not exist when the subject is merely instructed to learn. The crucial factor then is not the intent to learn per se but how the TBR item is processed. For example, Eagle and Leiter (1964) had found superior recognition for incidental learners over intentional learners where the incidental task had called for subjects to classify words in terms of their syntactic features (e.g., noun, verb, adverb and adjective). This finding has since been replicated many times (Craik & Tulving, 1975; Hyde & Jenkins, 1969, 1973; Seamon & Murray, 1976).

To illustrate, a number of studies by Jenkins and his associates indicated that items processed in terms of their semantic characteristics were recalled as well as an intentional learning control group and substantially superior to other incidental learning groups whose tasks involved processing items in terms of their physical attributes

(Hyde & Jenkins, 1969; Johnston & Jenkins, 1971). Hyde and Jenkins (1969) examined the effects of three different orienting tasks on the recall of an auditorally-presented list of 24 words. The three orienting tasks involved:

a) judging the pleasantness of each stimulus word
b) deciding whether each word in the list contained the letter "e"
c) estimating the number of letters in each word

Recall was greatest for those subjects engaged in the pleasantness judging task, even in relation to the intentional learning group. Superior recall was attributed to the fact that the subjects were required to view each word as a meaningful unit, whereas in other tasks it required the subjects to break each word into its component letters or elements. In a follow-up study Johnston and Jenkins (1971) used rhyming as an orienting task in an attempt to have subjects respond to each word as a whole unit as was the case for the pleasantness judging group in the Hyde and Jenkins study. The stimuli again were twelve highly associated word pairs and subjects in various orienting tasks were required to either write a rhyming word for each word in the stimulus list or, as in the case of the semantic orienting task, instructed to supply appropriate adjective modifiers for given nouns or nouns to be modified by given adjectives. Both these tasks (supplying nouns or supplying modifiers) led to higher recall and clustering as compared to control subjects who had performed no orienting task but had been given standard instructions to learn. The control group and the adjective-noun groups had superior recall to the rhyming groups.

In both studies the organization scores (associative clustering or items per category (IPC)) for subjects in the semantic processing group were greater as well, suggesting that the semantic relationships between the stimulus words are available for the organization of the subject's recall. The comparative effects of organization vis à vis type of processing will be taken up in more detail in the following section.

Following Craik and Lockhart's paper subsequent studies attempted to control the level and type of encoding in an effort to further explore the claim that memory for an item is directly related to the amount or depth of perceptual processing of the TBR item.

Craik (1973) described a series of five studies in which words were analyzed to different and presumably deeper levels as evidenced by increasing levels of recall. In one of the studies (Experiment (V) carried out in Craik's laboratory by Karl Egner, depth of processing was manipulated in a within subject design by asking 20 subjects one of the following questions prior to a brief tachistoscopic presentation of a word:

1) Is there a word present?
2) Is the word in capital letters/lower case?
3) Does the word rhyme with ________?
4) Is the word a member of the following category? __________
5) Does the word fit into the following sentence? __________

It was found that in addition to subjects requiring progressively more time to make a decision with each type of question, on a subsequent recognition test performance was a function of the initial processing depth.

Hyde and Jenkins (1973) examined a broad range of tasks within the incidental learning paradigm. Two new syntactic tasks were devised, in contrast to the semantic and orthographic tasks used previously. Specifically the tasks were a) determining the part of speech of the TBR words or b) deciding whether each TBR word did or did not fit into a particular sentence frame. The new syntactic tasks maintained the virtue of requiring the subject to attend to each TBR word and to execute a judgment with respect to the word as a whole. On the surface at least, these tasks resembled the semantic rating tasks and yet were quite distinct from the orthographic tasks used in earlier studies that merely called for the estimation of word lengths or required the detection of particular letters (and in which it might be argued that subjects could perform the orienting task without addressing the word as a whole). Hyde and Jenkins also included two semantic tasks requiring subjects to rate words for frequency of usage or rate them on a "pleasant-unpleasant" dimension. The so-called "non-semantic" task was checking the TBR words for the occurrence of the letters "E" or "G". In this experiment Hyde and Jenkins employed both associated words as well as semantically unrelated words. The results indicated that the semantic orienting tasks produced better recall in relation to the non-semantic tasks for the list of associated words than they did on the unrelated word list i.e. the effects of semantic orienting tasks were quite good with unrelated words (and better than non-semantic tasks) but the task effect was considerably enhanced with the list of associated words. Non-semantic tasks showed the same level of recall for both lists. Interestingly the pleasant-unpleasant rating produced the highest recall, superior even to the control group. Sentence-frame tasks produced the lowest recall. Semantic processing tasks also led to higher degrees of associative clustering than either the control condition or the non-semantic tasks.

One might question Hyde and Jenkins' classification of orienting tasks as semantic or non-semantic. Postman (1975) argues that while checking for the presence of particular letters within a word need not involve processing meaning attributes of the word, it is difficult to draw the same conclusions with respect to tasks requiring the identification of parts of speech or judgments of the appropriateness of the TBR word for a given sentence frame. Similarly, Postman claims that not very much thinking about the meaning of the word is necessary to decide its degree of familiarity and hence he questions the classification of judging the frequency of a word as a task involving deep semantic processing. To the present author the distinction between semantic and non-semantic encoding should be made on the following basis: semantic processing involves attending to both connotative (emotional, affective) and denotative (referential, lexical) attributes of the word (Herriot, 1974). Non-semantic processing involves graphemic or phonemic features of the stimulus. Hence it is suggested that judging a word's part of speech or whether it fits a given sentence frame involves processing of a semantic nature. Judgments relating to frequency of occurrence in the language are similarly placed in this category.

All of the above studies suggest that regardless of instructions to learn, retention is substantially determined by whether or not the subject engaged in semantic processing. When the orienting task was non-semantic, retention was considerably lower.

Hyde (1973) tried to answer the question of why semantic processing leads to better retention. Does it require more effort? Is it a function of the difficulty of the task or the amount of processing time required? Hyde varied the degree of processing difficulty within two general classes of orienting tasks. Because the aim of this experiment was to examine the relative effects of effort and type of task, only the incidental condition was utilized. One of the semantic groups rated the words (twelve medium associativelyrelated word pairs) as to their "pleasantness" or "unpleasantness" on a simple five-point scale. The second group rated the words on an "active-passive" dimension while another semantic group rated words on both semantic attributes. This last condition ostensibly required more effort than rating either attribute alone. Within the non-semantic groups subjects were either asked to search for the letter "e" in each of the words or were given the apparently more effortful task of noting the occurrence of both "e's" and "g's" in the stimulus words. One final group performed both semantic and non-semantic tasks on the stimulus words. Results showed that within each condition (semantic versus non-semantic) the difficulty of the orienting task made little difference in terms of either recall or associative clustering. Those subjects, however, engaged in non-semantic tasks showed both reduced recall and reduced clustering, thus adding substance to Hyde's claim that the degree of difficulty of the orienting task has little effect on either recall or associative clustering. It should be pointed out however, that the double task (pleasant-unpleasant; active-passive), was almost impossible to complete within the 5 seconds allotted for all tasks. Although, as we shall see, study time is not a crucial determinant of recall in and of itself, it might be argued that the effectiveness of rating a word along two semantic dimensions was not fairly evaluated since there was insufficient time to perform the relevant analyses.

Another study (Walsh & Jenkins, 1973), attempting to account for the superiority of semantic processing in memory performance, examined the hypothesis that orienting tasks which require considerable time and effort might lead to more interaction with the material and, hence, to greater learning than a task that was less demanding. The alternative hypothesis was that retention would be a function of the level of processing, with deeper levels leading to better recall. Hence, the "process" hypothesis was more concerned with the intrinsic nature of the task. For example, Hyde and Jenkins (1969, 1973) and Johnston and Jenkins (1971) found that tasks requiring the subjects to process words as meaningful units yielded higher levels of recall than tasks that required the subjects to process the orthographic, phonemic or syntactic aspects of the words. These earlier studies however, had provided no direct evidence bearing on the "difficulty" or "effort" hypothesis. Conceivably, the semantic tasks could simply require less time or more effort than the non-semantic tasks. Again, the three different orienting tasks employed were "pleasant-unpleasant," the "E" or "G" task (i.e., indicating whether the words involved "e's" or "g's") and estimating the number of syllables contained in each word. Words were read at the rate of one every four seconds and subjects engaged in one or more processing tasks. The pattern of results showed that retention of a list of 24 low-frequency English words was superior for groups employing the semantic processing task either singly or in combination with a non-semantic task over groups who employed non-semantic tasks either singly or in combination. In fact it was found that any combination of tasks used seemed to produce recall similar to that one would expect using the best of the combined tasks alone. Hence, it was concluded that it is the qualitative nature (or type of processing) associated with the task that accounted for the observed differences in recall in this as well as earlier studies cited above. Experiments II and III found essentially the same results.

Factors Underlying Depth of Processing

The Hyde (1973) and Walsh and Jenkins' (1973) studies were the first serious attempt to systematically delineate the variables that might underlie the concept of depth of processing. In the ensuing years a variety of explanations were advanced to account for the now reliable advantage of semantic over non-semantic orienting tasks on subsequent recall. The rest of the review section of this paper traces some of the major developments that have led to an elaboration of the original Craik and Lockhart notion. Why, for example, do certain classes of orienting tasks consistently yield higher recall than others? What are the factors that correlate most highly with the concept of depth? The order of discussion of each factor mirrors in part the chronological development of this concept and is therefore not meant to indicate the relative importance of each factor. The conclusion of the review section attempts to draw together some of the more potent variables. It is thought that these include some of the following: clustering at recall of associatedly related list items; degree of secondary organization; time required for processing; uniqueness or specificity of encoding of the TBR item; degree of trace elaboration in terms of the number and nature of features encoded at input; the interaction of variables such as level of meaningfulness with depth of processing; encoding and retrieval cues; amount of cognitive effort or conscious attention involved in the extraction of richer features of the TBR material.

Organization and Depth of Processing. This initial section questions whether various retention phenomena (e.g. free recall) are attributable to organizational variables or depth of processing variables. This is a most complex question. Although historically organizational theories precede depth of processing their origins were plagued with the same problems (relating to definition and clarification of the concept) as beset depth of processing today. Also the methods by which organization has been measured are varied and sometimes questionable. The early work (Bousfield, 1953; Cofer 1965) examined clustering (associative and categorical) in free recall and this was extended to non-categorized lists by Tulving (1962, 1964) with his emphasis on subjective units of organization. Mandler (1967, 1972) explored the relationship between number of categories and recall, concluding that total recall is a function of the number of categories used in sorting. Bower (1972) demonstrated large empirical effects which can be attributed to organizational processes. For example, Bower, Clark, Lesgold and Winzenz (1969) showed that recall was greatly facilitated if the TBR items were pre-organized for subjects in terms of their conceptual hierarchies.

Some of the empirical effects claimed by organization theorists, however, may be open to interpretation in quite different terms. For example, Lockhart, Craik and Jacoby (1976) suggest that sorting items into categories might involve deep-level processing which in turn leads to a more distinctive memory trace. Certainly the data of Wood (1972) support this notion.

To refocus on the original question, however, one might wonder in what ways the concept of organization and depth of processing are similar and also, how do they differ. Mandler (1967) presented a strong case that the availability of TBR words is not merely facilitated by the subject organizing the TBR list presented but that organization is a necessary condition for memory. What should be stressed is that the subject organizes the list at input in terms of intra-list associations, i.e., words may be semantically, phonemically or structurally related and hence are organized in such a way that these relationships facilitate recall.

The concept of organization is contrasted with Tulving's (1972) notion of encoding where a word is related to, and perhaps integrated with, the contents of semantic memory independently of other list words presented. Words can be encoded qua words (Craik & Tulving, 1975) or in terms of their extra-list associations and this is able to produce good recall. The suggestion is that elaborate stimulus processing leads to a new association - an association between the TBR item and information currently stored in memory. Whether this point of view represents a radical departure from organizational processes remains to be seen.

Certainly it seems incorrect to talk of semantic processing and organization as if they are mutually exclusive strategies. The results of the studies discussed earlier (Hyde & Jenkins, 1969, 1973; Hyde, 1973; Johnston & Jenkins, 1971) demonstrated that incidental learning tasks that orient subjects to the semantic aspects of words significantly affect both level of recall and degree of secondary organization (as measured by associative clustering). Several other experiments (Till & Jenkins, 1973; Till, Diehl & Jenkins, 1975) have provided additional evidence to suggest that subjects organize their recall if task demands permit. Till and Jenkins' (1973) independently manipulated associative structure and task structure to see the effects of task variables on organization of recall. In Experiment I, using unrelated words, there were no differences in amount of clustering for recalled words with respect to each type of orienting task. In Experiment II a list of high strength associative pairs was used. In one group task activities were made compatible with the associative structure of the list i.e., tasks were assigned so that members of each associative pair were processed with the same task. In another group the activities of the subject were in opposition to the associative structure. The results indicated reduced call for the different tasks condition, suggesting a strong dependence of recall on semantic organization in storage. To summarize, although associative clustering was greater than chance in both conditions, it was effectively reduced in the different task condition by the experimental manipulation of assigning different tasks to each member of the associative pair.

Subsequently, Till, Diehl, and Jenkins (1975) replicated the effects of the Till and Jenkins' study by presenting members of associative pairs in non-adjacent positions followed by orienting tasks that were again the same (both semantic or both non-semantic) or different for each member of the pair. The results showed that irrespective of the task, percentage of clustering was greater for pairs processed with the same task. The authors tentatively concluded that clustering is not automatic and may in fact require facilitating by the way in which the experimenter manipulates the task demands and the manner in which they are assigned to the associative pair members (p. 22). An alternative explanation to the above findings and one compatible with the current status of depth of processing is that related items that cluster at recall simply indicates that the initially recalled member serves as an efficient retrieval cue for the other (Horowitz & Prytulak, 1969). The effectiveness of each recalled word as a retrieval cue for its pair member was more potent following semantic than non-semantic processing. Jacoby and Goolkasian (1973) found essentially the same results for semantically related versus acoustically related pairs.

Bellezza, Cheesman and Reddy (1977) explored the relative effectiveness of thematic organization and semantic processing as determinants of free recall by attempting to manipulate semantic processing and organization independently. Subjects encoded alternate TBR words by either giving a meaningful sentence using the word ("Remember Words") or thematically organizing words ("Story Words") (Bower & Clark, 1969). Additionally, degree of semantic elaboration was manipulated by having subjects generate sentences of differing lengths on the assumption that longer sentences would provide more elaborate or unique encodings for each TBR words (Craik & Tulving, 1975, Experiment 7). Although there was no stronger input-output correlation for words woven into a story than "remember" words, as expected more "story" words were recalled than "remember" words. The expected effect of a greater degree of semantic elaboration associated with increasing sentence length did not materialize. Experiment 3 was a further attempt to demonstrate that organization is facilitative of recall beyond the effectiveness of semantically elaborating each TBR item. Both groups were instructed to form a visual image of each word and rate it on a 7-point scale of difficulty. The organization group (using Tulving's 1962 alphabetic mnemonic technique) was encouraged to form a composite visual image between the previously formed visual image and some visual image of the beginning letter (e.g., the letter B may be thought of as a fat lady). The definition group produced a short definition for each word following the rating task. Not surprisingly free recall was superior for the mnemonic group. Additionally, three measures of organization were used: (a) input-output correlation, (b) the correlation between order of recall with recalled words ranked in alphabetical order, (c) Roenker, Thompson and Brown's (1971) adjusted ratio of clustering (ARC) score was computed with categories defined as words beginning with the same letter. The only significant difference between the two groups was on the third measure. The conclusion of the present study that "semantic processing is necessary but not sufficient for optimal recall" (p. 542), however, remains suspect especially since the mnemonic group received additional instructions to form composite images at input, a strategy that is known to have potent effects on recall even beyond the normal imagery instructions (Bower, 1972). Furthermore, it is contended that both groups were engaged in semantic processing and it is highly doubtful that organization and semantic elaboration were manipulated independently.

Yet the overall pattern remains equivocal. Tversky (1973) found that instructing subjects to search for relationships among a list of to-be-presented words and to group associated words together had a faciliative effect on recall. This was particularly so when subjects received these instructions prior to list presentation as opposed to just prior to recall. Further, Jacoby (1974) suggested that requiring subjects to look back through memory of a list of recently presented items for categorically related items led to the formation of an implicit contiguity which facilitated recall. It was argued that "with semantically related items...the meaning that is encoded is a product of the interaction among items rather than the meaning of individual items" (p. 490). On the surface, these findings appear to lend support to the Bellezza, Cheesman and Reddy position. In Jacoby's study however, neither physical or implicit contiguity of acoustically related items was shown to be an effective determinant of recall. The effect was strongest when the items were processed semantically (i.e., categorically related).

In summarizing this section, it is worthwhile to re-examine the notion of organization. One thing appears clear. Organization is based on common semantic properties, so that to say subjects have organized a list of TBR items is to say they have processed them in some semantic way. In that case, levels of processing researchers have no trouble handling these data. Lockhart, Craik and Jacoby (1976) concede that where organization leads to increased depth of processing the resultant memory trace is more distinctive and hence retention is likely to be enhanced. It is suggested that during the formation of the episodic memory trace, where organization involves seeking out relationships in semantic memory between the TBR items the two processes (organization and depth of processing) are probably co-extensive.

Time. Even the most recent formulations of the mechanisms that are thought to underlie depth of processing, i.e., the concept of conscious attention or effort as proposed by Lockhart, Craik and Jacoby (1976), have to consider the possibility that time may be a correlate of depth. Conscious attention requires time. As mentioned earlier in this review, Craik (1973) tentatively suggested that deeper levels of analysis require more time. Craik and Tulving (1975) entertained this notion more seriously in their series of studies. Usually questions of three kinds were asked of subjects prior to presentation of a word. These questions required decisions about physical, phonemic, and semantic attributes of the words. The questions were: a) Is the word in capital letters? b) Does the word rhyme with ---? and c) "Is the word in the category ______?" (or "would the word fit in the sentence ________?"). Results were consistent in Experiments 1-4 revealing that both decision latencies and subsequent performance on tests of recognition or recall increased as a monotonic function of level of processing, i.e., as the questions moved from structural to phonemic to semantic levels. Hence there seemed to be strong indications that processing time might serve as an independent index of depth. However, subsequent re-examination of the data from Experiment 2 showed that this relationship was not critical. By looking for the recognition levels for subsets of words whose decision latencies were appreciably different it was found that the slowest responses led to no higher retention as measured on a recognition test than that subset containing the shortest decision latencies for each task. Subsequently, it was found (in Experiment 5) that a shallow (but difficult) task led to poorer retention than deeper (but easier) tasks despite the latter taking less time for processing.

Gardiner (1974) also examined the relationship between initial processing time and subsequent recall to see what support there was for time as an independent criterion of depth. Subjects were told to look for a target word in a series of to-be-presented words. In each case the target was defined semantically in terms of one (the target is a country) or two attributes (the target is a European country) or phonetically, again with one attribute (the target contains the sound /l/) or two attributes (the target contains the sounds /l/ and /n/). In each case the decision latency for the response to each target word was recorded. Subjects took longer to process the target words phonetically than in terms of semantic attributes and yet recall for semantically processed targets was superior to target processed in terms of phonetic attributes.

However, in the Gardiner study, it could be argued that subjects were not selectively attending to the whole word but identifiying target words on the basis of constituent phonemes. An earlier study by Mondani, Pellegrino and Battig (1973) however, required subjects to attend to a whole word (e.g., CHERRY) prior to either orthographic processing (choose the correct response from a set of alternatives where in all but one there was a letter not in the word e.g., CJUL, CXED, CHEP, CREH) or taxonomic processing (where the alternatives were DWELLING, ANIMALS, SPORTS, FRUITS). These authors found that, while orthographic processing took longer than taxonomic processing, recall was superior for words processed taxonomically. Jacoby and Goolkasian (1973, Experiment I) reported that while acoustic processing took less time than category processing, recall for the latter was superior. Seamon and Murray (1976) also found a negative correlation between decision time and recall as a function of type of processing and level of meaningfulness.

More recently, Goldman and Pellegrino (1977) added further evidence to the claim that retention differences are not simply a function of total processing time (providing one accepts reaction time as an indicator of level of processing). Specifically, they found that the latency for decisions about acoustic versus orthographic attributes of words was approximately the same (e.g., 610 msec vs. 630 msec.) although these tasks led to differing levels of recall. Further they found that negative decisions took longer than positive decisions but positive ones were better recalled (This phenomenon of the superiority of 'yes' over 'no' decisions will be taken up in more detail later in this review).

In summary it may be said that the duration of the encoding process may not be the critical factor for registering events in memory. The implication (to be tested in the present thesis) is that processing time may only be important to the extent that it permits deeper levels of analysis to be carried out.

Uniqueness or Distinctiveness of Encoding. It has long been accepted that decreased similarity among list items implies less retroactive interference at recall (Adams, 1967). In the discussion so far we have noted a consistent advantage for items encoded in terms of their semantic features as opposed to items processed in terms of phonemic, orthographic or other structural characteristics. Certainly encoding at the lower structure levels (i.e., checking a list of words for a "Es" or "Gs") does not lead to a particularly distinctive stimulus encoding. Perhaps the laws of interference could be invoked to explain the obtained retention differences. Semantically encoded items are likely to be more distinctive, since in a list of unrelated words there is not much likelihood of a severe overlap in terms of their meaning attributes. On the other hand, many of the so-called non-semantic tasks described so far have entailed redundant encoding to the extent that the same encoding question is asked of successive list items. The effect of this is to increase the probability of interference of similarly encoded items of recall.

Moscovitch and Craik (1976) explored this possibility, i.e., that the levels of processing notion is in accord with an encoding distinctiveness hypothesis. It was hypothesized that deeper analyses set the stage for more unique or distinctive encodings in memory. Before being exposed to a TBR word for 200 msec subjects were primed with one of the following questions:

(a) Does it rhyme with ______________?

(b) Is it in the category ____________?

(c) Does it fit in the sentence ___________?

On the basis of Craik's (1973) higher recognition levels for sentences frames over categories, the authors predicted retention would be a < b < c. It was suspected that sentence frames would provide a more unique encoding context than either of the other types of questions. As speculated earlier, shallow encodings having to do with typescript or rhyme, provide relatively redundant rather than specific encoding contexts for the TBR items. The distinctiveness of category names is also questionable, especially as the number of TBR items associated with a particular category increases (Tulving & Pearlstone, 1966). Since half of the subjects were supplied with the encoding questions as retrieval cues at recall there was also some interest as to whether the expected effect due to levels would be attenuated by the aid of a retrieval cue at recall.

The now reliable effect due to levels was obtained i.e., retention was in the order a < b < c. Providing subjects with the original questions as retrieval cues magnified the effects due to level of processing especially for the deeper encodings. This interaction of encoding and retrieval factors in cued recall will be dealt with more fully in the ensuing section.

To summarize the present study, leaving aside for the moment the cueing effects, it appears that uniqueness of encoding is important but only if the item is processed substantively (i.e., in terms of its underlying, deep features). Surface-level uniqueness, in and of itself, does not guarantee the memorability of an item. To clarify the distinction, consider Experiment III wherein semantic processing for a set of 6 TBR words were encouraged in terms of encoding questions that were unique in terms of surface structure (i.e., Is it a garment? Is it a form of apparel? Is it clothing? Can you wear it?) but obviously substantively similar (i.e., they all meant the same thing). In the shared condition all questions, relating to the same category of TBR words were the same. Since there was no differences between these two conditions in levels of cued recall it was concluded that for semantic processing formal or surface structure uniqueness was not conducive to greater retention. In the third condition in this study each TBR word had its own unique encoding question, in both form and substance. Recall was markedly superior for this condition only when the items were processed semantically.

One is tempted to conclude from the results of Experiment III that even under conditions of free recall deeper levels of processing presumably would create more unique or accessible memory traces and hence facilitate retention. Certainly in the Moscovitch and Craik study, under conditions of cued recall, encodings of TBR words had to be semantically distinctive before recall was enhanced. Lockhart, Craik and Jacoby (1976) underlined this point by speculating that phonemic encodings are somehow not as distinctive or unique as semantic encodings. Certainly the range of phonemic attributes is smaller and more redundant for a given list of TBR words than the set of semantic attributes that may be tapped. Being more discriminable it is understandable that tasks requiring analyses of these deep semantic features would provide a more distinctive encoding that would enhance retrieval of the list items.

In a recent study, Reed Hunt and Mitchell (1978) attempted to extend the distinctiveness hypothesis to several non-semantic domains (i.e., phonemic and orthographic) by arguing that "if the distinctiveness of the episodic memory trace is an important determinant of memory, orienting tasks drawing attention to highly specific, non-semantic features also should facilitate recall" (p. 122). These authors claimed that Moscovitch and Craik's failure to find a facilitative effect with unique rhyme encodings on recall may have been a function of the particular list of words used, which potentially rhymed with large numbers of other words. These authors, using low-rhyme words (e.g., devil, truth), obtained better retention in free recall compared to words that rhymed with large numbers of words (e.g., meat, pride). Within the orthographic domain, too, recall was better following processing of orthographically distinct words (e.g., lymph, gnaw) versus orthographically common words (e.g., harp, kennel).

Craik and Tulving (1975, Experiment 8) also found that when rhyme encodings were made more unique (by reducing the set size to be encoded to 4 in a set of 60 TBR words) their recall was enhanced to the point where they were equivalent to recall of 40 (out of 60) categorically encoded words. This uniqueness effect was not, however, obtained for case-encoded words or for the category group where recognition also did not vary substantially as a function of set size. The overall conclusions therefore remained tentative in this study as to explanatory powers of uniqueness of encoding.

Elaboration and Depth of Processing. The Craik and Tulving (1975) series of experiments began as an attempt to support the original Craik and Lockhart (1972) position that memory for an event is dependent upon the nature and number of perceptual analyses carried out at input. Accordingly, they embarked on a series of 10 experiments to fully explore the notion of depth of processing. These experiments will not be dealt with *in toto* at the present time but will be referred to in various sections to follow in order to bolster arguments or provide counter points of view to the extant mechanisms postulated to underlie the level of processing effect. The results of Experiments 1-4 replicated the earlier findings (cited above) that showed superior retention in both recognition and recall for items processed in terms of their semantic features irrespective of whether subjects were given incidental or intentional learning instructions. The authors were somewhat perturbed at the consistency of their findings even under intentional learning conditions since it was thought that irrespective of the orienting task subjects would engage in some semantic analyses thus tending to equalize retention across conditions.

The notion of 'depth' was retained since it connoted greater degrees of semantic analysis but the concept of amount of stimulus elaboration or spread of encoding was introduced to better characterize the empirical findings. This modification to the original depth notion was also invoked to explain the discrepancies in recall associated with Yes and No responses. The results of Experiments 1-4 had consistently yielded higher retention levels for encoding questions requiring a positive response. It was suggested that in the case of a positive response the TBR item and the encoding context formed a more coherent or elaborate unit since it was well integrated or congruent (Schulman, 1974). Although not entirely relevant to this section the Schulman study merits some discussion since the findings have been increasingly utilized by levels researchers to account for consistently higher recall patterns associated with positive decisions (e.g. Is a cork screw an opener?) as opposed to negative decisions (e.g. is spinach ecstatic?) The effect is quite pervasive having been obtained for both free and cued recall as well as recognition. It was Schulman's contention that congruous encodings foster better relational encoding. This explanation seems unsatisfactory especially since both encodings involve semantic processing.

The present author feels that congruous encodings entail processing more relevant (sometimes definitive) attributes of the TBR item. At retrieval greater use can be made of existing cognitive information in reconstructing the original encoding. This strategy is not available for the recovery of incongruities (i.e., we do not normally think of spinach as being ecstatic.)

To summarize, effective stimulus elaboration involves the analysis of attributes that are both unique with respect to TBR word and salient insofar as they form a connection or congruency between the task demands, the encoding context, the TBR word, and additionally, the contents of semantic memory.

These results might as easily be interpreted using the notion of uniqueness. Uniqueness as we have seen has no small part to play in accounting for the results of depth of processing studies (Moscovitch & Craik, 1976; Reed Hunt & Mitchell, 1978). Furthermore, we have seen that subjects instructed to form bizarre or unusual images between a pair of words are able to subsequently recall significantly more words than subjects given standard imagery instructions (Delin, 1968). Certainly incongruous and distinctive encodings are not mutually exclusive terms. Craik and Tulving themselves were able to wipe out the Yes/No differences by encouraging equally elaborate encodings in both cases. In Experiment 6 eight dimensions (size, length, width, height, weight, temperature, sharpness and value) were used and subjects were asked to respond "Yes" or "No" depending on whether the TBR word was "greater than" or "less than" a reference object on one of eight dimensions. In a cued recall procedure (where the original question was restated) subjects recalled both sets of TBR words equally well suggesting that original encodings were equally elaborate, irrespective of whether the decision called for was positive or negative.

Further consideration of degree of elaboration revealed that no simple explanation such as depth of processing could account for the total range of memory phenomena. To illustrate, Craik and Tulving (1975, Experiment 7) manipulated degree of elaboration within the semantic domain. Sentential contexts in which the TBR word was to be embedded, were designated "simple," "medium," and "complex" primarily in terms of the length of the sentence frame. Subjects in an incidental learning study responded "Yes" or "No" if the TBR word fitted the sentence frame. Results strongly supported the expectation that free recall levels would be higher as the degree of elaboration of encoding context increased. What is important to note is that the effect due to sentence complexity was magnified in cued recall when subjects were given the original sentence frames in an attempt to boost recall. These results were nicely replicated by the cued recall aspects of the Moscovitch and Craik (1976) study.

It has become clear to levels researchers that the qualitative nature of the mental operations during encoding has a decisive impact on the probability of recall. Of more importance, is the fact that the memorial consequences of various elaborative encoding strategies can be enhanced considerably by reinstating the original encoding context at recall, i.e., by providing cues at recall that had been part of the original encoding environment. Tulving (1974) had stressed the importance of certain cues available to the subject at recall and the way in which they might interact with, and assist in, the retrieval of the episodic memory trace. This interaction is the subject of the next section.

Levels of Processing and Cueing. While the beneficial effects of cued over non-cued recall has long been established (Mandler, 1967; Tulving & Pearlstone, 1966) only recently have researchers investigated the effects of cued recall as a function of different levels of processing. Mondani, Pellegrino and Battig (1973) in addition to demonstrating a superiority of taxonomic processing over orthographic processing showed that supplying cues after free recall was beneficial for the taxonomic processing group only.

Nelson, Wheeler, Borden and Brooks (1974) examined the combined effects of levels of processing with retrieval cue effectiveness in a cued recall study where presentation of 24 pairs of words (cue and target) was fast (1.2 sec.) or slow (3.0 sec.). Cues were either weak or strong in terms of their relationship to the TBR word and were either a rhyme or a synonym. Subjects were informed that the word on the left side of the pair was a cue that either "rhymed with" or "has a meaning similar to" the TBR word on the right. Cues were available either at study and test or at test only. Note that in the latter case, the relationship of the retrieval cue to the TBR word was explained only after list presentation.

The most notable outcomes of this study were that a) best cued recall was for synonyms when they were available at study and test for the slow rate b) providing cues during both study and test did not enhance recall for rhymes but strongly facilitated recall for synonyms. The relatively poor recall for synonyms where the presentation rate was fast and the cue present only at test is interpretable in light of a levels of processing viewpoint and also in terms of the encoding specificity principle (Tulving & Thomson, 1973). In terms of the former, it had been hypothesized that synonym recall might be less effective than rhymes at the fast rate since there would be insufficient time available to achieve the degree of elaboration or semantic analysis the task required. As stated earlier while time per se may not be an index of depth of processing, it is nevertheless important to the extent that it permits relevant analyses to be carried out on the stimulus material.

In terms of the encoding specificity principle the Nelson, Wheeler, Borden and Brooks study confirms the importance of the interaction between encoding and retrieval operations in cued recall when attempting to provide a complete account of memory processes. The principle which grew out of earlier work by Tulving and Pearlstone (1966) and Tulving and Thomson (1973) states that "specific encoding operations performed on what is perceived determine what is stored, [implying that different types of processing produce different types of memory traces], and what is stored determines what retrieval cues are effective in providing access to what is stored" (Tulving & Thomson, 1973, p. 369).

The interactive effects of these two variables has since been replicated many times (Craik & Tulving, 1975; Experiments 6 and 7 (for positive responses only); Fisher & Craik, 1977; Moscovitch & Craik, 1976) especially when the cue - target relationship is unique (Stein, 1977). Moscovitch and Craik (1976) had also demonstrated the importance of uniqueness of the retrieval cue as a facilitator of recall but this had been accomplished by varying the number of TBR words paired with each cue. Stein, however, sought to extend the generality of the concept by manipulating the uniqueness of the cue - target relationship by varying the encoding instructions to subjects. Specifically, subjects rated two unrelated words as "same" or "different" on a "hard-soft" dimension or rated the simile "an X is like a Y" as "easy" or "difficult." The hypothesis that the more unique (or elaborate) encoding established in the context of rating the word pairs as similes would subsequently lead to a higher level of effectiveness for one of the pair as a retrieval-cue at recall was confirmed in Experiment 1 for cued recall and replicated in Experiment 2 using a pair-recognition task. The results of this study were claimed to be more supportive of the encoding specificity principle than the levels of processing approach since it was subsequently shown that recognition (Experiment 3) and recall (Experiment 4) was equivalent for both types of encoding in the absence of any retrieval cues. Intuitively, both tasks appear to require relatively deep processing and hence the fact that both were recalled or recognized equally well is not a source of great concern for the levels of processing viewpoint.

Although the present review is concerned solely with the role of various encoding strategies and their memorial consequences for words, it is interesting to note that orienting tasks, especially those requiring semantic processing have been demonstrated to interact with retrieval cues to increase the magnitude of the difference in recall over non-semantic tasks both at the level of sentences (Barclay, Bransford, Franks, McCarrel & Nitsch, 1974; Till, Cormack & Prince, 1977) and paragraphs (Schallert, 1976). Mistler-Lachman (1974, 1975) in a series of studies also demonstrated the advantages of processing connected discourse for meaning. These findings at the supra word level are seen to demonstrate the pervasiveness and power of the concept of semantic processing.

Although most of the experiments discussed so far lend support to the levels of processing viewpoint a number of researchers were developing serious reservations about the interpretations of some of these studies. For example, Arbuckle and Katz (1976), although providing support at the level of their data for semantic over non-semantic processing, questioned whether the different types of orienting tasks really lead to the formation of qualitatively different memory traces. These authors suggested that "non-semantic orienting tasks may simply provide a less efficient means for encoding semantic information" (p. 362). Their argument had its origins in the failure of Walsh and Jenkins (1973) to find any evidence for addivity effects when semantic and non-semantic tasks were used in combination. As Postman (1976, p. 38) points out, however, it is inconceivable that recall remains unchanged irrespective of the number of tasks imposed on the subject. Furthermore, recent studies (Klein & Saltz, 1976; Goldman & Pellegrino, 1977) have marshalled support for the additivity effects of various encoding tasks. Arbuckle and Katz (1976) also overlooked the Levy and Craik (1975) study that demonstrated that semantic and acoustic codes combined additively to facilitate recall.

To return to the original point, it must be restated that Arbuckle and Katz (1976) and subsequent researchers (Morris, Bransford & Franks, 1977; Stein, 1978) had serious misgivings about the dichotomous nature of semantic versus non-semantic tasks and especially the appropriateness of the test for the TBR item given a particular mode of encoding. They argued that the type of test ought to be compatible with the original acquisition activity. Morris, Bransford and Franks presented subjects with sentence frames e.g., "The __________ had a silver engine" followed by a TBR word that was either compatible semantically or not (e.g., TRAIN vs. EAGLE) or a sentence frame e.g. "_________ rhymes with legal" followed by a TBR word that required a "Yes" or "No" in terms of its compatibility with the sentence frame (e.g., EAGLE vs. PEACH). Following acquisition subjects were given either a standard recognition test or a rhyming-recognition test (where the foil given in the recognition test (e.g., REGAL) rhymed with the one given in the acquisition (e.g., LEGAL) and both rhymed with the TBR word (e.g., EAGLE).

While the standard recognition test supported the levels of processing claim that greater depth of processing leads to better memory performance the rhyming recognition test revealed a different pattern of results. Rhyme acquisition was shown to be superior to semantic acquisition given a rhyming recognition test but only for 'Yes' decisions. This was accounted for by invoking Schulman's (1974) explanation of incongruity for the 'No' decisions.

The claim that earlier findings supporting the levels of processing notion were caused by a mismatch between the encoded information and the requirements of the retrieval environment does not hold up under scrutiny. Specifically, Morris, Bransford and Franks (p. 530) claimed that Craik and Tulving's Experiment 8 (1975) did not test what was actually learned. Fisher and Craik (1977) as well as others, have provided ample evidence to demonstrate that retention is highest when the TBR words are encoded and cued in the context of the same type of information, i.e., it was found that rhyme cues led to optimal recall for rhyme encoded words while category cues maximized recall for words encoded categorically, and finally words embedded in sentence frames were best recalled when cued by the sentence frames. It is stressed however that providing the retrieval cues at recall did not equate retention across conditions. Deeper semantic encodings still produced higher levels of recall than shallower levels of processing.

Despite the fact that Morris, Bransford and Franks acknowledged that semantic acquisition - semantic test conditions evidenced better recall than the rhyme acquisition - rhyme test conditions there was some reluctance to interpret these results within a levels of processing framework. Rather the authors opted for Jenkins'

(1974) explanation that memory for an event depends on whether the subject possesses and/or uses appropriate knowledge and skills to formulate an optimal encoding. If Morris, Bransford and Franks were suggesting that due to an impoverished semantic memory the subject is sometimes unable to form or retrieve episodic memory traces then there seems no great disagreement between their point of view and what is essentially the view of the chief proponents of the levels of processing approach (Lockhart, Craik & Jacoby, 1976).

Stein (1978), further questioned the assumption that non-semantic processing necessarily is less meaningful and leads to less durable memory traces. He joined ranks with Arbuckle and Katz (1976) and Morris, Bransford and Franks (1977) suggesting that given test-appropriate conditions, memory is a faithful measure of the interaction of the TBR material and the task demands (e.g., the level of processing). Specifically it was hypothesized that subjects differentiate successive inputs in terms of the task at hand. It is therefore appropriate to test for the retrieval of precisely that information which was encoded. Accordingly it was thought that case encoding questions (e.g., "________ has a capital I") should yield higher recognition performance than semantic encoding (e.g. The ________ rolled down the hill.) when the mode of testing was appropriate (e.g., a case-oriented recognition test). The data confirmed that "given an appropriate test, case encoding can yield superior retention of the stimulus than semantic encoding" (p. 168). Stein preferred to characterize these results in terms of "precision" of encoding rather than "levels of processing."

To the present author, the ideas advanced by Stein, or the claim that retention is optimal when test-appropriate conditions exist (Morris, Bransford & Franks, 1977), are not substantially different from Thomson and Tulving's (1973) encoding specificity principle - a point of view with which levels proponents would readily agree. All that has been demonstrated is that for recall to be maximal the original encoding context must be re-established.

Elaboration within a Domain (or Level) of Processing

An important aspect of the original levels viewpoint that has survived revisions is that the richness of the memory trace is essentially a function of the number and nature of the attributes analyzed during input (Lockhart, Craik & Jacoby, 1976). As a result of the rather large body of existing evidence these authors also advanced the notion that deeper, richer encodings are also more 'distinctive' and the resulting episodic memory traces are unique and therefore more easily recovered or reconstructed during both recognition and recall. There is also some suggestion that degree of elaboration (even within a domain) is related to both the number and the nature of the features analyzed (p. 81). However, one might emphasize that it is the quality rather than the quantity of the features analyzed that makes for a durable or rich memory trace. It is in the light of this statement that further research (within the semantic domain) is to be considered.

Levels researchers have increasingly utilized Tulving's episodic/semantic memory distinction (Craik & Jacoby, 1975; Craik &

Tulving, 1975; Lockhart, Craik & Jacoby, 1976) in an attempt to elucidate the processes thought to be operating in short-term retention. Tulving (1972) defines semantic memory as follows (p. 386):

> Semantic memory is the memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meanings and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of these symbols, concepts, and relations. Semantic memory does not register perceptible properties of inputs, but rather cognitive referents of input signals. The semantic system permits the retrieval of information that was not directly stored in it, and retrieval of information from the system leaves its content unchanged, although any act of retrieval constitutes an input into episodic memory.

Using Tulving's framework, the experiments investigating depth of processing have to do with the formation and retrieval of various episodic memory traces, since memories for lists of items presented to subjects in experimental situations are autobiographical by nature. Incoming stimuli may be recorded directly in episodic memory by way of various perceptual analyses (i.e., it is possible for the episodic system to act independently of the semantic system). The stimulus may be temporally tagged (Tulving, 1972) or encoded in terms of structural or other relatively shallow attributes. Additionally certain aspects or features of the stimuli may be related to, and integrated with, the stored contents of semantic memory.

Craik, Lockhart and Jacoby (1976) argue for a closer interdependence between the two systems than originally proposed by Tulving. For example, a person may encode incoming stimuli in terms of their perceptible properties and/or additionally access the contents of semantic memory. The resultant trace of the stimulus becomes an event in episodic memory. That the durability of the trace is a function of the elaborative aspects of the encoding operations has already been stated (Lockhart, Craik & Jacoby, 1976).

One might posit that recourse to the contents of semantic memory is only encouraged by deeper levels of encoding. There is some experimental support for this notion. Seamon and Murray (1976) speculated that level of processing might interact with serial position, with recall of non-terminal items being greater for semantic versus non-semantic tasks. Subsequent results confirmed the expectation of a primacy effect for semantic processing only, especially with the highly meaningful stimuli.

Further it is thought that, whenever the task demands require relatively shallow encodings and no inputs to the episodic memory system are derived from semantic memory, retrieval for that item is less likely than if the memory trace is enriched by extracting features from semantic memory. For example, proof-reading a manuscript for typographical or spelling errors can be done without processing the deep meaning of the sentences. Similarly deep meaning can be accessed (presumably from semantic memory) and retained without a great deal of surface structure analysis being consciously carried out. To illustrate this point Graesser and Mandler (1975) found that memory for strictly surface structure characteristics of sentences was poor when subjects were induced to concentrate primarily on semantic analyses. It has also been a personal observation of the author that bilingual people reading connected discourse for meaning seem not to notice changes at the surface structure level from one language to the other. e.g., "Would you please open the porte?" What is being put forward here is the proposition that subjects can directly access the target domain where analyses relevant to the task demands can be performed. In this sense the episodic memory trace can be formed either with the assistance of the contents of semantic memory or completely independently of it. The memory for that event is likely to be more or less durable depending on the degree of semantic memory involvement.

Levels within Levels - The Semantic Domain

Not a few researchers have pointed out the need to investigate the concept of levels within a broad domain (Baddeley, 1976, 1978; Bellezza, Cheesman & Reddy, 1977; Elias & Perfetti, 1973; Goldman & Pellegrino, 1977; Klein & Saltz, 1976; Mistler-Lachman, 1974; Nelson, 1977). Baddeley (1976) acknowledged a shortcoming in the Craik and Lockhart approach in that there was no clearly specified means of comparing the processing demands within the semantic domain (a point also advanced by Nelson, 1977). More recently Baddeley (1978) lamented the fact that no one had systematically explored the concept of a continuum of processing within a broadly defined domain. Baddeley also observed that Craik and Tulving (1975), after finding essentially no differences in levels of recall in Experiment 1 between two semantic tasks (category, sentence), confined further attention to only one semantic task in subsequent experiments.

Hyde (1973) also failed to find increased recall for subjects who rated words on two semantic dimensions as opposed to words rated along one dimension alone. However to the present author, it is inconceivable that all semantic processing leads to the same memory trace durability.

Other researchers fortunately have been more persistent. The question of greatest interest is that, if there is a gradation of memory performance associated with different types of semantic tasks, is it due to qualitatively different types of analyses or is the richness of the memory trace a function of the number of semantic features encoded? The latter view is appealing to some, since the memory trace is often conceived by many researchers as bundles of attributes (Underwood, 1969) or attribute dimensions (Saltz, 1971).

Elias and Perfetti (1973) found that recognition varied as function of type of semantic encoding. Subjects were asked to give as many free associates as possible to a TBR word in 10 seconds while another group was asked to give synonyms or words that were "very close in meaning" to the stimulus word. Although both groups outperformed a group of intentional learners there were no differences in the hit rate for old words between the synonym and associative groups. The hypothesis had been synonym > associative since it was argued that in the course of generating synonyms more semantic attributes would have to be searched for and tagged than in the course of producing associative responses. Although the predicted difference was in the right direction, a slightly different rationale might have predicted the same differences. It may be the case for example, that synonym encoding does not simply lead to more semantic features being tagged (or extracted from semantic memory) but qualitatively richer or more distinctive features which are more highly accessible during recognition. To develop this point further, consider Gardiner's (1974) study that employed two levels of semantic processing. He found that subjects classifying TBR words (e.g., Poland) as a country (general) or a European country (specific) showed better recall after the latter task that required more specific semantic processing.

Klein and Saltz (1976) attempted to differentiate levels of processing within the semantic domain. A preliminary rating study was carried out to establish the correlations between the various semantic attributes (e.g., pleasant-unpleasant, big-little). Following this a series of incidental learning studies was conducted wherein subjects were instructed to rate a list of nouns on a) a dimension irrelevant to the word's meaning (e.g. vowel counting), b) one relevant dimension (e.g., pleasant-unpleasant), c) two relevant dimensions that were correlated with each other (e.g., fast-slow, happy-sad) or d) two relevant dimensions that were poorly correlated with each other (e.g., big-little, pleasant-unpleasant).

With a single exception recall increased in concert with the change in levels of processing described above. The exception was that the "happy-sad" task produced recall that exceeded all other single dimension tasks and was not significantly different from any of the conditions where TBR words were rated on two semantic dimensions. The most important outcome, however, was that recall involving rating on two uncorrelated dimensions produced better recall than rating on two moderately correlated dimensions. The authors interpreted their overall findings in the light of Saltz's (1971) conceptual model of cognitive space. Briefly the theory postulates that a concept can be specified in space in terms of its representation on certain dimensions. The more clearly specified its location is (as would be the case with the intersection of two independent dimensions) the more clearly defined the concept is and hence it is less likely to be subject to interference effects as would be the case if the concept were more diffusely specified. Thus, Klein and Saltz (1976) account for their results by suggesting that deeper levels of processing within the semantic domain activate or specify the more relevant or meaningful attributes associated with the TBR item.

Fraise and Kamman (1974) provided further support for the proposition that specificity is a correlate of depth of processing. When they asked subjects to search for specific exemplars of a concept (e.g., a vegetable) they recalled more than if the search entailed locating an instance of the more general class (e.g., food). Note that the concept of specificity is not radically different from the idea of distinctiveness or uniqueness advanced earlier by Moscovitch and Craik (1976) to account for levels of processing results. Morevoer in another study alluded to earlier, (Reed Hunt, & Mitchell, 1978) facilitative effects were demonstrated for specific or distinctive encodings in both the semantic and non-semantic domains. Seamon and Murray (1976), in the context of demonstrating the superior retention effects of semantic processing over structural processing utilized a semantic orienting task similar to that used by Fraise and Kamman (1974), i.e., subjects decided if a TBR word was a general (TOOL) or a specific (HAMMER) instance of the same semantic category. Unfortunately, these authors did not report the data in sufficient detail to allow conclusions to be drawn about the specificity principle.

It is not clear from the Klein and Saltz (1976) study whether deeper processing involves the nature of the attributes activated or the number of relevant attributes activated or perhaps both. Certainly the anomalous finding of the heightened recall for the single relevant dimension task "happy-sad" does not permit the drawing of any strong conclusions attributing the effect to the number of encoded attributes.

Evidence for the additivity effects of multiple encoding attributes of TBR words (within and across various domains) can be derived from a study by Goldman and Pellegrino (1977). Unfortunately the multiple encodings within a domain were not along various attribute dimensions but in fact amounted to three repetitions of the same encoding question. Still results showed strong additivity effects even across domains (orthographic, acoustic and semantic) for recall and recognition. Perhaps the most significant general finding was that multiple encodings were most beneficial when they involved 'Yes" responses in the semantic domain.

Epstein, Phillips and Johnson (1975) in an ingenious procedure contributed solid evidence for the existence of differing levels of semantic processing. Moreover, they added strength to the growing conviction that depth of processing is not only a function of the task demands but also the way these demands interact with the nature of the TBR material.

To begin with, it was postulated that finding a difference between semantically related words would require deeper processing than finding a similarity. The converse argument was put for unrelated pairs, i.e., finding a similarity would call for a deeper level of processing if the pair was unrelated. The hypothesized interaction was confirmed and is interpreted here as strong evidence for a levels within levels viewpoint. The findings are entirely compatible with Craik, Lockhart and Jacoby's (1976) statement that "the operations carried out on the stimulus depend on the interpretive task given to the subject...if the stimulus is rare, unexpected or must be discriminated from similar stimuli, many features are analyzed and a rich memory trace results" (p. 81).

The only discomforting aspect of the above analysis is that the emphasis still appears to be on the number of features analyzed. The present author prefers to interpret the Epstein, Phillips and Johnson findings as reflecting the consequences of qualitatively different operations, with the deeper task (i.e., finding a similarity between an unrelated word pair) producing better recall because it required more conscious effort or attention to perform the task. Greater conscious effort is needed because, as Lockhart, Craik and Jacoby (1976) suggest, the contents of semantic memory are accessed and altered as a result of the deeper processing.

The concept of conscious attention or effort as an index of depth of processing has a great deal of intuitive appeal. Consider a study recently conducted by Schnur (1977) where subjects in one condition were instructed to encode exemplars of various categories using a design similar to that of Craik and Tulving (1975). The independently manipulated variable was the relationship between the encoding context (category type) and the TBR word. This was varied by presenting either high, medium or low exemplars of various categories as derived from norms supplied by Rosch (1975). For example a high ranking exemplar of the category "weapon" was knife; a medium ranked exemplar was bullet and a low ranking exemplar was car.

It was found that low ranking exemplars were classified more slowly but showed better retention on a subsequent recognition test (for 'Yes' responses only) than high ranking exemplars. In Experiment 2 using a paired associate paradigm again, low ranking exemplars were again better recalled than high ranking pairs. Although the results were interpreted in terms of the encoding elaboration hypothesis in that each TBR word and its category formed a unique or highly elaborated unit particularly in the case of low ranking exemplars it is also arguable that the process of classifying low ranking exemplars required more conscious attention or effort. To the extent that conscious effort was required to formulate unusual encodings (e.g., car as a type of weapon), then the findings are consonant with the explanation given here and recently advanced by Lockhart, Craik and Jacoby (1976).

In a just published study (Auble & Franks, 1978), subjects were read sentences that were generally incomprehensible (e.g., "the street was full of holes because the turning stopped.") until an appropriate cue was provided (e.g., cement mixer). For one group the cue was made available 5 seconds after the sentence had been read. Other groups received the cue immediately following input and then had 5 seconds in which to elaborate the stimulus or alternatively had the cue directly integrated within the sentence. Recall was best for the first 2 groups since active or conscious effort was required to integrate the cue with the sentence. Where the cue was already meaningfully integrated few analyses were required to extract meaning and hence trace durability was not as great. In conclusion, it is clear that not all semantic processing results in equivalent recall. It is equally clear that the question of whether semantic encoding is better thought of in quantitative terms remains unresolved. Once access is gained to the semantic domain the durability of the memory trace may be a function of the number of relevant attributes activated. The alternative centers on the notion that what is more crucial is the qualitative nature of the mental operations performed. A third alternative may involve some combination of both processes.

The Case Against Levels of Processing

While the levels of processing hypothesis has met with reasonable empirical support there have been several attacks on the fundamental principle itself or some aspects of its elaboration.

Nelson (1977) expressed some serious concerns relating to the inherent circularity of the whole notion of depth of processing.

The basic objection was that whenever obtained differences in retention were found as a result of manipulating type of processing the retention differences were ascribed to greater depth of stimulus processing or elaboration. In other words, the absence of independently defined criteria that permit a priori predictions to be made about memorial outcomes as a result of various levels of analyses rendered the concept scientifically empty. Hence, Nelson (1977) and Nelson and Vining (1978) preferred to attribute differences in memory performance to different kinds (rather than levels) of processing.

Seamon and Virostek (1978) overcame this basic objection by providing an independent assessment of depth of processing and memory performance. Subjects rank ordered Nelson's (1977) thirteen exemplary classification tasks according to the degree of difficulty associated with each task. Tasks ranged over four broad categories: orthographic classifications, acoustic-phonological classifications, category-based classifications and meaning or object classifications. The subject defined depth of processing correlated significantly with free recall scores for each task after incidental learning by a separate group of subjects. Hence, the data consolidated earlier claims as to the advantage of semantic over non-semantic processing since classification questions involving category or meaning referent judgments, in addition to being judged to entail deeper processing, consistently yielded higher recall scores than the orthographic classifications. More importantly this study dispelled the notion that depth of processing could not be assessed independently of memory performance. The data, however, do not suggest a continuum either within or across processing domains. In some instances tasks evidencing little difference in subject defined 'depth' were found to differ markedly in probability of recall.

Nelson (1977) also attempted to refute a claim made in Craik and Lockhart's (1972) original paper that Type 1 (or maintenance rehearsal, Craik & Watkins, 1973) processing does not increase the durability of the memory trace i.e., repetition of a particular level of analysis would not lead to increased memory performance. Nelson's results showed that two repetitions of a TBR word requiring either the same or a different phonemic decision with respect to that word led to better recall than one repetition at the phonemic level. The conclusion drawn was that repetition at a constant depth of processing does in fact enhance recall despite earlier research that reported contradictory findings (Craik & Watkins, 1973; Jacoby, 1973). While this conclusion is not basic to the present enterprise, it might be pointed out that repetition, in the sense that the subject was presented with the TBR item more than once, should not be interpreted to mean that the same level of analysis was repeated by the subject each time even though the orienting task might require two different but still constant depth decisions (as in the case of two repetitions with a different phonemic question each time). It is almost impossible to say with any certainty that the subject is restricting his level of analysis to that specified by the orienting task on the second repetition, especially when the phonemic decision could have been made well within the allotted presentation time of four seconds per item. However, there is evidence presented earlier to the effect of multiple encodings (i.e., repetition of the same analysis) can lead to increased recall and recognition (Goldman & Pellegrino, 1977). In addition, Nelson (1977) questions how, if same depth analysis does not enhance memory, the levels of processing proponents account for increased retention across trials in a multi-trial learning situation. Most likely the proponents of levels of processing would respond by indicating that the subject engages in deeper levels of stimulus analysis on each trial (or repetition).

The assumption that deeper levels of elaboration ensures more durable traces has also been challenged (Nelson & Vining, 1978). In the present survey of the literature it has not been at all clear whether effects of levels of processing operate primarily at acquistion or additionally at retention or perhaps both. Nelson and Vining (1978) conceded that levels of processing exerts a considerable effect on acquisition but has little or no effect on long term retention. After conducting pilot research to determine a suitable retention period (1 week) and the number of trials to equate level of learning (2 vs.10) between groups that acquired items via semantic ("does the word represent something you could hold in your hand?") or structural ("does the word contain more than one syllable?") orienting tasks subjects were presented with a list of unrelated words. The initial recall test, given immediately after list presentation, showed no significant differences between groups. However, there were differences in retention favouring the semantic group. These data were explained away by claiming that the slight differences at acquisition, though not statistically significantly, did favour the semantic group versus the structural processing group (.71 vs. .64) and hence subtle carry-over effects may have been responsible for the (somewhat larger) retention differences. Furthermore, it was claimed that there was a lack of substantial power in at least one of the comparisons.

It was resolved, therefore, to repeat the experiment reducing the retention interval to four days and changing both the nature of tasks so the semantic question "Does the word represent something that occurs naturally in the world?" was asked once after each TBR word whereas the structural processing group saw the list ten times responding to the same question, "Does the word contain a long vowel sound?" Another change was that separate groups were tested at acquisition versus retention. The results again showed no differences at acquisition. In contrast to Experiment 1 the semantic processing condition showed no worse retention after four days than the structural processing group. It was concluded that overall there is "no consistent advantage of semantic processing (or vice versa)" (p. 208) after groups have been equated for level of learning. While it is agreed that level of learning between groups was indistinguishable at immediate recall it is hardly conceivable that the subjects' credibility was preserved over ten incidental trials with the same encoding question for the structural processing group. It is suggested that processing of a more elaborate nature may have also taken place. If so, then the results are less clear cut.

Reference is made to two studies where delayed recall also differentially affected rate of information loss depending on the nature of the original encoding.

At the level of sentences, Treisman and Tuxworth (1974) showed that although immediate recall levels were equal for two monitoring tasks (i.e., searching for semantic or phonemic anomalies) after a 20 second filled delay interval, recall favoured subjects who were engaged in the semantic monitoring tasks. McMurray and Duffy

(1972) also equated for level of acquisition across groups by varying presentation rate. The TBR stimuli were letter strings that were pronounceable-nonmeaningful (e.g., PAC), meaningful but unpronounceable (e.g., CTV) or nonmeaningful and unpronounceable (e.g., HGT). Interestingly under intentional learning conditions recall, after 18 seconds of interpolated activity, was best for the last type of material. It is argued that while the stimuli in the last condition were novel or unusual the amount of conscious effort required to carry out the relevant analyses (i.e., raise the level of immediate recall to that of the other two groups) resulted in a more durable memory trace.

Baddeley (1976) expressed some serious misgivings about the status of depth of processing as a heuristic tool for investigating memory phenomena. These concerns are elaborated further (Baddeley, 1978) and some of them (e.g., lack of evidence for continuum within a domain; the problem of circularity in the concept itself; "levels" versus "degree of elaboration"; the necessity of deep processing for durable memory traces) have been met and dealt with elsewhere in this review. One criticism that remains most cogent is that depth of processing is too global and lacks a precise definition that would allow predictions to be made about the relative effectiveness of various processing tasks, particularly within a single domain. Another criticism is that too many of the studies have been essentially replicative in nature, a point with which the present author tends to agree.

Conclusions

As we have seen not all retention differences are accountable by postulating single or simple mechanisms like depth of processing or degree of elaboration (Jacoby, Bartz & Evans, 1978). On the basis of their experimental work these authors concluded that retrieval cues and the retrieval context generally interact with encoding operations and meaningfulness of the material to affect both recall and recognition. To reinforce their conclusions, we suggest that the complex interactive effects of some of the following variables still require clarification: cue-target uniqueness; level of encoding; number of repetitions; level of scaled meaningfulness of TBR material (and its compatibility with the existing cognitive structures of the learner); frequency; degree of organization and sufficient time for the subject to perform relevant analyses. No simple account of memory and the manner in which it functions awaits us.

Perhaps the most promising aspect of the levels approach is its focus on processes which has taken us away from the rigid notions of the multistore models of memory (although see Glanzer & Koppenaal, 1976 who propose the two approaches differ mostly in terminology and points of emphasis). Murdock (1974) acknowledged the problems that plague that multistore theorist, especially as they concern each of the stages and the need to describe them in terms of their coding, capacity and retention characteristics. Murdock (p. 261) as a result favours a continuity view of memory and is in basic agreement with the levels of processing approach in that he prefers not to think of memory in terms of discrete temporal partitions.

Now the present emphasis in memory research is on processes (as opposed to stages) - specifically the processes of encoding, storage and retrieval. Although this paper has touched on all three, the emphasis has been squarely on the nature of the encoding operations the learner employs in relating to and processing various aspects of the TBR material in accord with the particular demands of the experimental situation.

Much of what has been reviewed can be accomodated by Craik and Jacoby's (1975) process view of short-term retention or primary memory. In terms of its broad characteristics, their concept of primary memory is not radically different from Waugh and Norman's (1965) or even Baddeley's idea of working memory (see Baddeley & Hitch, 1974). The point of emphasis for Craik and Jacoby (1975) is the role of attention as the central processor. It is acknowledged to be similar to Treisman's (1969) concept of attention where deeper analyses successively require more conscious attention for their completion. Memory performance is seen as an increasing function of these levels of processing that occur within a more or less continuous system of perceptual/cognitive analyses. The model does not however, preclude the idea of a secondary system or what James (1890) called "memory proper." More recently this system has been referred to as semantic memory (Tulving, 1972) and is regarded as a repository of generalized knowledge about the world.

An item can be maintained in primary memory by continued attention to the item and as a result that encoded trace forms the latest addition to episodic memory. Rather than being displaced when conscious attention is diverted from it, it is simply pushed further down the line, i.e., items are temporally ordered in episodic memory. It is important to note however, that in Craik and Jacoby's model, the semantic memory system can provide independent inputs into episodic memory. Accordingly, memory can be a function of two types of input i.e., either external or internal or both (Norman & Bobrow, 1977). To clarify, the newly encoded episodic memory trace can be partially formed on the basis of elaborate cognitive processes where relevant features or attributes are extracted from semantic memory as well as being the by-product of perceptual analyses of the sensory input. This represents a refinement of Craik and Lockhart's (1972) original view. Now processing can involve the use of cognitive structures at all levels of perceptual analysis.

Retrieval for recently encoded events is effected via a backwards scanning mechanism for temporally tagged items. After a long delay between presentation and test, extensive reconstructive strategies involving existing cognitive structures are required to retrieve the TBR item. Craik and Jacoby (1975) conducted a study to underline the need to postulate two retrieval processes-backwards scanning and quided reconstruction. Using a Craik and Tulving (1975) intentional learning deisgn (i.e., case, rhyme, semantic processing), encoding trials were interspersed with recognition trials where the TBR item was re-presented after a lag of 0-23 trials. Results showed no differences in recognition rate for each type of encoding (presumably because all items were relatively recent and were retrievable via the backwards scanning mechanism). In a subsequent free recall test, however, the now familiar pattern of results emerged, i.e., semantic > phonemic > structural. The differences were thought to be due to the differential use subjects were able to make of reconstructive strategies where now the contents of semantic memory come into play.

In summary, this paper has been about human memory performance and any theory that attempts to place memory in its proper perspective as an integral part of an information processing analysis of human cognition needs to keep an important distinction in mind - between the contents of memory and the operations performed upon it. The part of the memory that has received the bulk of attention in this review has been the qualitative nature of mental operations and processes carried out by the learner while encoding incoming information. The central processor postulated to underlie short-term retention is conscious attention (Craik & Jacoby, 1975). William James (1890) said that memory involved thinking about things which we have experienced but about which we were not thinking immediately before. Today, almost 100 years later, we are re-examining conscious attention as one of the chief control processes in memory.

APPENDIX B:

Practice Sheet for Nine Semantic Processing Tasks

EXPERIMENT I

Practice Sheet for Nine Semantic Processing Tasks

	Pliers	Crab
A. Decide if each word is Living or Nonliving	______	______
B. Decide if each word is Pleasant or Unpleasant	______	______
C. Decide if each word is Strong or Weak	______	______
D. Write 2 free associates for each word	______	______
	______	______

E. Define each word

1. ______________________________

2. ______________________________

F. Write a sentence for each word

____

____

____

____

G. Write a few words saying how the 2 words are <u>different</u>

____

____

H. Write a single sentence for each word

____

____

I. Write a few words saying how the 2 words are <u>similar</u>

____

____

APPENDIX C:

Examples of Paired Comparison Booklets

EXPERIMENT I

Example of a Paired Comparison Booklet

Page 1

Mark with a X the task that requires more conscious effort or attention.

Write a few words saying how the 2 words are <u>different</u> ________
Decide if each word is Living or Nonliving ________

Decide if each word is Strong or Weak ________
Write a sentence for each word ________

Decide if each word is Pleasant or Unpleasant ________
Write a sentence for each word ________

Write a few words saying how the 2 words are <u>similar</u> ________
Write a few words saying how the 2 words are <u>different</u> ________

Decide if each word is Living or Nonliving ________
Write 2 free associates for each word ________

Write a <u>single</u> sentence using both words ________
Decide if each word is Pleasant or Unpleasant ________

Define each word ________
Decide if each word is Living or Nonliving ________

Decide if each word is Strong or Weak ________
Write a <u>single</u> sentence using both words ________

Decide if each word is Pleasant or Unpleasant ________
Write a few words saying how the 2 words are <u>different</u> ________

Page 2

Write a sentence for each word ________
Write a <u>single</u> sentence using both words ________

Write a few words saying how the 2 words are similar ________
Decide if each word is Living or Nonliving ________

Write 2 free associates for each word ________
Define each word ________

Decide if each word is Strong or Weak ________
Write a few words saying how the 2 words are different ________

Write a single sentence using both words ________
Write a few words saying how the 2 words are similar ________

Write a few words saying how the 2 words are similar ________
Decide if each word is Pleasant or Unpleasant ________

Define each word ________
Decide if each word is Strong or Weak ________

Write a2 free associates for each word ________
Write a few words saying how the 2 words are different ________

Decide if each word is Living or Nonliving ________
Decide if each word is Strong or Weak ________

Page 3

Decide if each word is Pleasant or Unpleasant ________
Decide if each word is Living or Nonliving ________

Write a single sentence using both words ________
Write 2 free associates for each word ________

Define each word ________
Decide if each word is Pleasant or Unpleasant ________

Decide if each word is Strong or Weak ________
Write a few words saying how the 2 words are similar ________

Write a <u>single</u> sentence using both words ________

Write a few words saying how the 2 words are <u>different</u> ________

Decide if each word is Living or Nonliving ________

Write a <u>single</u> sentence using both words ________

Define each word ________

Write a few words saying how the 2 words are <u>different</u> ________

Page 4

Decide if each word is Strong or Weak ________

Decide if each word is Pleasant or Unpleasant ________

Write a few words saying how the 2 words are <u>similar</u> ________

Write a sentence for each word ________

Write a sentence for each word ________

Write 2 free associates for each word ________

Decide if each word is Living or Nonliving ________

Write a sentence for each word ________

Decide if each word is Pleasant or Unpleasant ________

Write 2 free associates for each word ________

Write a

Write a few words saying how the 2 words are similar ________

Define each word ________

Write a few words saying how the 2 words are <u>different</u> ________

Write a sentence for each word ________

Decide if each word is Strong or Weak ________

Write 2 free associates for each word ________

Define each word ________

Write a <u>single</u> sentence using both words ________

EXPERIMENT IV

Example of Paired Comparison Booklet

Page 1

Mark with an "X" the task in each pair below that required more conscious effort for you to carry out. Remember it may have not been the longer task but simply required more conscious or elaborate thinking on your part.

Deciding if each word was used in a good, meaningful sentence. ________
Deciding if each word was a good associate to the word at the top. ________

Deciding if each word was a good associate to the word at the top. ________
Deciding if each word was a good example of a given category. ________

Deciding if each word had a good definition. ________
Deciding if each word was a good example of a given category ________

Deciding if each word was Living or Non-living. ________
Deciding if each word was used in a good, meaningful sentence ________

Decide if each word was a good example of a given category. ________
Decide if each word was Living or Non-living ________

Page 2

Deciding if each word was used in a good, meaningful sentence. ________
Deciding if each word was a good example of a given category. ________

Deciding if each word was Living or Non-living. ________
Deciding if each word was a good associate to the word at the top. ________

Deciding if each word had a good definition. ________
Deciding if each word was used in a good, meaningful sentence. ________

Deciding if each word had a good definition. ________
Deciding if each word was Living or Non-living ________

Deciding if each word was a good associate to the word at the top. ________
Deciding if each word had a good definition. ________

APPENDIX D:

Response Patterns for Paired Comparisons Data

EXPERIMENT I

Response Patterns for Paired Comparisons

Between Nine Semantic Tasks

Subject No.1
ENSEMBLE - ABDUCTION

S	1	2	3	4	5	6	7	8	9	R_j
1		1	1	1	1	1	1			6
2							1			1
3		1			1		1			3
4		1	1		1	1	1			5
5		1								1
6		1	1		1		1			4
7					1					1
8	1	1	1	1	1	1	1			7
9	1	1	1	1	1	1	1	1		8

Subject No.2
RESIDUE - NEPHEW

S	1	2	3	4	5	6	7	8	9	R_j
1		1	1	1	1	1	1	1	1	8
2										
3		1		1	1	1	1	1		6
4		1			1	1	1	1		5
5		1					1			2
6		1			1		1			3
7		1								1
8		1			1	1	1			4
9		1	1	1	1	1	1	1		7

Subject No.3
BUFFOON - AMAZEMENT

S	1	2	3	4	5	6	7	8	9	R
1		1	1	1	1	1	1		1	7
2										0
3		1					1			2
4		1	1			1	1			4
5		1	1	1		1	1		1	6
6		1	1				1			3
7		1								1
8	1	1	1	1	1	1	1		1	8
9		1	1	1		1	1			5

Subject No.4
EXTERMINATION - WENCH

S	1	2	3	4	5	6	7	8	9	R
1		1	1		1	1	1		1	6
2										0
3		1					1			2
4	1	1	1		1	1	1		1	7
5		1	1			1	1			4
6		1	1				1		1	4
7		1								1
8	1	1	1	1	1	1	1		1	8
9		1	1		1		1			4

Subject No.5

CHARLATAN - IMPOTENCY

S	1	2	3	4	5	6	7	8	9	R
1		1	1	1	1	1	1		1	7
2							1			1
3		1					1			2
4		1	1		1		1			4
5		1	1				1			3
6		1	1	1	1		1		1	6
7										0
8	1	1	1	1	1	1	1		1	8
9		1	1	1	1		1			5

Subject No.6

MAKER - OPINION

S	1	2	3	4	5	6	7	8	9	R
1		1	1		1		1	1		5
2			1				1			2
3										0
4	1	1	1		1		1	1		6
5		1	1				1			3
6	1	1	1	1	1		1	1	1	8
7			1							1
8		1	1		1		1			4
9	1	1	1	1	1		1	1		7

Subject No.7

WHOLESALER - DECEIT

S	1	2	3	4	5	6	7	8	9	R
1		1	1	1	1	1	1	1	1	8
2			1				1			2
3							1			1
4		1	1		1		1			4
5		1	1				1			3
6		1	1	1	1		1			5
7										0
8		1	1	1	1	1	1		1	7
9		1	1	1	1	1	1			6

Subject No.8

ORIGINATOR - MALICE

S	1	2	3	4	5	6	7	8	9	R
1		1	1	1	1	1	1		1	7
2					1					1
3		1			1					2
4		1	1		1	1	1		1	6
5										0
6		1	1		1		1			4
7		1	1		1					3
8	1	1	1	1	1	1	1		1	8
9		1	1		1	1	1			5

Subject No.9
FOLLY - ARMADILLO

S	1	2	3	4	5	6	7	8	9	R
1		1	1		1	1	1			5
2					1					1
3		1			1					2
4	1	1	1		1	1	1		1	7
5										0
6		1	1		1		1			4
7		1	1			1				3
8	1	1	1	1	1	1		1	1	8
9	1	1	1		1	1	1			6

Subject No.10
FALLACY - ARBITER

S	1	2	3	4	5	6	7	8	9	R
1		1	1		1		1			4
2					1					1
3		1			1		1			3
4	1	1	1		1	1	1		1	7
5										0
6	1	1	1		1		1			5
7		1			1					2
8	1	1	1	1	1	1	1		1	8
9	1	1	1		1	1	1			6

Subject No.11
PACIFISM - HENCHMAN

S	1	2	3	4	5	6	7	8	9	R
1		1	1		1	1	1		1	6
2			1				1			2
3							1			1
4	1	1	1		1	1	1		1	7
5		1	1				1			3
6		1	1		1		1		1	5
7										0
8	1	1	1	1	1	1	1		1	8
9		1	1		1		1			4

Subject No.12
BARD - CHARTER

S	1	2	3	4	5	6	7	8	9	R
1		1	1	1	1	1	1		1	7
2										0
3		1					1			2
4		1	1			1	1			4
5		1	1	1		1	1			5
6		1	1				1			3
7		1								1
8	1	1	1	1	1	1	1		1	8
9		1	1	1	1	1	1			6

EXPERIMENT IV

Response Patterns for Paired Comparisons Between Five Semantic Tasks

Subject No.1

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-LN						0

Subject No.2

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-LN						0

Subject No.3

	FA	SENT	CAT	DEF	L-NL	R_j
FA						0
SENT	1		1		1	3
CAT	1				1	2
DEF	1	1	1		1	4
L-LN	1					1

Subject No.4

	FA	SENT	CAT	DEF	L-NL	R_j
FA						0
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-NL	1					1

Subject No.5

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1		1	3
CAT	1				1	2
DEF	1	1	1		1	4
L-NL						0

Subject No.6

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1	1	1	3
SENT	1		1	1	1	4
CAT					1	1
DEF			1		1	2
L-NL						0

Subject No. 7

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT		1	1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No. 8

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1		1	3
CAT	1				1	2
DEF	1	1	1		1	4
L-NL						0

Subject No. 9

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1			1	2
SENT				1	1	2
CAT	1	1		1	1	4
DEF	1				1	2
L-NL						0

Subject No. 10

	FA	SENT	CAT	DEF	L-NL	R_j
FA						0
SENT	1		1		1	3
CAT	1				1	2
DEF	1	1	1		1	4
L-NL	1					1

Subject No. 11

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1	1	1	3
SENT	1		1		1	3
CAT					1	1
DEF		1	1		1	3
L-NL						0

Subject No. 12

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1	1	1	4
CAT	1			1	1	3
DEF	1				1	2
L-NL						0

Subject No. 13

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No.14

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1		1	3
CAT					1	1
DEF	1	1	1		1	4
L-NL						0

Subject No. 15

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-NL						0

Subject No. 16

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1	1		1	3
SENT			1		1	2
CAT				1	1	2
DEF	1	1			1	3
L-NL						0

Subject No. 17

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1	1	1	1	4
SENT				1	1	2
CAT		1		1	1	3
DEF					1	1
L-NL						0

Subject No. 18

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1		1	3
CAT	1				1	2
DEF	1	1	1		1	4
L-NL						0

Subject No. 19

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1		1	3
CAT					1	1
DEF	1	1	1		1	4
L-NL						0

Subject No.20

	FA	SENT	CAT	DEF	L-NL	R_j
FA						0
SENT	1		1		1	3
CAT	1				1	2
DEF	1	1	1		1	4
L-NL	1					1

Subject No. 21

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1				1	2
CAT	1	1			1	3
DEF	1	1	1		1	4
L-NL						0

Subject No. 22

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1	1	1	1	4
SENT			1	1	1	3
CAT					1	1
DEF			1		1	2
L-NL						0

Subject No. 23

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-NL						0

Subject No. 24

	FA	SENT	CAT	DEF	L-NL	R_j
FA						0
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-NL	1					1

Subject No. 25

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1	1	1	3
SENT	1		1	1	1	4
CAT					1	1
DEF			1		1	2
L-NL						0

Subject No. 26

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No. 27

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-NL						0

Subject No. 28

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No. 29

	FA	SENT	CAT	DEF	L-NL	R_j
FA				1	1	2
SENT	1		1	1	1	4
CAT	1			1	1	3
DEF					1	1
L-NL						0

Subject No. 30

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No. 31

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No. 32

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-NL						0

Subject No. 33

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1			1	2
SENT			1		1	2
CAT	1				1	2
DEF	1	1	1		1	4
L-NL						0

Subject No. 34

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1		1	1	3
SENT						0
CAT	1	1			1	3
DEF		1	1		1	3
L-NL		1				1

Subject No. 35

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1		1	3
CAT	1				1	2
DEF	1	1	1		1	4
L-NL						0

Subject No. 36

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1	1	1	4
CAT	1				1	2
DEF	1		1		1	3
L-NL						0

Subject No. 37

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1	1	1	1	4
SENT			1	1	1	3
CAT				1	1	2
DEF					1	1
L-NL						0

Subject No. 38

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1		1	3
CAT					1	1
DEF	1	1	1		1	4
L-NL						0

Subject No. 39

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1				1	2
CAT		1			1	2
DEF	1	1	1		1	4
L-NL						0

Subject No. 40

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1	1	1	1	4
SENT					1	1
CAT		1		1	1	3
DEF		1			1	2
L-NL						0

Subject No. 41

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1	1	1	1	4
SENT				1	1	2
CAT		1		1	1	3
DEF					1	1
L-NL						0

Subject No. 42

	FA	SENT	CAT	DEF	L-NL	R_j
FA						0
SENT	1		1			2
CAT	1					1
DEF	1	1	1		1	4
L-NL	1	1	1			3

Subject No. 43

	FA	SENT	CAT	DEF	L-NL	R_j
FA		1	1	1	1	4
SENT				1	1	2
CAT		1		1	1	3
DEF					1	1
L-NL						0

Subject No. 44

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No.45

	FA	SENT	CAT	DEF	L-NL	R_j
FA					1	1
SENT	1		1		1	3
CAT	1			1	1	3
DEF	1	1			1	3
L-NL						0

Subject No. 46

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No. 47

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT						0
DEF	1		1		1	3
L-NL			1			1

Subject No. 48

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

Subject No. 49

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1	1	1	3
SENT	1		1	1	1	4
CAT					1	1
DEF			1		1	2
L-NL						0

Subject No. 50

	FA	SENT	CAT	DEF	L-NL	R_j
FA			1		1	2
SENT	1		1	1	1	4
CAT					1	1
DEF	1		1		1	3
L-NL						0

APPENDIX E:

Coefficients of Concordance for Experiments I and IV

EXPERIMENT I

Coefficient of Concordance of Ranks Assigned
to 9 Semantic Tasks for 12 Raters

	Task									
	SIM	DIFF	DEF	2-WS	1-WS	2FA	S-W	P-UP	L-NL	
TASK #	8	1	9	4	6	5	3	7	2	WORD RATED
1	2	3	1	4	5	8	6	8	8	Ensemble-Abduction
2	5	1	2	4	6	7	3	8	9	Residue-Nephew
3	1	2	4	5	6	3	7	8	9	Buffoon-Amazement
4	1	3	5	2	5	5	7	8	9	Extermination-Wench
5	1	3	5	2	4	6	8	9	7	Pacifism-Henchman
6	1	2	4	5	3	6	7	9	8	Charlatan-Impotency
7	5	4	2	3	1	6	9	8	7	Maker-Opinion
8	2	1	3	5	4	6	8	9	7	Wholesaler-Deceit
9	1	2	4	3	5	9	7	6	8	Originator-Malice
10	1	4	3	2	5	9	7	6	8	Folly-Armadillo
11	1	5	3	2	4	9	6	7	8	Fallacy-Arbiter
12	1	2	3	5	6	4	7	8	9	Bard-Charter
R_j	22	32	39	42	54	78	82	94	97	

The mean of R_j = 540 ÷ 9 = 60

Kendall's coefficient of concordance: with tied ranks (see Siegel, 1956, pp. 229-238).

$$W = \frac{S}{\frac{1}{12} k^2 (N^3 - N) - k \sum_T T}$$

$$S = (22 - 60)^2 + (32 - 60)^2 + (39 - 60)^2 + (42-60)^2 + (54-60)^2 + (78 - 60)^2 + (82-60)^2 + (94-60)^2 + (97-60)^2$$

$$= 6,362$$

In the 12 rankings S^1 had 3 tied rankings

$$T = \frac{\Sigma(t^3 - E)}{12} = \frac{3^3 - 3}{12} = 2$$

S_4 also had 3 tied rankings ∴ $T_4 = 2$

$$\therefore \quad W = \frac{6362}{\frac{1}{12}(12)^2(9^3 - 9) - 12(4)}$$

$$= \frac{6362}{8592}$$

$$= .74$$

Using a chi square to test the significance of W

$$\chi^2 = k\,(N-1)\,W \qquad df = N-1$$

$$= 12\,(8)\,.74 \qquad = 8$$

$$= 71.13$$

EXPERIMENT IV

Coefficients of Concordance of Ranks Assigned to Five

Semantic Tasks by Fifty Raters

Subject #	SENT 2	DEF 4	FA 1	CAT 3	L-NL 5	Subject #	SENT 2	DEF 4	FA 1	CAT 3	L-NL 5
1	1	2	3	4	5	26	1	2	3	4	5
2	1	2	3	4	5	27	1	2	4	3	5
3	2	1	5	3	4	28	1	2	3	4	5
4	1	2	5	3	4	29	1	4	3	2	5
5	2	1	4	3	5	30	1	2	3	4	5
6	1	3	2	4	5	31	1	2	3	4	5
7	1	2	3	4	5	32	1	2	4	3	5
8	2	1	4	3	5	33	3	1	3	3	5
9	3	3	3	1	5	34	5	2	2	2	4
10	2	1	5	3	4	35	2	1	4	3	5
11	2	2	2	4	5	36	1	2	4	3	5
12	1	3	4	2	5	37	2	4	1	3	5
13	1	2	3	4	5	38	2	1	3	4	5
14	2	1	3	4	5	39	3	1	3	3	5
15	1	2	4	3	5	40	4	3	1	2	5
16	3.5	1.5	1.5	3.5	5	41	3	4	1	2	5
17	3	4	1	2	5	42	3	1	5	4	2
18	2	1	4	3	5	43	3	4	1	2	5
19	2	k	3	4	5	44	2	2	4	2	5
20	2	k	5	3	4	45	1	2	3	4	5
21	3	1	4	2	5	46	1	2	3	5	4
22	2	3	1	4	5	47	1	2	4	3	5
23	1	2	4	3	5	48	1	2	3	4	5
24	1	2	5	3	4	49	1	3	2	4	5
25	1	3	2	4	5	50	1	2	3	4	5
						R_j	89.5	102.5	156.5	161.5	240

Kendall's coefficient of concordance: with tied ranks (see Siegel, 1956, pp. 229-238).

$$W = \frac{S}{\frac{1}{2}k^2(N^3-N) - k\ {}_T T}$$

$$S = (89.5-150)^2 + (102.5-150)^2 + (156-150)^2 + (161.5-150)^2 + (240-150)^2$$

$$= (60.5)^2 + (47.5)^2 + (6)^2 + (11.5)^2 + (90)^2$$

$$- \ 14,184.75$$

$$T_x = \frac{(t^3-t)}{12}$$

$$T_9 = \frac{3^3 - 3}{12} = 2$$

$$T_{11} = \frac{3^3 - 3}{12} = 2$$

$$T_{16} = \frac{2^3 - 2}{12} + \frac{(2^3 - 2)}{12} = 1$$

$$T_{33} = \frac{3^3 - 3}{12} = 2$$

$$T_{34} = \frac{3^3 - 3}{12} = 2$$

$$T_{39} = \frac{3^3 - 3}{12} = 2$$

$$T_{45} = \frac{3^3 - 3}{12} = \underline{2}$$

$$\therefore \quad {}_r T = 13$$

$$\therefore \quad W = \frac{14,184.75}{(50)^2\ (5^3-5) - (50 \times 13)}$$

$$= \frac{14{,}184.75}{\frac{1}{2}(50)^2(5^3-5) - (50\text{x}13)}$$

$$= \frac{14{,}184.75}{\frac{1}{2}\ 2500\ (120) \quad - 650}$$

$$= \frac{14{,}184.75}{24.350}$$

$$= .58$$

Using a chi square to test the significance of W

$$\chi^2 = k\ (N-1)\ W \qquad df = N-1$$

$$= 50\ (4)\ .58 \qquad = 4$$

$$= 116.5$$

APPENDIX F:

Word Lists

Word List - Experiment II

Stimulus Words and Their Meaningfulness and Imagery Ratings - selected From Paivio, Yuille and Madigan (1968).

		M	I
abduction		4.80	4.93
amazement		4.50	4.47
arbiter		3.71	2.33
armadillo		4.38	3.77
bard		4.80	4.90
buffoon		4.50	4.33
charlatan		3.19	3.43
charter		4.89	3.83
deceit		4.92	3.30
ensemble		4.80	4.63
extermination		4.96	4.03
fallacy		3.72	2.30
folly		4.40	2.93
henchman		3.80	4.67
impotency		3.16	3.40
maker		5.00	3.57
malice		4.56	3.30
nephew		4.48	4.30
opinion		4.96	3.23
originator		4.48	3.30
pacifism		3.80	3.90
residue		4.48	4.87
wench		5.08	5.30
wholesaler		4.88	3.87
	$\overline{X}$	4.43	3.87
	s.d.	.57	.80

Word List - Experiment III

Stimulus Words and Their Meaningfulness and Imagery Ratings - Selected From Paivio, Yuille and Magigan (1968).

	M	I
artist	6.65	5.93
author	5.24	4.53
bravery	6.44	4.40
captive	5.39	5.27
colony	5.00	5.10
creature	5.50	4.60
demon	5.04	4.70
dweller	5.09	4.37
explanation	5.80	2.90
fantasy	5.06	3.70
genius	5.50	4.27
homocide	6.24	4.17
instructor	5.29	5.70
kerosene	6.04	5.77
madness	5.16	4.03
mirage	5.63	4.97
opium	5.88	4.60
prestige	4.56	3.67
retailer	6.76	4.33
sickness	6.36	5.13
socialist	4.54	3.13
utensil	6.08	5.47
venom	6.40	4.23
weapon	7.84	5.73
$\overline{X}$	5.77	4.61
s.d.	0.78	0.81

Word List - Experiment IV

Stimulus Words Selected from the Michigan Restricted Association Norms (Riegel, 1965).

anger	nephew
armadillo	ocean
baby	opal
barn	opium
doctor	orchid
hand	porthole
hawk	scissors
homicide	soldier
instructor	square
king	sun
knife	thief
lettuce	town
mattress	typhoon
moth	uranium
music	whiskey

Buffer items

* *

Primacy	Recency
foreigner	path
sheep	socialism

Word List - Experiment V

Stimulus Words Selected from the Michigan Restricted Association Norms (Riegel, 1965).

armadillo	ocean
baby	opal
barn	opium
doctor	orchid
hand	porthole
hawk	scissors
homicide	soldier
instructor	square
lettuce	thief
mattress	town
moth	typhoon
nephew	whiskey

Buffer Items

**

Primacy	Recency
music	sheep
sun	king

Definitions Used in Experiments IV and V

1. Anger is a strong emotion or feeling, like hate, where often the adrenalin runs high.
2. An armadillo is an animal with a hard shell-like body that likes to burrow in the ground and is usually found in South America.
3. A baby is a tiny infant who is cuddled and fed from birth by it's mother.
4. A barn is usually a wooden building located on a farm to store hay and provide shelter for animals.
5. A doctor is a qualified medical practitioner who tends to the needs of sick people.
6. The hand is that manipulative part of the body at the end of the arm having five fingers.
7. A hawk is a predatory bird with a strong beak and sharp talons that likes to prey on smaller animals.
8. Homicide, like murder, involves the killing of human beings.
9. An instructor is a trained person who earns his living by teaching certain skills and knowledge to students.
10. A king is a royal monarch whose nobility and sovereignty empowers him to rule over his subjects.
11. A knife is a cutting implement that usually has a steel blade with a sharp edge.
12. A lettuce is a leafy green vegetable that adds crispiness to any salad.
13. A mattress is a type of bedding or case stuffed with feathers and springs and upon which people sleep.
14. A moth is a small grey insect, like a butterfly, that likes to eat woollen clothes.
15. Music is the art of composing and combining sounds, or notes, in a harmonious way so that it may be played or sung according to some arrangement.

Continued...

16. A nephew is a relative, the son of either your sister or your brother.
17. An ocean is a great body of water that is salty and houses a variety of fish and planets.
18. An opal is a beautiful bluish-green stone that is found mostly in Australia.
19. Opium is a kind of drug made from the juice of the poppy flower, and is smoked or eaten as a stimulant.
20. An orchid is an exotic flower of brilliant colors, usually, white, red, or purple.
21. A porthole is a small round hole for looking out of the side of a ship.
22. Scissors are sharp steel instruments used for cutting paper or material.
23. A soldier is a military person who enlists in the Army in times of war and goes off to fight the enemy.
24. A square is a four-sided figure whose sides are all equal and whose angles are all ninety degrees.
25. The sun is a bright yellow star whose hot rays shine down and warm the planet Earth.
26. A thief is a type of criminal, like a robber, who breaks the law by stealing from other people.
27. A town is a large collection of houses or a community where a group of people live, work, and grow together.
28. A typhoon is a violent storm, like a hurricane, with winds in excess of 75 m.p.h.
29. Uranium is a radioactive white metal that can be used as a source of atomic energy.
30. Whiskey is an alcoholic spirit or drink that is usually distilled from grains like corn or rye.

Sentences Used in Experiments IV and V

1. The <u>anger</u> of the crowd grew in intensity over the referee's decision.
2. The <u>armadillo</u> scurried down its burrow in an effort to escape its pursuer.
3. The noisy <u>baby</u> cried all night long for its bottle.
4. The animals were kept in the <u>barn</u> overnight when the temperature dropped below zero.
5. The <u>doctor</u> came at once in response to the emergency call.
6. He extended his <u>hand</u> in a gesture of friendliness toward the newcomer.
7. The <u>hawk</u> swooped on the little field mouse and carried it away.
8. The detective in charge of <u>homicide</u> expected an early arrest.
9. The <u>instructor</u> spent extra time tutoring the slower students in his statistics class.
10. By the <u>king</u>'s royal authority the man was executed.
11. The butcher cut his finger on the sharp <u>knife</u>.
12. The <u>lettuce</u> leaves were greatly appreciated by the pet rabbit.
13. The customer purchased the <u>mattress</u> at Eaton's for her new bedroom.
14. The <u>moth</u> was attracted by the light of the fire and flew to its death.
15. The haunting melody of the <u>music</u> drifted across the water to the shore on the other side.
16. The <u>nephew</u> went duck shooting with his favorite uncle.
17. The mood of the <u>ocean</u> changed from day to day throughout the voyage.
18. The prospectors spent weeks digging for precious <u>opals</u>.
19. The illicit importation of <u>opium</u> into North America has increased noticeably over the last decade.
20. The young man presented his date with a beautiful <u>orchid</u> as a corsage.
21. The angry sea came crashing in through the broken <u>porthole</u>.
22. The dressmaker was lost without her favorite pair of <u>scissors</u> for cutting out her patterns.
23. The <u>soldier</u>'s gun and uniform were wet from wading across the swollen streams.
24. The old peasant went to do her shopping in the village <u>square</u>.

Continued...

25. The fierce desert sun took its toll of the weary travellers.

26. The thief broke into the safe and escaped with the jewels.

27. The farmer visited the town once a week to shop and carry out his business.

28. The typhoon left a trail of destruction and desolation in its wake as it lashed the tiny coastal village.

29. People are not yet sure of the potential radioactive threat posed by the current disposal methods of uranium waste.

30. The alcohol level of the whiskey is too much for the novice drinker.

Words and Questions Used In Experiments IV and V

Stimulus Word	Free Associate	Superordinate Category
ANGER	hate	a type of emotion
ARMADILLO	tortoise	a type of animal
BABY	child	a type of human being
BARN	cow	a type of building
DOCTOR	nurse	a type of occupation
HAND	finger	a part of the body
HAWK	wing	a type of bird
HOMICIDE	murder	a type of crime
INSTRUCTOR	professor	a type of occupation
KING	queen	a type of ruler
KNIFE	fork	a type of utensil
LETTUCE	salad	a type of food
MATTRESS	bed	a type of furniture
MOTH	butterfly	a type of insect
MUSIC	note	a type of sound
NEPHEW	uncle	a type of relative
OCEAN	water	a type of natural earth phenomenon
OPAL	diamond	a type of precious stone
OPIUM	drug	a type of plant
ORCHID	rose	a type of flower
PORTHOLE	ship	a type of window
SCISSORS	paper	a type of tool
SOLDIER	war	a type of person
SQUARE	circle	a type of shape
SUN	moon	a type of star
THIEF	robber	a type of criminal
TOWN	city	a type of community
TYPHOON	hurricane	a type of weather phenomenon
URANIUM	iron	a type of metal
WHISKEY	drink	a type of drink

Word List - Experiment VI

Stimulus Words and their Meaningfulness and Imagery Ratings - Selected From Paivio, Yuille and Madigan (1968).

		'm'	'I'
animal		7.00	6.10
bandit		6.92	5.83
blacksmith		7.44	6.17
breeze		7.33	5.87
butterfly		7.80	6.63
cotton		7.13	6.00
disease		7.44	4.87
doctor		7.32	6.40
fisherman		7.84	6.50
forest		9.12	6.63
golf		8.16	6.70
horse		8.67	6.80
infant		7.20	6.33
injury		7.32	5.70
judge		7.04	6.27
lord		6.68	4.63
meadows		8.00	6.43
ocean		8.76	6.77
pollution		7.40	4.63
pudding		7.31	5.27
sultan		7.24	5.57
tobacco		7.84	6.27
troops		7.54	6.13
vessel		7.82	5.90
	$\bar{X}$	7.60	6.06
	s.d.	.78	.62

APPENDIX G:

Instructions

EXPERIMENT I

SPECIFIC INSTRUCTIONS GIVEN PRIOR TO RATING TASK

"There is a theory in human information processing that goes something like this: If you are given a task to do with respect to a particular word your memory for that word will vary depending on the nature of the task. For example, if I asked you whether each word in a list of words had the letter "e" in it as opposed to asking you whether it rhymed with another word, the theory might suggest memory for the word would be better following the seoond task because somehow it required more conscious attention or a deeper kind of analysis in terms of the mental operations involved. Similarly, if I asked you to judge whether a word was in upper or lower case letters versus, say estimating how pleasant the sound of the word was, which do you think might require a deeper kind of processing in order to carry out the task?

I'm going to show you some tasks that all have one thing in common - they all involve thinking in some way about some aspect of a word's meaning. Some of the tasks could be carried out with a minimum of conscious effort or attention while others might cause you to direct your attention to deeper or more important features of the word. Some of the tasks will involve your relating to two words in order to carry out the task. I still want you to keep the same criterion in mind. Which task requires more conscious attention or effort for you to carry out.

I don't want you to feel that the time it takes for you to do the task is the important thing. People doing research in the area seem to think that while time is obviously an important factor, it is what you are doing with the time you have that is most important. So you could say some of these tasks are qualitatively different from others in that they require you to process, if you like, deeper aspects of the word.

We will have some practice, first of all, doing the tasks (E gives out practice sheets) so that you can get a feel for what is involved in carrying out each task. Again, try not to think in terms of how hard or easy the task is or how much time it takes. Work through the sheet with these words in mind (E writes several practice examples on the board, e.g., pliers - crab).

. . . .

All right, now I'm going to give you a booklet that has each task paired with all other tasks but in scrambled order. I want you to work through the booklet and put a cross beside the task in each pair that requires more conscious attention or effort to carry out. If you are not sure about a particular pair I don't want you to look back to earlier ratings - just try to keep the criteria we've talked about in mind. That is, which task directs your attention to deeper or more important features of the word or words?

Are there any questions? All right, you may begin now."

EXPERIMENTS II AND III

GENERAL INSTRUCTIONS

"One of the purposes of this study is to collect information on how people use words. As you know language is continually changing and we are interested in seeing how people perceive certain aspects of words today.

A series of slides will be shown to you with English words printed on them. Each of you has a booklet face down in front of you containing instructions telling you what to do as each word (two words*) is presented. Remember your task will always be the same for each word (two words*) although other people in the room may be performing different tasks. You can turn over your booklet now and read your instructions...

Are there any questions?

To give you an idea of what your task entails and particularly to show you that you will not have a great deal of time to carry out the task for each word (two words*) we will have a practice slide followed by a blank. During this interval I will stop the projector if anyone is uncertain as to what to do or has any questions. Remember we will begin with the practice slide followed by the blank interval. Following that there will be the series of slides for 10* seconds while you carry out your task. When this part of the experiment is over there are one or two other things we have to do before we are finished.

Are you all ready?"

*The time and task appropriate to each condition is stressed here.

At completion of the presentation of the slides and after an interval of 2 1/2 - 3 minutes the subjects were issued with written free recall instructions (see Part 3 of this Appendix).

EXPERIMENTS II AND III

2. SPECIFIC INSTRUCTIONS RELATIVE TO EACH SEMANTIC PROCESSING TASK

30 Second Tasks

1. SIMILARITIES

Your task is to write a sentence explaining how the two words that appear on the screen might be similar in meaning. Try to make the sentence meaningful or creative in some way.

e.g., if the two words are scavenger and stubbornness you might write:

A similarity between scavenger* and stubbornness* is that both are persistent kinds of things.

*be sure to write the 2 words.

A sentence like "A similarity between scavenger and stubbornness is that they are both nouns or they both have 3 syllables." does not really show possible similarities in meaning.

You will have 30 seconds to make up and write down the way in which the 2 words are similar in meaning and then turn the page in your booklet in readiness for the next two words. If you are uncertain about this procedure or have any questions please raise your hand.

2. DIFFERENCES

Your task is to write a sentence explaining how the two words that appear on the screen differ in meaning. Try to make the sentence meaningful or creative in some way.

e.g., if the two words are meteor and purse you might write

A difference between meteor*and purse* is that purse is a container for money and meteor is a flaming shooting rock.

*be sure to write the 2 words.

A sentence like: A difference between meteor and purse is that meteor has 3 syllables and purse has 1. does not really show possible differences in meaning.

You will have 30 seconds to make up and write down the way in which the 2 words differ in meaning and then turn the page in your booklet in readiness for the next 2 words. If you are uncertain about this procedure or have any questions please raise your hand.

3. 2-WORD SENTENCE

Your task is to make up a single sentence using both words that appear on the screen. Try to make the sentence as meaningful as you can.

If the two words are humility and escapade you might write:

The famous explorer exhibited unusal humility* over his latest escapade*.

Be sure to write each word in your sentence. Do not point out things that the two words have in common, e.g., "Both words are nouns" or "Both words have three syllables." Instead, try to construct a meaningful sentence.

You will have 30 seconds to make up and write your sentence for each word pair and turn the page in your booklet in readiness for the next two words. If you are uncertain about this procedure or have any questions please raise your hand.

20 Second Tasks

4. DEFINITION

Your task is to define each word as it appears on the screen. e.g., if the word is alimony you might write:

Alimony* is a payment of money by a man to his ex-wife for support.

*be sure to write each word in your definition.

Often a definition calls for more than just using it in a sentence, e.g., for alimony you might write:

The man paid alimony every month.

but this doesn't really say what alimony is.

Similarly, a definition for robot would be something like:

A robot is a mechanical creature that may have some of the same appearances and functions of man, i.e., try to include important features and characteristics in your definition.

You will have 20 seconds to make up and write down your definition for each word, and turn the page in your booklet in readiness for the next word. If you are uncertain about this procedure or have any questions raise your hand.

5. 1-WORD SENTENCE

Your task is to make up a sentence using the word that appears on the screen. Try to make the sentence as meaningful as you can.

If the word is draught you might write:

The draught* coming through the open window made the room feel chilly.

*be sure to write each word in your sentence.

Sentences like: "A draught is a noun", or "draught has seven letters" would not be a very meaningful sentence.

You will have 20 seconds to make up and write down your sentence for each word and turn the page in your booklet in readiness for the next word. If you are uncertain about this procedure or have any questions raise your hand.

15 Second Task

6. 2 FREE ASSOCIATES

Your task is to free associate to each word, i.e., write down the first two words you think of when you look at the word as it appears on the screen.

If the word is harp and you think of musician and string then write the following on lines provided.

*harp - musician

*harp - string

*be sure to write the word each time. You will have 15 seconds to make up and write down your responses for each word and turn the page in your booklet in readiness for the next word. If you are uncertain about this procedure or have any questions please raise your hand.

10 Second Tasks

7. STRONG-WEAK

Your task is to decide if each word is, in your opinion, generally Strong or Weak.

As each word appears on the screen you are to write down your decisions as follows:

If the word is bridge and you generally regard it as something strong, then write the following on the line provided:

Bridge* is strong

*be sure to write down the word.

or if the word is apathy and you generally regard it as weak write down:

Apathy is weak

You will have 10 seconds to decide and write down your response for each word and turn the page in your booklet in readiness for the next word. If you are uncertain about this procedure or have any questions please raise your hand.

8. PLEASANT-UNPLEASANT

Your task is to decide if each word is, in your opinion, generally Pleasant or Unpleasant.

As each word appears on the screen you are to write down your decision as follows:

Honey is Pleasant*

*be sure to write the word

or if the word is horror and you generally regard it as unpleasant write down:

Horror is Unpleasant

You will have 10 seconds to decide and write down your response for each word and turn to the page in your booklet in readiness for the next word. If you are uncertain about this procedure or have any questions please raise your hand.

9. LIVING-NONLIVING

Your task is to decide if each word is a Living or Nonliving thing in the sense that it actually lives. Some words, e.g., generosity seem like living things since some people are generous by nature. Nevertheless, there is no such living thing as generosity.

As each word appears on the screen you are to write down your decision as follows:

*Tormentor is a Living thing

*be sure to write down the word

<u>or</u> if the word is <u>prejudice</u> write down

<u>Prejudice is a Non-Living thing</u>

You will have 10 seconds to decide and write down your responses for each word and turn the page in your booklet in readiness for the next word. If you are uncertain about this procedure or have any questions please raise your hand.

EXPERIMENTS II AND III

3. FREE RECALL INSTRUCTIONS

Now that you have finished the first part of the experiment I would like you to try to recall as many of the words I presented to you on the screen.* Write down as many words as you can remember. Try to guess if you can't remember them all. Remember it is not important what order you write them down in just so long as they were the words you saw on the screen.* You will have approximately 3 minutes for this task. If you are uncertain about the procedure or have any questions please raise your hand.

______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________

*Instructions were altered slightly here for Experiments IV and V to read:

".....words I presented to you at the top of each page...so long as they were the words you saw at the top of each page."

EXPERIMENTS IV AND V

2. Specific Instructions for Each Semantic Processing Task

NEPHEW

A nephew is a relative, the son of either your sister or your brother.

GOOD DEFINITION — POOR DEFINITION

NEPHEW

The nephew went duck shooting with his favourite uncle.

GOOD SENTENCE — POOR SENTENCE

NEPHEW

The first word that comes to mind when you see the above word could be UNCLE

LIKELY — UNLIKELY

NEPHEW

The word above is a type of relative. It is a

GOOD EXAMPLE — POOR EXAMPLE

NEPHEW

The word above is an example of a

LIVING THING — NONLIVING THING

APPENDIX H:

Raw Data

EXPERIMENT II

FREE RECALL SCORES

S #	Living Nonliving	Pleasant Unpleasant	Strong Weak	2 Free Associates	1-Word Sentence	2-Word Sentence	Definitions	Differences	Similarities
1	5	7	6	3	6	11	8	6	7
2	6	11	5	10	11	8	8	10	17
3	4	7	9	13	11	14	9	12	5
4	3	7	8	7	12	7	12	12	12
5	3	3	4	12	12	12	9	13	8
6	5	11	9	3	7	12	8	7	2
7	6	5	9	8	13	19	4	6	8
8	4	3	6	8	9	13	8	14	6
9	5	12	10	6	6	10	10	13	4
10	9	11	6	6	10	9	5	9	7
11	6	10	3	6	17	13	5	11	6
12	5	9	14	8	9	14	4	8	6
13	7	2	4	10	11	12	11	9	5
14	7	4	10	7	9	9	8	5	5
15	3	8	5	7	10	10	12	6	4

EXPERIMENT III

FREE RECALL SCORES

S #	Living Nonliving	Pleasant Unpleasant	Strong Weak	1 Free Associate	1-Word Sentence	2-Word Sentence	Definitions	Differences	Similarities
1	7	5	11	5	10	7	12	12	9
2	9	9	9	11	5	14	6	11	18
3	8	5	10	3	9	14	11	10	11
4	11	12	7	8	11	18	14	12	9
5	10	3	8	7	6	12	13	10	12
6	8	8	4	11	15	4	17	8	14
7	3	7	9	3	8	13	14	14	17
8	4	9	8	13	7	8	15	12	10
9	2	9	7	9	12	10	14	7	10
10	5	9	7	7	6	8	10	11	16
11	8	8	8	10	8	12	9	13	15
12	6	6	10	11	5	12	9	10	13
13	2	5	8	8	9	7	7	12	8
14	7	8	7	9	9	9	13	11	5
15	6	10	6	6	4	10	10	10	10

EXPERIMENT IV

Mean Recall Scores

Subject No.	Living Nonliving	Category Judgment	Free Associate	Definition Rating	Sentence Rating
1	1	2	1	2	1
2	3	2	2	2	0
3	3	3	4	1	2
4	2	1	4	4	0
5	3	3	1	1	0
6	3	2	2	3	2
7	3	2	2	2	1
8	3	4	3	4	1
9	0	1	0	0	0
10	2	2	1	3	0
11	1	1	3	2	0
12	5	2	5	1	1
13	3	2	3	2	2
14	3	4	3	3	1
15	2	2	5	3	2
16	3	2	2	3	2
17	4	3	4	2	4
18	3	2	1	0	3
19	4	2	2	3	0
20	3	3	2	1	0
21	2	0	2	1	0
22	2	2	1	0	4
23	3	1	2	2	0
24	5	3	4	1	2
25	3	3	4	1	0

Continued...

Subject No.	Living Nonliving	Category Judgement	Free Associate	Definition Rating	Sentence Rating
26	3	0	3	3	0
27	4	0	2	0	2
28	4	1	1	0	1
29	2	3	2	3	2
30	3	0	2	3	2
31	4	2	3	3	1
32	4	0	2	4	0
33	3	3	3	1	2
34	3	2	4	4	1
35	2	2	1	1	0
36	1	0	2	2	0
37	3	3	2	2	0
38	3	1	4	3	2
39	3	1	3	0	2
40	3	1	2	2	2
41	1	2	1	2	0
42	4	0	1	1	2
43	2	1	2	3	3
44	2	1	1	1	0
45	3	2	3	4	1
46	2	1	5	3	3
47	3	2	2	3	0
48	1	3	3	1	0
49	2	3	3	3	2
50	3	5	4	2	2

EXPERIMENT V

Free Recall Scores and Processing Times

S #	Living Nonliving	Total Time (sec)	Category Judgment	Total Time (Sec)	Free Associate	Total Time (sec)	Definition Rating	Total Time (Sec)	Sentence Rating	Total Time (Sec)
1	14	111	6	230	8	180	10	245	6	285
2	10	95	16	150	5	258	6	208	7	375
3	8	92	11	100	4	252	13	210	2	300
4	12	145	6	195	12	254	14	212	5	417
5	12	130	14	141	3	175	9	208	5	236
6	13	120	10	175	12	210	8	215	1	297
7	5	82	10	175	9	185	14	413	4	235
8	7	125	11	150	9	188	11	390	3	195
9	9	91	10	160	10	150	13	264	11	210
10	12	110	17	183	8	250	10	275	11	241
11	10	138	10	272	12	252	16	175	9	225
12	10	143	12	248	12	198	11	205	4	215
13	5	95	15	258	8	190	12	205	4	320
14	11	206	12	217	9	290	13	330	7	223
15	9	200	13	205	10	161	8	325	5	210

EXPERIMENT VI

Free Recall Scores

S#	Living Nonliving	1-Word Sentence	Definitions	Similarities
1	13	12	19	21
2	13	11	11	13
3	14	14	11	14
4	14	13	17	16
5	16	11	13	20
6	13	20	17	11
7	14	10	12	10
8	6	12	12	18
9	14	14	19	16
10	8	12	15	12
11	12	13	10	15
12	11	18	12	20
13	12	12	15	15
14	13	8	14	15
15	13	14	14	8
16	14	16	9	18
17	12	12	16	5
18	10	14	11	10

APPENDIX I

Analyses of Variance

EXPERIMENT II

Analysis of Variance of Mean Levels of Recall as a Function of Type of Semantic Processing with Low 'm' Stimuli

Source of Variation	df	MS	F
Between Groups	8	58.21	6.65
Error	126	8.75	
Total	134		

.01 F(8, 120) = 4.95

EXPERIMENT III

Analysis of Variance of Mean Levels of Recall as a Function of Type of Semantic Processing with Medium 'm' Stimuli

Source of Variation	df	MS	F
Between Groups	8	58.97	7.31
Error	126	8.06	
Total	134		

.01 F(8,120) = 4.95

EXPERIMENT IV

Analysis of Variance of Mean Levels of Recall as a Function of Type of Semantic Processing with High 'm' Stimuli

Source	df	MS	F
Between Tasks	4	17.88	15.68
Between Subjects	49	2.15	
Error	196	1.14	
Total	249		

.01 F(4,120) = 13.6

EXPERIMENT V

Analysis of Variance of Mean Levels of Recall as a Function of Type of Semantic Processing with High 'm' Stimuli

Source of Variation	df	MS	F
Between Groups	4	85.62	10.1
Error	70	8.42	
Total	74		

.01 F(4,60) = 13.7
.05 F(4,60) = 5.69

EXPERIMENT V

Analysis of Variance of Times Required as a Function of Type of Semantic Processing with High 'm' Stimuli

Source of Variation	df	MS	F
Between Groups	4	48,663.25	16.02
Error	70	3,037.69	
Total	74		

.01 F(4,60) = 13.7

EXPERIMENT VI

Analysis of Variance of Mean Levels of Recall as a Function of Type of Semantic Processing with High 'm' Stimuli

Source of Variation	df	MS	F
Between Groups	3	19.79	2.19
Error	68	9.04	
Total	71		

.10 F(3,60) = 5.15
.25 F(3,60) = 2.47