1981

Role of Rating Value of Words and Displaced Rehearsal in Semantic Rating Tasks

Manju Karmeshu
Wilfrid Laurier University

Role of rating value of words and displaced rehearsal in semantic rating tasks

by

Manju Karmeshu

A thesis
submitted to the Department of Psychology
in partial fulfillment of
the requirements for the
Degree of Master of Arts.

Wilfrid Laurier University
Waterloo, Ontario,
CANADA

1981

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my supervisor Dr. Keith Horton for his valuable guidance and encouragement during the course of this research. Thanks are also due to Dr. Angelo Santi for his constant advice and help, and to Dr. Robert St.Claire-Smith for his useful suggestions.

Abstract

Two experiments were carried out to investigate the role of rating task, displaced rehearsal, and the value of words on the retention of items processed at the semantic level. In Experiment I, the subjects rated medium value words for either pleasantness or frequency under free rehearsal conditions. The results revealed no retention differences between words judged for pleasantness and those judged for frequency. In Experiment II, high and low pleasantness and high and low frequency values were factorially combined within a single list of words. In the free rehearsal condition, words were judged on either the pleasantness or frequency rating dimensions. In the forced rehearsal condition I, subjects compared the target words with a pair of intralist comparison words. In the forced rehearsal condition II, subjects compared the target words with a pair of extralist words. Recall was found to be significantly higher for words judged on the pleasantness than the frequency dimension in all three rehearsal conditions. This indicated the superiority of the pleasantness rating task over the frequency rating task when high and low value words were used. Existence of retention differences between the pleasantness and frequency tasks in the forced rehearsal conditions I and II also indicate that factors other than

displaced rehearsal may be operating. Better recall of high pleasantness words than the low pleasantness words for low frequency value but not for high frequency value suggested that item properties (values) interact.

Table of Contents

Tables

Figures

INTRODUCTION

Craik and Lockhart (1972) described the levels of processing framework for the study of memory by proposing that memory traces are a by-product of the perceptual and cognitive operations performed on the stimuli. The durability of the trace was viewed as a positive function of "depth" of processing, where depth referred to a greater degree of semantic involvement.

Subsequent experiments explored the levels of processing framework in greater analytic detail. Craik and Lockhart (1972) operationalized depth in terms of processing time on the assumption that the deeper the analysis, the longer it takes to carry it out (p. 676). However, Craik and Tulving (1975, Experiment 5) observed that the high retention of items processed at the semantic level did not depend simply on processing time. For the nonsemantic task in their experiment, the subjects judged the pattern of vowels and consonants which made up the word (for example, CCVVC was the sequence of consonants and vowels constituting the word 'Brain'), whereas for the semantic task the subjects' task was to decide whether the word would fit the sentence: The man threw the ball to the ___ 'child'. Thus the comparison of a nonsemantic and difficult task with a semantic and easier one revealed higher retention for the latter even

though it took less time. The results led them to suggest that memory performance depends on the qualitative nature of the task, and processing time by itself is not a good predictor of retention; thus the latter was discarded as an index of depth. Furthermore, they found that "yes" responses to the semantic task questions led to better retention than "no" responses, especially at the deeper levels. Positive and negative decisions presumably required the same level of processing. Further analysis of this relationship led them to the conclusion that the encoding was richer or more elaborate when the to-be-remembered item was congruent rather than incongruent with the question. Thus, the first major change in the original levels of processing model was signalled by Craik and Tulving (1975) when they proposed that the data on depth of processing could be interpreted in terms of "the idea that memory performance depends on the elaborateness of the final encoding" (p. 291). The difference between spread (or elaboration) of encoding and depth is that depth implies that encoding operations are carried out in a fixed sequence from one level to the next and spread "leads to a more flexible notion that the basic perceptual core of the event can be elaborated in many different ways" (p. 291).

The earlier view of processing of a stimulus in terms of a continuum of analysing operations (Craik & Lockhart, 1972) was further modified by Lockhart, Craik, and Jacoby (1976).

According to the earlier view, the physical and structural features of a stimulus are analysed first, then the stimulus is subjected to progressively more elaborate semantic analyses. The phrase 'greater depth' referred to these later semantic, associative operations. Lockhart et al. (1976) proposed that physical, phonemic, and semantic characteristics of words exist in different dimensions or domains. 'Greater depth' may refer to two somewhat distinct changes in processing. "First, the domains themselves may be thought of as a hierarchical organization proceeding from shallow, structural domains to deep, semantic domains. Second, at one depth in this sense, the stimulus may be further analysed or elaborated by carrying out additional operations within one qualitatively coherent domain" (p. 78).

Though according to the notion of domains, processing typically proceeds through a fixed series of qualitatively distinct stages or domains, this does not mean that all possible analyses are laboriously carried out in each domain. Only those analyses required to provide critical evidence for deeper levels of processing are carried out. In general, processing proceeds until the domain relevant to the present task is reached and, quite often, it is only at the 'target' domain that sufficient processing occurs for conscious awareness of the results of the processing operations. The apparent automaticity of encoding depends

on the number of analyses which must be performed on the word before its meaning is extracted. This in turn depends on such factors as the materials, practice, context, and set (Lockhart et al., 1976, p. 79).

An important question that remained was whether, or to what extent, the variations in memory performance were attributable to distinctiveness or uniqueness of encoding operations. Moscovitch and Craik (1976) moved toward such a formulation in explaining the interactions they observed between encoding operations and the ratio of the to-be-recalled items (TBRIs) to retrieval cues. In their experiment 2, the encoding questions asked during the study trial were presented as the retrieval cues. Under the unique condition, each TBRI was associated with a different cue; under the shared condition, ten TBRIs were associated with each of six cues. When the cues were semantic (category names or sentence frames), cue sharing led to a pronounced decline in retention. By contrast, a shift from unique to shared cues had no effect on retention at the phonemic level. According to Moscovitch and Craik, cue overlap has little or no effect when the traces are inherently similar, as is presumably the case for words encoded in terms of a limited number of physical or phonemic features. Cue overlap becomes an effective variable for semantically encoded words when "the forms of encoding are virtually limitless and, speculatively, these semantic encodings are less overlapping

in their content than are physical and phonemic encodings" (p. 452). Therefore, items encoded to shallow levels would suffer relatively small decrements in memorability from the cue-sharing manipulation. On the other hand, the beneficial uniqueness of semantically encoded words should be affected to a larger extent.

Postman, Thompkins, and Gray (1978), however, argued against the explanatory principle of distinctiveness suggested by Moscovitch and Craik. They pointed out that, as each semantic trace, compared to each phonemic trace, has the benefit of uniqueness, the explanation offerred by Moscovitch and Craik does not show a strong association between depth and uniqueness as determinants of retention. Postman et al. (1978, Experiment 1B) demonstrated that when the pre-experimental relations between the cue words and the TBRIs were controlled, cue sharing significantly increased the amount recalled and the interaction failed to reach significance in both the semantic and the nonsemantic conditions. This finding is at variance with the hypothesis and results of Moscovitch and Craik (1976) that distinctiveness is beneficial only when processing is semantic. It is observed that under some conditions at least, distinctiveness can influence performance equally, both when processing is semantic and nonsemantic.

Jacoby and Craik (1979) further examined the role of distinctiveness and a broad notion of encoding distinctiveness was stressed. Distinctiveness was used to denote the extent to which a particular processing strategy differentiates an item or a set of items from other information in memory. Thus, "a chair is a chair, but it is equally a piece of furniture, a thing, a wooden artifact, and any number of other descriptions, depending on what the chair is to be distinguished from. Similarly, the meaning of a given word in a given context depends on distinctions that are to be conveyed by that word in that context" (Jacoby & Craik, 1979, p. 2). As developed by Jacoby and Craik, distinctiveness and depth are not unrelated, since encodings that stress word meaning should have greater potential for developing distinct codes than should encodings that stress structural or auditory features. However, distinctiveness also depends on previously encoded events, the elaborateness of the code developed, and the retrieval cues present at retrieval. Thus, distinctiveness is always relative to some particular set of conditions, and an encoding that is distinctive in one retrieval context may not be in another. In a related study, Begg (1978) found a higher level of recall following contrastive processing (e.g., if the word pair is 'Beer-Wine', subjects are asked to list features in which the items differ from each other), which could be attributed to differential trace

discriminability in response production. Response production is referred to as a stage of retrieval in which the memory trace contacted by cue serves as the source of overt response. He suggested that "features appropriate for a given study task are of little value for discriminating among items sharing those features but of considerable value for discriminating items possessing those features from items not possessing them" (p. 517). The suggestion that deeper codes are more discriminable, and that this greater distinctiveness is the crucial factor underlying superior retention, has also been put forth by Eysenck (1978).

Jacoby and Craik (1979) also proposed that a difficult initial decision will usually be associated with higher levels of retention since difficulty necessitates more extensive processing, which then results in the formation of a more distinctive trace. They examined the effects of initial decision difficulty on subsequent recall and recognition, while also varying the degree of association between the "decision" word and the word used later as the retrieval cue. The subjects were asked to study a 'focus' word printed on one side of the card and then pick out from the two words printed on the reverse side, that word which was more highly related to the focus word. The words on the reverse side were either high or low associates of the focus word. The combinations of the words formed on the reverse were High-High, High-Low, High-Unrelated, Low-Low,

Low-Unrelated, and Unrelated-Unrelated. Decision difficulty was assumed to depend on the relative degree of association of the two words to the focus word; thus, difficult decisions would be involved in the High-High, Low-Low, and Unrelated-Unrelated cases. The selected item was referred to as the 'target' word.

The results showed that both initial decision difficulty and the strength of pair association between the focus and target words had strong effects on retention performance. Further, each of those variables interacts with the form of the retention test. The effects of both decision difficulty and prior associative strength were more pronounced in cued recall than in recognition. Cued recall scores were higher than recognition scores for the highly associated focus-target words but this superiority of cued recall dropped for low associates and reversed for unrelated words. Jacoby and Craik suggested that "retention level is a function of both of the nature of encoding and of the effectiveness of the retrieval information to enable formation of mental operations that will match the trace" (p. 14). Decision difficulty is assumed to affect the distinctiveness of the encoded trace, but the ease with which the focus word can facilitate reconstruction of the focus-target complex is important too.

The concept of distinctiveness has been explored in several studies to explain its role within the levels of

processing framework but a consistent operational definition is still needed. Recent manipulations of distinctiveness have included the following: (a) the proportion of words in a list encoded at the same level (Craik & Tulving, 1975, Experiment 8); (b) the ratio of cues to responses in a list (Moscovitch & Craik, 1976, Experiment 2; Postman et al., 1978, Experiments 1A and 1B) (c) the degree of correlation between the attribute dimensions of words (Battig & Einstein, 1977; Klein & Saltz, 1976); (d) the number of words rhyming with the TBRI (Hunt & Mitchell, 1978, Experiment 1); (e) the typicality of orthographic structure (Hunt & Mitchell, 1978, Experiments 2-4); (f) the degree of initial decision difficulty (Jacoby & Craik, 1979, Experiment 1); (g) the number of decisions required for a word (Johnson-Laird, Gibbs, & deMowbray, 1978; Ross, 1981). Thus, the notion of distinctiveness has been manipulated in various ways by different researchers. A generally acceptable operational definition is yet to be specified.

In empirical tests conducted by Craik and his associates, the effect of depth was found to interact strongly with other variables (Craik & Tulving, 1975; Fisher & Craik, 1977; Moscovitch & Craik, 1976). This led them to progressively modify the original hypothesis and the emphasis shifted from depth per se to such characteristics as elaboration and distinctiveness. In addition, it was also felt that the earlier view of memory focused primarily on

the encoding processes operating at the time of input. Moscovitch and Craik (1976) noted that Craik and Lockhart's (1972) formulation lacked any clear hypothesis concerning retrieval. In order to overcome this limitation, and thereby extend the range of phenomena to which a "levels" approach could be applied, a series of experiments was conducted to explore the effects of retrieval factors on memory. Moscovitch and Craik emphasized that memory must be viewed as a joint function of stored information (memory trace) and information provided to the subjects at retrieval rather than just the level of processing. They found that a shift from free to cued recall led to much larger gains in retention for semantically than for phonemically processed items (Experiment 1). On the basis of their results, they suggested that the level of processing "may set an upper limit on recall and recognition; how near the subjects' performance approaches the upper limit for a given level of processing will depend on the effectiveness of the retrieval environment" (p. 450).

Further support for this conclusion was presented by Fisher and Craik (1977) who investigated the effects of the compatibility between encoding operations and retrieval cues on retention. They factorially varied the levels (semantic or phonemic) of the input and output cues. For example, in the case of rhyme encoding context, if the presented pair was HAIL(Pail), the identical cue was "rhymes with pail,"

the similar retrieval cue was "rhymes with bail," and the different cue was "associated with snow." If the encoding context was HAIL(sleet), the identical cue was "associated with sleet," the similar cue was "associated with snow," and the different retrieval cue was "rhymes with bail." The results showed a significant effect of depth of processing in that, at each level of similarity between encoding context and cue, the semantic encoding yielded a higher level of retention. With the level of encoding constant, recall was lower when the test cue was similar to the input cue than when it was identical with the input cue (Experiment 3). The results also showed that superiority of semantic over phonemic processing was greatest when identical cues were used which created optimal conditions of retrieval. This is in line with the encoding specificity effect (Tulving & Thomson, 1973).[1] On the basis of their results, Fisher and Craik emphasized that both the qualitative nature of the encoding and the degree of compatibility between the encoding and test cues are important to an adequate account of memory processes. The coherence or congruence of retrieval cues with the encoded trace was stressed in subsequent experiments by Jacoby and Craik (1979, Experiments 1-2).

In an analytic review of the levels of processing framework, Eysenck (1978) pointed out that the effects of encoding depth appear to be greater on tests of recall than

on tests of recognition. This interaction can be explained within the framework of Anderson and Bower's (1972, 1974) theory of free recall and recognition. According to their theory, free recall involves a retrieval component and a decision component, whereas recognition involves primarily a decision component. Eysenck argued that the small effects of depth on recognition test performance may be because "depth affects the retrievability of information more than the decision or recognition based upon retrieved or presented information" (p. 163). Lockhart et al. (1976) argued that recall and recognition involve processes that reflect different aspects of the same retrieval system. The same two modes of retrieval exist for both recall and recognition, namely, reconstruction and scanning. The difference is that recall is conceptualized as the guided reconstruction of the original encoding from the basic information provided by the retrieval information. On the other hand, recognition relies less on the reconstructive efforts of the system since more retrieved information is provided by the stimulus. Since recall and recognition reflect different questions being asked of the system, the beneficial effects of depth of encoding on tests of recall may be because deeper encodings are distinctive and unique. As a consequence, resulting episodic traces are more easily contacted and the richer information may then provide more adequate feedback to guide further reconstruction.

In summary, the version of the levels of processing view advanced by Craik and Lockhart (1972) has evolved and changed in many respects over the last several years. Modifications through empirical tests have emphasized the additional concepts of "elaboration" and "distinctiveness" of encoding. "Depth" has been used to denote qualitatively different encodings whereas "elaboration" refers to greater amounts of processing of the same general type. In other words, the latter often refers to the addition of further information, so that the trace becomes richer and more detailed. The depth to which a stimulus is processed, in conjunction with its degree of elaboration, gives rise to an encoding that is more or less "distinctive" and thus discriminable from other memory traces in the system. Distinctiveness has been considered as somewhat different from depth and elaboration; "Whereas the latter terms describe the operations carried out during encoding, distinctiveness describes the similarity of the product of these operations to other memory traces" (Craik, 1979, p. 449). Thus, in using "distinctiveness," the emphasis is laid on the contrastive value of information in the trace. Finally, the relationship between encoding and retrieval operations has also been strongly stressed in the revisions.

The original levels of processing framework, as formulated by Craik and Lockhart (1972), also claimed that the orienting task acts to select particular attributes of

an event for encoding. An attribute (e.g., the sound of a word) will be encoded only if the orienting task requires the subject to deal with that attribute. There is certainly a good deal of evidence to suggest that this original selective encoding position is too extreme; subjects either consciously or unconsciously encode attributes in addition to those required by the orienting task. Jacoby and Craik (1979) noted that even the original levels of processing experiments provide evidence that this is the case. Those experiments demonstrated that retention was higher after decisions about the meaning of a word (e.g., Does the word refer to an animal?) than after decisions about the physical characteristics of a word (e.g., Is the word in upper or lower case?). The important point was that retention in the conditions where subjects judged the "case" of presented words was substantially above zero. Jacoby and Craik pointed out that if, in making case decisions, subjects had encoded the words only in terms of whether they appeared in upper case or lower case, retention should have been essentially zero, because remembered information regarding case alone would be of no help in the later memory test. The non-zero level of retention provides evidence that information beyond the minimal amount necessary to accomplish the orienting task must have been accessed. Nelson (1977) gave details of several experiments that support this view. Other studies also found that information which does not appear to be

required by the orienting task, is nonetheless encoded (Bird & Roberts, 1980; Coltheart, 1977; Klein & Saltz, 1976; Nelson, 1979; Postman et al., 1978).

The levels of processing framework has been viewed critically by several investigators. Nelson (1977), Baddeley (1978), and Eysenck (1978) have pointed out that one major limitation is the absence of an independent measure of depth. The principle of 'spread of encoding' (Craik & Tulving, 1975) or 'further elaboration within an encoding domain' (Lockhart et al., 1976) also lacks operational definition and independent indices. However, Johnson-Laird et al. (1978) suggested that elaboration can be operationalized in terms of amount of semantic processing and conducted a series of experiments to support the notion. (These shall be discussed in detail below.) But so far, with the the exception of Nelson (1977), no operational definition of depth of processing has been offered by any investigator. Further experimental and theoretical analysis is needed in this direction. Despite these shortcomings, Jacoby and Craik, Nelson, Eysenck, and Battig (Cermak & Craik, 1979) agree with the basic idea that "input processing of an event can be elaborated to a greater or lesser degree; that the degree of elaboration depends on such factors as amount of practice, task-induced processing, and processing produced spontaneously by the subject; and that elaborative processing typically results in a

distinctive encoding. Further, if the appropriate information is provided (e.g., a cue) at retrieval, these distinctive encodings are associated with high levels of subsequent retention" (Craik, 1979, p. 447-448).

The levels of processing framework proposed by Craik and Lockhart (1972) has spurred new interest in semantic processing as a factor in memory, particularly with regard to free recall following incidental learning (i.e., subjects are not informed of the subsequent memory test). Their formulation has focused on studying memory differences between the levels of processing (semantic and nonsemantic). Little attention has been directed to the different types of operations involved within a processing domain (either semantic or nonsemantic) which may account for large retention differences. Some researchers have recently applied the ideas derived from the levels of processing framework and its subsequent modifications to explore differential memorability of items processed within the semantic domain. Efforts are being directed to investigating the memorial consequences of different operations to understand human memory and provide a sound data base that can serve as a foundation for a theory of memory. However, studies conducted so far in this direction have produced different results.

Investigators have attempted to explore the effect of 'elaboration' within a given level of processing. Some

studies have investigated differences by varying the number of ratings required for an item. Hyde (1973) included three semantic orienting tasks. Two groups of subjects rated items on one semantic scale, either pleasant-unpleasant or active-passive, while a third group rated items on both scales. The recall between the three groups did not differ. Klein and Saltz (1976) required subjects to rate fifteen animate and nine inanimate nouns on a single attribute dimension (pleasant-unpleasant, happy-sad, fast-slow), two moderately correlated dimensions (pleasant-unpleasant, fast-slow), or two highly correlated dimensions (happy-sad and fast-slow; pleasant-unpleasant and happy-sad). Recall was found to be better for words rated on two dimensions than for words rated on a single dimension. Within the former condition, higher recall was observed for words rated on moderately correlated than highly correlated dimensions. They suggested that recall would be better, the greater the specificity of a concept in the cognitive space. Cognitive space is defined as the set of dimensions on which a person can react to the stimuli (Saltz, 1971, p. 33). According to Klein and Saltz, the moderately correlated dimensions specified the encodings more precisely and distinctively in the "cognitive space" as compared to highly correlated dimension. This approach is complementary in many ways to the views of Anderson (1976), Anderson and Reder (1979), and Craik (1979) that a greater amount of elaboration yields superior memory performance.

Another factor responsible for differences in recall among tasks within the semantic and nonsemantic processing domains was suggested by Postman (1976; Postman & Kruesi, 1977). He pointed out that the amount of displaced rehearsal attendant upon the rating of different attributes may be one of the major factors responsible for the higher recall of words judged on pleasantness than the same words judged on the frequency dimension. In the Postman and Kruesi (1977) study, the subjects were asked to rate words for either pleasantness or frequency in either the semantic or phonemic domains. Higher recall was obtained for the items judged for pleasantness than for those judged for frequency. In order to account for these results, they suggested that pleasantness represents a subjective dimension where ratings of a particular word are based upon comparison with previouly rated words. Such comparisons entail displaced rehearsal and are conducive to the development of interitem associations. Frequency ratings, on the other hand, are presumed to represent more objective comparisons with information external to the list. The use of external anchors would curtail the number of intralist comparisons and, hence, the amount of displaced rehearsal. Thus, they offered an explanation for the retention differences between items judged for pleasantness and those judged for frequency in terms of variations in the amount of displaced rehearsal. However, Postman and Kruesi did not include any specific

test to evaluate the hypothesis of differential displaced rehearsal with the various rating scales.

Further evidence regarding the role of displaced rehearsal was reported by Shaughnessy (1979). In the incidental learning, the subjects were asked to rate 40 items on a 7-point scale. One group rated each item in terms of its association to a colour cue (with eight words corresponding to each of five colour categories: cued condition). The categories were blocked; that is, the instances of each category were presented in successive list positions. A second group rated each item in terms of a different concept label (noncued condition). For example, items "coffee" and "coal" were rated in terms of the concept "black" in the cued condition and the concepts "type of beverage" and "type of fuel" in the noncued condition. Higher retention and primacy effects within categories were obtained in the cued as compared to the noncued condition. Shaughnessy suggested that these results provide support for the hypothesis that displaced rehearsal occurs in situations where the task requires the comparison of a new item with an earlier one on the same dimension. The within-category primacy effect was seen as a result of items presented early within each category being accorded a greater amount of displaced rehearsal as they tended to be used more often (compared to later items within each category) as the basis for interitem comparisons for items presented later in the

category. No such primacy effect was found in the noncued condition when a different concept label was used for the rating of each item.

Shaughnessy's experiment provides strong support for a displaced rehearsal hypothesis but, again, displaced rehearsal was not directly monitored. More importantly, his work does not demonstrate displaced rehearsal differences within the pleasantness and frequency rating tasks used by Postman and Kruesi (1977).

A more direct assessment of the use and effect of displaced rehearsal was carried out by Walther and Horton (Note 2). Subjects engaged in either pleasantness or frequency ratings of words at either the phonemic or semantic level. Two rehearsal conditions were used. In the free rehearsal condition, subjects rated the items individually whereas, in the forced rehearsal condition, the rating of the target word was made in the context of the rating given to the immediately preceding word. The results showed that under the free condition, superior retention was observed for items judged for pleasantness than for frequency in the semantic task. These results were in agreement with Postman and Kruesi's (1977) findings. However, when the subjects were forced to engage in displaced rehearsal (in the forced condition), retention for items judged for frequency increased to the level of those judged for pleasantness. The rehearsal condition

manipulation had no effect on the retention of items judged for pleasantness. Similar effects were obtained for the phonemic condition. Thus, they demonstrated that when items were rated on the basis of interitem comparisons, the retention advantage for items judged for pleasantness was no longer evident.

Another experiment that demonstrated superior memory for words processed for pleasantness as compared to a variety of other semantic processing dimensions was conducted by Packman and Battig (1978). They compared free recall and recognition memory following processing on the seven semantic dimensions of concreteness, imagery, categorizability, meaningfulness, familiarity, number of attributes, and pleasantness. The subjects rated 50 words representing two different levels (high and low) of scale values across all seven of these dimensions. Both recall and recognition scores were higher for pleasantness than for all other six dimensions. A comparison of retention for high and low value words showed overall recall performance to be nearly twice as high for the high as for the low value words.

In a subsequent experiment, Weiss, Packman, and Battig (cited in Packman & Battig, 1978) found that the superior recall for words judged on pleasantness can be eliminated when exclusively neutral rather than pleasant or unpleasant words are used. On the basis of the Weiss et al. data,

Packman and Battig (1978) suggested that "the type(s) of word(s) may be an important determinant of any memory superiority for pleasantness over other types of semantic processing with highly pleasant or unpleasant ratings representing more distinctive encodings than neutral ratings" (p. 506). In suggesting this hypothesis, they point out further that the factors underlying the superior memory and/or greater distinctiveness produced by pleasantness encodings remain to be elucidated by further research.

In a recent study, Karmeshu and Horton (Note 1) have obtained findings very similar to those of Weiss et al. They investigated the effects of depth of processing and displaced rehearsal and adopted a procedure similar to that used by Walther and Horton (Note 2). The subjects judged the pleasantness or frequency of words within the semantic or phonemic domains under free (rated independent words) or forced rehearsal condition (rated target words in comparison with another word paired with it). The free condition was included to replicate the results obtained in the Walther and Horton study. Notably, the forced condition differed from theirs in the sense that Walther and Horton had required the subjects to compare each item with a previously rated item in order to enhance displaced rehearsal. Karmeshu and Horton, in the forced condition, induced the subjects to make judgements on the basis of external

referents. It was hypothesized that the use of external anchors should curtail the number of intralist comparisons and hence reduce the amount of displaced rehearsal. Accordingly, it was predicted that the recall of words judged for pleasantness would decline, whereas the recall of words judged for frequency would remain unaffected, as these judgements are normally made on the basis of extralist comparisons (Postman & Kruesi, 1977).

Karmeshu and Horton (Note 1) found that, in the free rehearsal condition, the retention of words judged on pleasantness did not differ significantly from the recall of the same words judged on frequency. These results differed from those obtained by Walther and Horton but were consistent with Hyde and Jenkins (1973). Hyde and Jenkins compared five orienting tasks, two semantic and three nonsemantic. The semantic orienting tasks involved pleasantness and frequency ratings. They found no significant differences in the recall of words judged on these two dimensions for unrelated lists. One possible reason for the differences between the Walther and Horton and the Karmeshu and Horton results may be the nature of the specific words selected. In the Walther and Horton study, there was a possibility that the words rated by the subjects were either highly pleasant or unpleasant whereas in the Karmeshu and Horton study, the words were mostly neutral. In order to test this interpretation of the conflicting

findings, research is currently underway to ascertain pleasantness ratings for the items used in the two studies.

In the forced condition of the Karmeshu and Horton study, retention for words judged for pleasantness as well as frequency dropped markedly. This could be due to extensive interference caused by the extralist items as each target item was paired with a different extralist word.

Differences within a level have been investigated in a few studies by involving a more subtle manipulation of processing, namely specificity. Frase and Kammann (1974) had subjects search lists for instances of a general category (e.g., foods) or a more specific category (e.g., vegetables). In all three of their experiments, free recall was higher for the words searched for within the more specific categories. Bock (1976) included a similar manipulation and obtained the same results.

Recently, Johnson-Laird and his associates (Johnson-Laird & Bethell-Fox, 1978; Johnson-Laird et al., 1978) proposed the concept of amount of processing as an unambiguous measure of elaboration within a processing domain. They defined amount of processing as "the number of decisions about an item that yield pertinent information to the task at hand" (p. 342). They further suggested that the amount of processing involved in a task determines its memorability on the assumption that the more components a word has in common with the target category, and hence the

more components that may have to be checked, the better it will be remembered. The subjects were asked to classify words as denoting either a positive or negative instance of the category of substances that were consumable, solid, and natural. The words were of four sorts: target words with three components (e.g., 'apple'), words with two of the components (e.g., 'coal'), words with only one of the components (e.g., 'sweat'), and words with none of the components (e.g., 'paraffin'). The words were balanced for length and for frequency of use from the Kučera and Francis (1967) norms. The results were as predicted: Overall recall was highest for words with three components followed by those with two components, which was followed by those with one component. The lowest recall was for the words with no components of the target category.

Although the notion of amount of processing does not distinguish between levels of processing (Johnson-Laird et al., 1978, p. 374), Johnson-Laird and his associates believed that it accounts for those findings that led the proponents of the depth of processing framework to introduce their additional concept of the "elaboration" of an encoding (Craik & Tulving, 1975). 'It has been shown experimentally that items responded to positively are better remembered than items responded to negatively: For example, when subjects are asked "Is a shark a type of fish?," and "Is a heaven a type of fish?" they remember "shark" better than

"heaven" (cf. Shulman, 1974). The notion of 'elaboration' suggests that where a positive response is made, the encoding question and the target word form a more coherent and integrated unit' (Johnson-Laird et al., 1978, p. 374). Johnson-Laird et al. claimed that elaboration can be operationalized in terms of amount of relevant processing. Deciding about semantically close conjectures (e.g., a whale is not a fish) would require a greater amount of semantic processing (number of property decisions) than would deciding about a semantically remote conjecture (e.g., a heaven is not a fish). Since whale and fish have so many preexisting relations, the trace would be far more elaborated than would the heaven-fish trace. "Merely recollecting that a whale swims in the sea does not suffice, one must recall that it has no gills and must surface to breathe, and so on, whereas any aspect of heaven suffices to reject it. Hence, subjects should remember 'whale' better than 'heaven' after they have rejected both of them as varieties of fish" (Johnson-Laird et al., 1978, p. 375).

Ross (1981) employed the notion of amount of processing, as operationalized in the Johnson-Laird et al. procedure, to investigate memorability within a level. He investigated whether the three variables processing time, retrieval aids, and the number of positive decisions could be confounded with the number of decisions explanation. To examine the possibility that processing time rather than number of

decisions could be the main determinant of memory performance, he used (Experiment 1) a procedure similar to that of Johnson-Laird et al., except reaction times were taken for each item rather than for the whole list. Words were selected from 8 categories defined by the combinations of consumable or nonconsumable, liquid or solid, and natural or artificial properties. Words again varied in terms of the number of target properties exhibited: 0, 1, 2, or 3. The subjects responded 'Yes' if an item contained all the three properties and 'No' if it contained only 0, 1, or 2 properties. The results showed that the more target properties a word had, the more likely it was to be recalled. Within a given property level, longer processing time was not generally associated with greater recall. Thus, the results replicated the findings of Johnson-Laird et al. that, while the reaction times within a property level were not predictive of memory performance, number of decisions within a property level was a major determinant of recall.

Ross (1981, Experiment 3) also examined whether the variables 'retrieval aid' or 'number of positive decisions' were confounded with the number of decisions in the Johnson-Laird et al. experiments. The retrieval aid explanation proposes that since subjects have classified items using three properties, these properties may be used at retrieval, either singly or in combination, to aid recall. For example, subjects searching a list of words

having the three properties of being consumable natural solids (e.g., apple) would use the three properties as retrieval cues at the time of recall. Subjects may initially use each property separately to try to recall and, then, as the single cue's effectiveness wanes, they may use a combination of two or three properties as retrieval cues.

The number of positive decisions hypothesis claims that only properties that match target properties will improve the memorability of an item. Ross included two target properties instead of three. Subjects were divided into three groups. Subjects in the "All" group responded 'yes' if an item had all the target properties and 'no' otherwise. Subjects in the "Any" group responded 'yes' if the item contained any of the two properties and 'no' if it contained none of them. Subjects in the "Each " group recorded two separate responses for each item, a yes-no decision about each of the two properties. The recall results were again very similar to those predicted by the number of decisions hypothesis and were inconsistent with the number of positive decisions and retrieval aid hypotheses. As the number of target properties was increased, recall increased in the All group, decreased in the Any group, and remained relatively constant in the Each group. Ross interpreted his results as providing strong support for the notion of the number of (semantic) decisions about a word as a major determinant of its memorability by unconfounding the number of decisions

idea from the other hypotheses. He further suggested that the number of decisions made about a word may be interpreted as corresponding most closely to elaboration and indeed can be used as an index of elaboration. In fact, Ross suggested that "the number of decisions idea is preferable to the elaboration idea because of its greater specificity" (p. 31). According to this hypothesis, the more properties of an item that are checked, the more links or associations will be formed with the context (target categories) and, hence, the more likely it will be recalled. Thus the number of decisions made about an item affects later recall by increasing the number of strengthened connections between context and the item.

The hypothesis of number of decisions proposed by Johnson-Laird and his associates has been further explored by McClelland, Rawles, and Sinclair (1981) to investigate whether a retrieval explanation could yet be an alternative hypothesis. In the Johnson-Laird et al. experiments, the target categories used were the extreme properties of a word like consumable-nonconsumable and liquid-solid. By deciding that an item does not have the component "consumable," a subject might note that it is, by default, "nonconsumable." Similarly, an item that is not "solid" is "liquid." Thus the subjects could remember the four components and then use them as retrieval cues at the time of free recall. McClelland et al. manipulated the nature of the search task

and the recall test. They used sixteen semantically independent categories (Experiment 2) rather than four semantically related categories (Experiment 1). One group received target-search instructions (the criterion used by Johnson-Laird et al.), under which subjects searched for target items. They were asked to put a checkmark by category members and a cross by non-members. The other group received component-search instructions. For example, if the category used was "consumable solid," then examples of the two, one, and zero component words might be "BREAD," "SHERRY," and "SHAMPOO." In this situation, subjects were asked to write a 2, 1, or 0 beside each word in the list to indicate of the number of semantic components each word had in common with the target category. They found that cued recall was superior to free recall under both sets of instructions. An interaction between number of components and type of retention test was also observed. A comparison of recall proportions under target-search conditions revealed that in comparison with 1 and 0 component words, 2 component items were associated with higher scores than would be expected solely on the basis of number of encodings formed during classification. Retention differences between 2 component words with 1 and 0 were not as large in the component-search condition as those observed in the target search condition. The retention of 2 component words in the component-search condition was lower than the recall of 2 component words in the target-search condition.

These results suggest that, when access is readily gained to the retrieval cues which are then used in free recall situations (Johnson-Laird et al., 1978; Ross, 1981; McClelland et al., 1981, Experiment 1), number of decisions is a potent variable affecting performance. However, number of decisions is no longer an effective variable when access to retrieval cues is not easy (McClelland et al., 1981). Thus, access to retrieval cues is a very critical overriding variable in the free recall conditions used by these investigators.

In reviewing the concepts of displaced rehearsal and number of decisions, it may be suggested that the two are not entirely different. In fact, the number of decisions may be used as an index of the amount of displaced rehearsal whenever the decisions involve intralist comparisons exclusively.

Now the question arises, can the concepts of displaced rehearsal or value of words be used to account for the memory differences observed by earlier investigators using different semantic rating scales. Like Postman and Kruesi (1977), Walther and Horton (Note 2) found higher recall of words following a pleasantness rating task than a frequency rating task. Packman and Battig (1978) have demonstrated higher recall for pleasantness over imagery, concreteness, familiarity, categorizability, meaningfulness, and number of attributes whereas Weiss, Packman, and Battig (cited in

Packman & Battig, 1978) have reported no superior memory for pleasantness. Moreover, Hyde and Jenkins (1973) and Karmeshu and Horton (Note 1) found no superior memory for pleasantness. Thus, conflicting results have been obtained in these studies. In the proposed research, a further attempt is made to ascertain whether the retention differences can again be obtained for words judged on pleasantness and frequency rating dimensions and, if so, whether the two concepts mentioned above can be used to account for these differences.

The second issue that requires attention is whether the concept of displaced rehearsal can also be used to account for the memory differences associated with various values of words within semantic rating dimensions. As mentioned earlier, Packman and Battig's study also demonstrated that high and low value words lead to different levels of recall within a semantic rating dimension. In their experiment, markedly higher recall was obtained for high than for low value words in all the seven semantic dimensions used, although superiority of pleasantness over the other six dimensions was consistent across the high and low subsets. The high value words on Kučera and Francis (1967) frequency counts were noted to be over four times that of the low words, and potentially these large differences between high and low words may have led to the higher retention for the former category of words (cf. Gregg, 1976). Packman and

Battig observed that recall superiority for high over low words is not limited to homogeneous lists (separate lists of high and low words), but rather it can also be found in mixed lists (consisting of high and low words). Gregg (1970; cited in Gregg, 1976), on the other hand, observed superior recall for high frequency words from homogeneous lists only. The main point is that value of words may play an important role in memory differences observed within as well as between various semantic rating dimensions.

A purpose of the proposed research was to investigate the possibility that subjects use more than one strategy to process different classes of words. They may employ the strategy of comparing high value items with previously rated items (i.e., displaced rehearsal) hence more decisions (i.e., more comparisons) may be made for these words. Alternatively, subjects may compare target items with other external items. These strategies may be used in different rating tasks or for words of different characteristics, such as high and low values of words.

Earlier studies (Deese, 1960; Gregg, 1976; Sumby, 1963) using pure lists (containing only high or low frequency words) demonstrated superior free recall of lists containing only high frequency words to recall of lists composed of only low frequency words. The high frequency advantage in pure lists has been explained in terms of organizational processes taking place during list presentation. High

frequency lists provide an opportunity for subjects "to encode useful interitem relationship during presentation of lists" (Gregg, Montgomery, & Castano, 1980, p. 241). However, Packman and Battig (1980) obtained superiority of high over low words in mixed lists across all the seven semantic dimensions used by them. The greater frequency count of high value words, led Packman and Battig to conclude that superiority of high over low words may be found in mixed lists also. The general superiority of high words, as observed in the Packman and Battig study, could be due to high value words requiring more displaced rehearsal. Accordingly, if the superior retention of high value words versus low or neutral value words, on the dimension being rated, is due to greater displaced rehearsal of the former, then requiring subjects to rate all target words in comparison with intralist words of similar value should provide an equal opportunity for displaced rehearsal for all values of words. If recall for high words is still greater than that for other words, then it would indicate that factors other than displaced rehearsal may account for the recall differences across different values of words.

An alternative hypothesis for interpreting the superior recall of high versus neutral and low value words on the dimension being rated involves factors operating at the time of retrieval. It is known that retention is influenced by retrieval factors as well as by encoding operations (Tulving

& Thomson, 1973; Tulving, 1974), a point which has been emphasized in recent modifications of the levels of processing approach (Craik & Moscovitch, 1976; Fisher & Craik, 1977; Jacoby & Craik, 1979). Although the retention differences across various classes of words rated on various semantic dimensions can be interpreted as reflecting encoding variations, the possibility remains that encoding may be equivalent for different values of words and the differences arise at retrieval, perhaps due to differences in accessibility of cues. One possibility is that at retrieval, differently encoded traces are all still present in memory and can potentially give rise to equal levels of retention, provided the optimal retrieval environment is present. As mentioned earlier, McClelland et al. (1978, Experiment 2) demonstrated that retrieval cues derived from the earlier orienting task may be used by the subjects in both free and cued recall. Thus, in the present research, it is hypothesized that when subjects compare the target words with a limited number of intralist or extralist words, these words are likely to be used subsequently as retrieval cues.

In the present research, two experiments are reported. Experiment I was designed in two parts. In part I, medium value words were rated on either the pleasantness or frequency rating dimensions under free rehearsal conditions. In part II, the target words were to be compared with a pair

of intralist words (forced rehearsal condition I) or a pair of extralist words (forced rehearsal condition II). Administration of part II depended on the results obtained in part I. In the free rehearsal condition, it was hypothesized that retention for words judged on the frequency dimension should not be lower than the words judged on pleasantness dimension if the type of words is an important determinant of recall and superiority of pleasantness is eliminated when neutral words are used (Packman & Battig, 1978). On the other hand, according to the displaced rehearsal hypothesis, the retention of words judged for pleasantness should be higher than that of words judged for frequency. If so, then, in the forced rehearsal condition I, retention differences between frequency and pleasantness dimensions should no longer be present since displaced rehearsal is encouraged for words rated on the frequency dimension also. In the forced rehearsal condition II, recall for words judged for pleasantness should decline to that of words judged for frequency, as the use of extralist words should curtail the number of intralist comparisons and hence the amount of displaced rehearsal in the pleasantness rating task. Part II was to be conducted if retention differences were observed between the pleasantness and frequency tasks, as this would indicate that factors other than word value were operating. Since the results obtained in part I indicated that there was no

retention difference between the two rating tasks, part II was not carried out.

Experiment II was an extension of Experiment I using high and low value words. In the free rehearsal condition, the words were rated either for their pleasantness or frequency. In addition, the target words were compared with either 2 intralist words (to encourage displaced rehearsal) or 2 extralist words (to discourage displaced rehearsal). An attempt was made to experimentally control for differences in decision difficulty, since, as noted earlier it has also been found to be a contributing factor to retention differences (Begg, 1978; Fisher, Craik, & Begg, 1979; Jacoby & Craik, 1979). Thus, the values of pairs of comparison words were both either high or low on the pleasantness dimension in the pleasantness rating condition and in the frequency rating condition they were both either high or low on frequency value.

EXPERIMENT I

Method

Subjects. Twenty-four undergraduate and graduate students from the University of Guelph were tested in groups of 2-12.

Materials and Design. The between-subject variable was the semantic rating dimension (pleasantness, frequency) presented in the free rehearsal condition. The number of words recalled by each group was the dependent variable.

A list of 36 single target words, plus 4 buffer items (2 primacy and 2 recency) of medium value, was selected based on the pleasantness ratings documented in the Toglia and Battig (1978) norms. The mean pleasantness rating for these words ranged from 3.40 to 4.50 on a 1-7 scale and the standard deviation was below 1.50 for each item. The frequency count ranged from 19-26 occurrences per million in the Kučera and Francis (1967) norms.

The words were presented in booklets. Each word appeared in capital letters on a separate page above a horizontally arranged five-point scale.

Procedure. Incidental learning instructions specified the item property to be judged. The subjects were informed that immediately below each word would be a scale of numbers 1 to 5. The subjects' task was to rate each word on the 1-5 scale in terms of the pleasantness of the meaning or the frequency of usage of the word. For the pleasantness rating

task, the points of the scale were labelled (from 1 to 5) very unpleasant, unpleasant, neutral, pleasant, and very pleasant. For the frequency rating task, the corresponding points were labelled very infrequent, infrequent, average, frequent, and very frequent. The subjects were asked to read each word silently to themselves before rating it. They were instructed to put a check mark on the small line above the number which they considered the appropriate rating.

The subjects proceeded through the booklets at a rate of one word every 5 seconds, paced by the sound of an automatic timer. After all the items had been rated, the booklets were collected. There was a one minute interval between the end of the rating task and the free recall test. During this time, the subjects were asked to fill in some administrative information (e.g., name, I.D. number, address) on a sheet of paper handed to them. After that, the instructions for the free recall test were read. The subjects were allowed 5 minutes for recall test. They were asked to write down as many of the words as they could remember in any order.

Results and Discussion

The mean number of words recalled in the free rehearsal condition for items judged for pleasantness and frequency was equal ($M = 8.92$). This replicates the earlier finding

(Packman & Battig, 1978) that "superiority of pleasantness memory can be eliminated when exclusively neutral rather than pleasant or unpleasant words are used" (p. 506).

EXPERIMENT II

As mentioned earlier, Experiment II was an extension of Experiment I. In Experiment I, the retention of medium value words judged for pleasantness and frequency was the same. In Experiment II, high and low value words were used to investigate the effect of value of words and displaced rehearsal on retention differences between these two rating tasks as well as within each task.

It was predicted first that, in the free rehearsal condition, retention of words judged in the frequency rating task should be lower than that for words judged in the pleasantness rating task for both high and low value words. This prediction is based on the hypothesis that in the frequency rating task extralist anchors are used in rating the target words, and therefore minimal displaced rehearsal occurs. If so, in the forced rehearsal condition I, the difference in retention between the frequency and pleasantness rating tasks should be offset when displaced rehearsal is encouraged (Walther & Horton, Note 2), since displaced rehearsal is conducive to the development of interitem associations.

It was also predicted that the retention of high value words should be higher than other values of words on the dimension being rated in the free rehearsal condition (Packman & Battig, 1978). Finally, if the retention of high value words in the free rehearsal condition, is higher than low value words due to more displaced rehearsal, then the recall differences between these two classes of words should be reduced in the forced rehearsal condition I when equal opportunity of displaced rehearsal is provided for all the words. If displaced rehearsal accounts for the differences between high and low value words, then comparison with extralist words should reduce the differences in the amount of displaced rehearsal between the two types of words Earlier data suggest that high frequency words may involve wider encoding options (cf. Gregg, 1976). Inducing subjects to compare target words with two extralist words (forced rehearsal condition II) should provide an opportunity for making an equal number of decisions for high and low value words. This would discourage displaced rehearsal which is one of an important factors in retention differences. Thus, performance would be expected to decline in both the pleasantness and frequency tasks.

With regard to extralist words, comparison words of similar values, either High-High or Low-Low were used to equate for initial decision difficulty, as this variable is associated with differences in retention (cf. Craik &

Jacoby, 1979). Similarly, in the forced rehearsal condition I, decision difficulty was controlled by equating the normative value of the comparison words selected for judgements about each target. The only difference between the comparison words used for the two rehearsal conditions was that in the forced rehearsal condition I they were intralist words whereas, in the forced rehearsal condition II, the comparison words were extralist words.

If the recall for high words remains superior to that of low words, even in forced rehearsal condition II, it would indicate that other uncontrolled factors are still operating.

Method

Subjects. Seventy-two undergraduate and graduate students from the University of Guelph, Wilfrid Laurier University, and the University of Waterloo were tested in groups of 2-12, with each group assigned to one of the six between-subjects conditions.

Materials and Design. The design was a 2 x 3 x 2 x 2 mixed factorial, with the two between-subject factors of semantic rating task (pleasantness, frequency) and rehearsal condition (free, forced I, forced II). The within-subject factors were pleasantness value of words (high, low) and frequency value of words (high, low).

The stimulus material consisted of 36 target words, plus six buffers and four extralist words, selected from the Toglia and Battig (1978) norms. Exactly half the words had a high pleasantness rating and half a low pleasantness rating in the norms. The mean for the high pleasantness value words was 5.40 or above and for the low pleasantness value words 2.50 or below. The standard deviation for the rating of each word was less than 1.50. The frequency of usage of these words was derived from the Kučera and Francis (1967) norms so that the high and low values on the pleasantness dimension were factorially combined with two levels of frequency. The high frequency value was defined as more than 40 occurrences and low frequency words 5 or fewer occurrences per million. Thus, the factorial combination of pleasantness and frequency values of words yielded four categories, viz., Low on pleasantness, Low on frequency (L-L); Low on pleasantness, High on frequency (L-H); High on pleasantness, Low on frequency (H-L); and High on pleasantness, High on frequency (H-H).

For the three rehearsal conditions, the target and buffer items remained the same. Input positions for high and low value classes of words were set up such that the means of input positions were approximately equal across the four types of words (mean of H-H and L-L was 18.56, and mean of H-L and L-H was 18.44). In the free rehearsal condition, lists comprised single words. For each rehearsal condition

two lists were prepared with no target item occupying the same input position twice.

In the forced rehearsal condition I, the words were presented along with a pair of intralist words. One word of the pair was the word immediately preceding the target whereas the other word was one of the other list words that had been presented previously. For the pleasantness task, the two intralist comparison words were equated on pleasantness value. For example, if a target word high on pleasantness value was immediately preceded by a word low on pleasantness value, then the other comparison word selected was also low pleasantness. Frequency values of these comparison words were not equated within the pleasantness rating task. Similarly, in the frequency task, the intralist comparison words presented with the target words were matched on frequency value only. For example, if a word immediately preceding the target was high in frequency value, the other comparison word was also high frequency. The values of comparison words on the rated dimension were matched to keep the initial decision difficulty uniform for all the target words. In the forced rehearsal condition I, the first two words were presented singly. Each target word was shown twice except the last target item. In the forced rehearsal condition II, words were presented along with one pair of extralist words, either two words high on pleasantness and frequency values (H-H) or two words low on pleasantness and frequency values (L-L).

The lists were presented in booklets. In the free rehearsal condition, each word was printed in capital letters and appeared on a separate page above a horizontally arranged five-point scale. In the forced rehearsal conditions I and II, the target word was presented along with a pair of intralist and extralist comparison words respectively. The target words were printed in capital letters on the left side of the page whereas the two comparison words were printed entirely in small letters, one below the other, to the right side of the target word.

Procedure. Incidental learning instructions were used. For the free rehearsal condition, the procedure was identical to that adopted in Experiment I. The subjects were asked to rate each target word on the 1-5 scale as to the pleasantness of the meaning of the word or the frequency of usage.

For the forced rehearsal condition I, the subjects were instructed to compare the capitalized (target) word on the left with the two (intralist comparison) words presented to the right of it. The subjects were asked to circle from the two comparison words the word that was closest to the target word in terms of pleasantness or frequency of usage, depending on the assigned task. The subjects were instructed not to compare the words presented on the first two pages but to read them carefully as they would be used later on. This was necessitated by the procedure used to select intralist comparison words.

In the forced rehearsal condition II, the subjects were instructed to compare the target word with two extralist comparison words. The procedure was the same as that adopted in forced rehearsal condition I.

In all three rehearsal conditions, the target words (either alone or along with the appropriate comparison words, as the case may be) were presented for 5 seconds, paced by the sound of an automatic timer.

After rating all the words in the list, the subjects were given the same one-minute filler task as described in Experiment I. Immediately after its completion, the instructions for the free recall test were read.

In the free and forced I rehearsal conditions, the subjects were asked to write down as many of the words as they could remember. In the forced rehearsal condition I, the subjects were instructed to write the words once even if they had seen them more than one time. In the forced rehearsal condition II, the subjects were asked to recall as many of the words as they could remember of those seen on the left side of each page in the booklets. All subjects were instructed to write the words in any order.

Results

The level of significance for all statistical tests was set at $p < .05$.

The mean number of words recalled as a function of the type of rating task, rehearsal condition, pleasantness

value, and frequency value of words is presented in Table 1. A split-plot analysis of variance revealed that the main effect of task was significant, $F(1, 66) = 11.43$, $MSe = 2.74$, supporting the observation that recall of items judged for pleasantness was superior to that of items judged for frequency. There was a main effect of rehearsal condition, $F(2, 66) = 13.30$. The pairwise a posteriori comparisons using the Tukey's test indicated that recall in the forced rehearsal condition II was significantly lower than recall in the free and forced I rehearsal conditions, which did not differ. The interaction of task x rehearsal condition did not attain significance, $F(2, 66) < 1$.

Figure 1 shows the mean number of words recalled as a function of frequency and pleasantness values. The interaction of pleasantness value x frequency value was significant, $F(1, 66) = 15.33$, $MSe = 1.88$. The main effects of pleasantness and frequency value of words were also significant, $F(1, 66) = 8.25$, $MSe = 1.67$, and $F(1, 66) = 86.39$, $MSe = 1.26$, respectively. A simple main effects analysis was performed on the two-way interaction. The effect of the pleasantness value of words was significant for low frequency items, $F(1, 66) = 23.26$, $MSe = 1.77$, but not for high frequency items, $F(1, 66) < 1$. The effect of frequency value of words was significant for both low and high pleasantness values, $F(1, 66) = 79.42$, $MSe = 1.56$, and $F(1, 66) = 8.18$, respectively. This indicates that

retention of high frequency value words was significantly superior to that of low frequency value words, both at low and high pleasantness values. Thus, the results show that frequency value does play an important role in influencing the retention of words. No other effects were significant in this main analysis.

On the basis of Packman and Battig's (1978) findings, it was expected that high pleasantness value items would yield higher recall than low pleasantness value items in the pleasantness rating task and high frequency value items would yield higher recall than low frequency value items in the frequency rating task in the free rehearsal condition. It was also hypothesized that, if the recall advantage for high value items over low value items was due to differential displaced rehearsal, the recall differences should be reduced in the forced rehearsal condition I. The data did not fully confirm this prediction. A priori comparisons were made, using Dunn's test, for the pleasantness and frequency tasks separately. For the pleasantness task, it was found that the recall of low and high pleasantness value items did not differ significantly in any of the three rehearsal conditions. However, in the frequency task, recall of high frequency value items was significantly greater than that of low frequency value items in the forced rehearsal condition I, but not in the other two rehearsal conditions.

Thus, the results of the present study are consistent with those obtained by earlier investigators with mixed lists (Craik & Gregg, cited in Gregg, 1976). A priori comparisons revealed no significant recall difference between high and low frequency words in the free rehearsal condition. However, as mentioned earlier a split-plot analysis performed on task, rehearsal condition, pleasantness value, and frequency value of words showed an interaction of pleasantness and frequency values. Analysis of the effect of frequency value indicated greater recall of high frequency relative to low frequency value words.

Recent studies (Begg, 1978; Fisher & Craik, 1979) have demonstrated that difficult decisions are associated with high levels of subsequent retention. Though decision difficulty was not manipulated in the present study, it was hypothesized that greater decision difficulty may be involved when the pleasantness or frequency value of target words is the same as that of the preceding word. In the free rehearsal condition, it was assumed that subjects compare the target words with the preceding words (Walther & Horton, Note 2). Therefore, higher recall may be observed for those target words whose value is the same as that of the preceding word, compared to recall of target words whose preceding word had a different scale value. Thus, low and high value target words preceded in the list by low and high value words, respectively, might be expected to be recalled

better than the low and high value words preceded by high and low value, respectively. Similar assumptions were made for the forced rehearsal condition I, since each target word was compared with the immediately preceding word paired with another word that had been presented earlier. To assess this effect, the mean proportion of words recalled in the free and forced I rehearsal conditions was calculated. The data are presented in Table 2 as a function of the value of the preceding word on the judged dimension.

A split-plot analysis of variance of these data revealed that the main effect of task was significant, $F(1, 44) = 7.31$, $MSe = .04$. The main effect of rehearsal condition was not significant, $F(1, 44) < 1$. The interaction between target word value and the preceding word value was significant, $F(1, 44) = 5.24$, $MSe = .02$. The interaction of task x rehearsal condition x target word value x preceding word value was significant, $F(1, 44) = 4.76$, $MSe = .20$. Separate analyses were performed for frequency task, pleasantness task, free, and forced I rehearsal conditions to extract information about the source of this four-way interaction. It was found that in the free rehearsal condition, only the main effects of task and target word value were significant, $F(1, 22) = 5.32$, $MSe = .03$, and $F(1, 22) = 14.14$, $MSe = .02$, respectively. This indicates that the recall for words judged for pleasantness was higher than those judged for frequency and the high value target words

were recalled better than low value words. In the forced rehearsal condition I, the main effect of target value was significant, $F(1, 22) = 7.91$, $MSe = .02$, indicating that high value target words were recalled better than low value target words.

The results in the frequency rating task showed only the main effect of target word value was significant, $F(1, 22) = 22.72$, $MSe = .02$, indicating superior recall of high value target words to that of low value words. On the other hand, in the pleasantness task, the interaction of rehearsal condition x target word value x preceding word value was significant, $F(1, 22) = 6.57$, $MSe = .02$. Because the triple interaction was significant, simple-simple main effects and simple interaction effects were calculated. The analyses indicated that, in the forced rehearsal condition I, the effect of preceding word value (high versus low) was not significant for low pleasantness targets, $F(1, 22) = 1.72$, $MSe = .02$. Thus, the data show that decision difficulty, as indexed by preceding word value, in this way is not uniformly associated with higher levels of recall. Other variables may also affect performance.

To assess the output of words other than the target words, the number of extralist comparison items and unpresented items (considered as errors) recalled by the subjects in the forced rehearsal condition II were noted. A split-plot analysis of variance was performed on these

errors. The results showed that the occurrence of extralist errors was the same in the pleasantness and frequency tasks. The main effects of task and type of error were not significant, both F's <1. A comparison of the number of occurences of unpresented item errors in different rehearsal conditions was also made. The mean number of errors observed in the pleasantness and frequency tasks in the three rehearsal conditions is presented in Table 3. A completely randomized factorial analysis conducted on unpresented item errors revealed no main effects of task and rehearsal condition, $F(1, 66) < 1$. and $F(2, 66) = 1.02$, MSe = 4.13, respectively. The interaction between task and rehearsal condition was also not significant, $F(2, 66) < 1$. Occurrence of unpresented items did not vary as a function of any of the experimental manipulations.

Clustering. Output organization during the recall test may be observed on the basis of the clustering seen from the ratings assigned to the words. The words that are assigned the same ratings may be recalled in a group if the judged attribute becomes a salient feature of the representation of the items in memory (cf. Postman & Kruesi, 1977). To assess this, in the free rehearsal condition the words were divided into five categories on the basis of the ratings they had received. For any subject, the number of items recalled in each of these categories depended on his or her distribution of ratings. Clustering was measured by using a Z score

derived from the number of categorical runs (Frankel & Cole, 1971). The degree of clustering in a group's recall was obtained by using the formula $\sqrt{n}.\bar{Z}$, where n is the number of subjects in the group and $\bar{Z}$ is the mean Z score for the group. The calculated values for this statistic for the subjects who rated the words for pleasantness and frequency dimensions were -2.23 and -1.14, respectively. Since clustering is defined as the presence of significantly "too few" runs as defined by chance, the tabled value was -1.645 at the .05 level. Hence, a clustering effect was observed for the pleasantness task. The items rated alike were not retrieved together by the subjects when the frequency task was rated.

Subjects' actual ratings as related to recall. In order to determine the degree of agreement among the subjects in ranking the target words, the Kendall's coefficient of concordance (Siegel, 1956) was determined. Particular ratings assigned by the subjects to each word were noted for the pleasantness and frequency tasks separately. The observed values of χ^2 for the pleasantness and frequency rating tasks were 246.45 and 202.63, respectively. The significant values of χ^2 for both the pleasantness and frequency rating tasks indicate that the subjects in the two tasks agreed substantially in the ratings assigned to the words. These results are in agreement with those of Toglia and Battig (1978, pp. 14-15).

The ratings used by the subjects were analysed further to evaluate any retention differences between the two rating tasks. Trend analyses were performed on the proportion of rated items that were correctly recalled, considering only the rating categories actually used by the subjects. Missing observations in the data for each individual subject were defined as those rating categories which were not used by that subject. In the pleasantness task, there was only one missing observation, which was estimated by the formula given by Kirk (1968, p. 281). In the frequency task, there were five missing observations and, therefore, an iterative process (Kirk, 1968, p. 146) was used to estimate them. Packman and Battig (1978) noted a significant linear trend component for the pleasantness task. Figure 2 shows the trend components observed in the pleasantness and frequency tasks. The data show some evidence of a linear increase in recall from low to high individual word ratings of the pleasantness task but the test statistic did not attain an acceptable level of significance, $F(1, 11) = 2.87$, $MSe = .05$, $p = .12$. Similarly no significant trend was found in the frequency task, $F(1, 11) = 2.31$, $MSe = .03$, $p = .16$, although there was some suggestion that recall of words identified as high frequency was somewhat greater than that of words identified as low frequency. Packman and Battig (1978) could not carry out trend analysis for their familiarity rating task (which they considered closest to

the frequency dimension) since most subjects did not use the low familiarity rating categories on their 7-point rating scale.

Discussion

In the present study, the major concerns were the findings relating to type of rating task, value of words, and rehearsal condition. When medium value words were used, recall in the free rehearsal condition was equal for the pleasantness and frequency tasks. These data are consistent with those obtained by Weiss, Packman, and Battig (cited in Packman & Battig, 1978). The superior recall of items judged for pleasantness versus frequency was obtained for high and low value words. These results conform with Packman and Battig's (1978) conclusions that "the type(s) of word(s) being rated may be an important determinant of any memory superiority for pleasantness over other types of semantic processing" (p. 506).

It was hypothesized that recall differences between the pleasantness and frequency tasks would be minimized in the forced rehearsal condition I because the comparison of target words with intralist words should encourage displaced rehearsal in the frequency task, as demonstrated by Walther and Horton (Note 2). However, different results have been obtained in the forced rehearsal condition I in the present study. Although the recall of words in the frequency task

increased in the forced rehearsal condition I as compared to the free rehearsal condition, performance in the pleasantness condition was still significantly higher than that observed in the frequency condition. In the Walther and Horton (Note 2) study, when subjects were forced to engage in displaced rehearsal, retention of items rated for pleasantness did not change but retention of items rated for frequency increased virtually to the level of the pleasantness condition. The basis for the differences between the two studies needs empirical clarification. one possible factor may be the nature of task performed in the two studies. In the forced rehearsal condition of the Walther and Horton study, subjects rated the target word in comparison with the immediately preceding word on either the pleasantness or frequency dimension. In the present study, subjects compared each target word with two intralist comparison words (one immediately preceding the target and one that had preceded the target earlier) to determine which one of these two words was closer to the target on pleasantness or frequency of usage. Moreover, even though the subjects saw each target item thrice in contrast to twice in Walther and Horton's study, retention differences remained. Thus, the reasons for the retention differences between the two studies remain unresolved.

As regards the present study, in terms of elaboration, it may be expected that since each target word may involve

greater amount of processing in the forced rehearsal condition I compared to the free rehearsal condition, richer and more detailed traces should be formed which aids retention (cf. Craik & Tulving, 1975). However, it seems that forced encoding may not necessarily lead to better retention. In the present study, the free rehearsal condition may be similar to subject-generated encoding condition whereas the forced rehearsal condition I may be similar to experimenter-induced processing condition. Several investigators (Bobrow & Bower, 1969; Graf, 1980; McFarland, Frey, & Rhodes, 1980; Schwartz, 1971; Slamecka & Graf, 1978) have obtained results which support the hypothesis that subject-generated encodings prove significantly more beneficial than the equivalent ones supplied by the experimenter.

The clustering data provide support for the superiority of the pleasantness task over the frequency task on the basis of organization. However, Winograd and Smith (1978) suggested that the single-trial recall experiment lacks sensitivity to organizational encoding. They compared free recall performance following either pleasantness, imagery, or number of associations rating tasks. A fourth group received intentional learning instructions and no orienting task. For half the subjects in each of these four groups, the list was presented three times followed by a single recall test, while for the other subjects, a recall test was

given after each list presentation. For the former subjects, no differences were observed on the free recall test administered after the third study trial. For the latter subjects, no differences occurred on the first test trial, but recall was superior for the intentional learning group by trial 3, compared to the three orienting task groups combined. On the basis of these results, Winograd and Smith suggested that "organizational coding may take more than a single trial to get under way, perhaps because potentiation from retrieval processes taking place during output is necessary to induce relational encoding" (p. 165). However, a closer inspection of the data suggests a somewhat different conclusion. The mean recall across three trials by subjects performing the pleasantness rating tasks was 14.7, according to data presented by the authors. Based on Figure 1 of their report, the mean recall across three trials for the standard intentional learning group is 14.77 (estimated values for Trials 1, 2, and 3 are 9.05, 15.35, and 19.90, respectively). Thus, intentional groups may perform better than groups performing some orienting tasks but not necessarily all such tasks. The pleasantness task may fit the latter category.

Another finding of interest in the present research is the significant drop in performance observed in the forced rehearsal condition II for both rating tasks. According to the displaced rehearsal hypothesis, it was expected that, as

the pleasantness task normally entails displaced rehearsal, there would be a significant drop in performance if subjects were forced to compare target items with extralist items (forced rehearsal condition II). On the other hand, performance in the frequency task would remain unaffected if items rated for frequency were normally compared with external items. Contrary to the expectations, the data indicated that, irrespective of the type of task, forcing the subjects to make extralist comparisons reduced performance, presumably by discouraging displaced rehearsal. Recall was significantly lower in the forced rehearsal condition II for the pleasantness as well as the frequency task in comparison with the free and forced I reherasal conditions. The results showed that use of extralist comparison words interferes with the retention of target words. It also implies that, in the frequency task, the target words are not compared exclusively with external anchors (as hypothesized by Postman & Kruesi, 1977), at least given the current procedure of defining extralist anchors. One possible interpretation of the lack of an interaction between task and rehearsal condition, as anticipated, may be that the frequency task entails displaced rehearsal to the same extent as the pleasantness task. Accordingly, the retention differences observed between the pleasantness and frequency tasks may be the results of other factors operating either at encoding or retrieval, or both.

According to the logic of the hypothesis of decision difficulty, it was hypothesized that, in the free rehearsal condition, the retention of low and high items preceded by an item of similar value might be higher than that of low and high items preceded by items of different values. The recall data did not confirm this hypothesis. The results show that the initial decision difficulty does not necessarily seem to be associated with high levels of retention. In the free rehearsal condition, retention of low value target items preceded by low items was not higher than the low value target items preceded by high items. In fact, recall of the former was somewhat lower than that of the latter target items. Moreover, retention of high items preceded by high ones was also not superior to retention of high items preceded by low ones. Similar results were observed in the forced rehearsal condition I. Jacoby, Craik, and Begg (1979) suggested that to understand the retention differences within the domain of semantic processing, the factors of "decision difficulty, type of test, and the preexperimental strength between some aspect of the context and the target item must all be borne in mind" (p. 595). The results of their four experiments point to the fact that other factors, in addition to the difficulty of initial processing, must be taken into account. Jacoby (1978, Experiment 2) demonstrated that problem difficulty interacts with the spacing of

presentations. The two levels of problem difficulty, easy and difficult, produced nearly equivalent levels of cued recall at 0-spacing while higher performance was observed for more difficult problems than the easy ones at the greater spacings. Though the results of the present study cannot be compared directly with earlier studies which have investigated decision difficulty, the results indicate that initial decision difficulty per se may not be associated with high levels of performance.

The next concern was the effect of item value on retention across the different rating tasks. On the basis of the results obtained by Packman and Battig (1978), it was anticipated that in the free rehearsal condition when the words were judged for pleasantness, high pleasantness value words would be recalled more than low pleasantness value words, and similarly, in the frequency task, high frequency value words would yield higher recall than low frequency value words. The results of a priori comparisons were different from these expectations. In the free rehearsal condition, the recall of high pleasantness and high frequency items was not significantly greater than that of low pleasantness or low frequency items, respectively. A similar pattern was observed for high and low pleasantness value words for the pleasantness task in the forced rehearsal conditions I and II. Thus the results are at variance with Packman and Battig's earlier findings of

superior recall for high than for low value items. These differences may be attributed to the fact that, in their study the high pleasantness words were over four times as frequent in the Kučera and Francis (1967) frequency count than the low pleasantness words. In the present study, this confounding of pleasantness and frequency was eliminated. Similarly, Packman and Battig's conclusion that the superiority of high over low value words is not limited to homogeneous lists but can be found in mixed high and low frequency lists is also not supported unequivocally as a result of this confound. Their conclusion was also in contrast to the generalizations of Gregg (1976). In the free rehearsal condition, the results of a priori comparisons in the present study were again in agreement with Gregg's finding that the superior retention of high frequency value words is not observed in mixed high and low frequency lists.

In this study, it was observed that word frequency was a potent factor regardless of whether frequency was used as the rating dimension. Analysis of the interaction of pleasantness value x frequency value of words revealed that high pleasantness value words were recalled significantly more than low pleasantness value words when factorially matched with low frequency value but the difference between low and high pleasantness value words was not observed in the high frequency condition. This may also provide a basis

for the superior recall of high pleasantness over low pleasantness words observed by Packman and Battig (1978). As mentioned above, the low pleasantness items in their experiment were also low on frequency count. The results of the present study are in agreement with their findings, if low frequency is kept constant across high and low pleasantness value. However, when low pleasantness words were averaged over frequency value the results of a priori comparisons did not conform with the Packman and Battig findings. The reason for this inconsistency may be that the low pleasantness value words with high frequency count might have reduced the difference between low and high pleasantness value words. On the other hand, the retention of high frequency value words was higher than that of low frequency value words both at low and high pleasantness values. It indicates that frequency value of words is an important factor in memory.

On the basis of the results, it may be concluded that the value of words plays an important role in the recall differences observed between pleasantness and frequency rating tasks. The results of Experiment I indicated that the superiority of the pleasantness task over frequency task is eliminated when medium value words are used. However, in Experiment II, retention differences between the two tasks were evidenced when extreme value words (i.e., high and low) were used. The performance on the frequency rating task was

lower than that of the pleasantness task across all rehearsal conditions. Encouragement of displaced rehearsal could not eliminate the differences between the two rating tasks. Furthermore, a significant drop in performance was observed for both the pleasantness and the frequency tasks when extralist comparisons were used (to discourage displaced rehearsal). This indicates that variables other than displaced rehearsal may be operating which may contribute to retention difference between these two rating tasks.

REFERENCE NOTES

1. Karmeshu, M., & Horton, K.D. Depth of processing and displaced rehearsal. Unpublished manuscript, Wilfrid Laurier University, 1981.

2. Walther, B., & Horton, K.D. Rating tasks as manipulations of depth of processing: Effects of displaced rehearsal. Paper presented at the Canadian Psychological Association annual meeting, Calgary, June, 1980.

REFERENCES

Anderson, J.R. Language, memory, and thought. Hillsdale, N.J.: Erlbaum, 1976.

Anderson, J.R., & Bower, G.H. Recognition and retrieval processes in free recall. Psychological Review, 1972, 79, 97-123.

Anderson, J.R., & Bower, G.H. A propositional theory of recognition memory. Memory & Cognition, 1974, 2, 406-412.

Anderson, J.R., & Reder, L.M. An elaborative processing explanation of depth of processing. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, 1979.

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

Battig, W.F., & Einstein, G.O. Evidence that broader processing facilitates delayed retention. Bulletin of the Psychonomic Society, 1977, 10, 28-30.

Begg, I. Similarity and contrast in memory for relations. Memory & Cognition, 1978, 6, 509-517.

Bird, C.P., & Roberts, R. An examination of orienting task relationship in a proactive interference paradigm. Memory & Cognition, 1980, 8, 468-475.

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

Bock, M. The influence of instructions on feature selection in semantic memory. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 183-191.

Cermak, L.S., & Craik, F.I.M. Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, 1979.

Coltheart, V. Recognition errors after incidental learning as a function of different levels of processing. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 437-444.

Craik, F.I.M. Levels of processing: Overview and closing comments. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, 1979.

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

Deese, J. Frequency of usage and number of words in free recall: The role of association. Psychological Reports, 1960, 7, 337-344.

Eysenck, M.W. Levels of processing: A critique. British Journal of Psychology, 1978, 68, 157-169.

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

Frankel, F., & Cole, M. Measures of category clustering in free recall. Psychological Bulletin, 1971, 76, 39-44.

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

Graf, P. Two consequences of generating: Increased inter- and intraword organization of sentences. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 316-327.

Gregg, V. Word frequency, recognition and recall. In J. Brown (Ed.), Recall and recognition. London: Wiley, 1976.

Gregg, V.H., & Montgomery, D.C., & Castano, D. Recall of common and uncommon words from pure and mixed lists. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 240-245.

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 effect 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. Recall of 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. On interpreting the effects of repetition: Solving a problem versus remembering a solution. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 649-667.

Jacoby, L.L., & Craik, F.I.M. Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, 1979.

Jacoby, L.L., Craik, F.I.M., & Begg, I. Effects of decision difficulty on recognition and recall. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 585-600.

Johnson-Laird, P.N., & Bethell-Fox, C.E. Memory for questions and amount of processing. Memory & Cognition, 1978, 6, 496-501.

Johnson-Laird, P.N., Gibbs, G., & deMowbray, J. Meaning, amount of processing, and memory for words. Memory & Cognition, 1978, 6, 372-375.

Kirk, R.E. Experimental design: Procedures for the behavioral sciences. Belmont, Cal.: Brooks/Cole, 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, 2, 671-679.

Kučera, H., & Francis, W.N. Computational analysis of present-day American English. Providence, R.I.: Brown University Press, 1967.

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

McClelland, A.G.R., Rawles, R.E., & Sinclair, F.E. The effects of search criteria and retrieval cue availability on memory for words. Memory & Cognition, 1981, 9, 164-168.

McFarland, C.E., Jr., Frey, T.J., & Rhodes, D.D. Retrieval of internally versus externally generated words in episodic memory. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 210-225.

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

Nelson, D.L. Remembering pictures and words: Appearance, significance and name. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing in human memory. Hillsdale, N.J.: Erlbaum, 1979.

Nelson, D.L., & Friedrich, M.A. Encoding and cueing sounds and senses. Journal of Experimental Psychology: Human Learning and Memory, 1980, 6, 717-731.

Nelson, D.L., & McEvoy, C.L. Encoding context and set size. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 292-314.

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

Packman, J.L., & Battig, W.F. Effects of different kinds of semantic processing on memory for words. Memory & Cognition, 1978, 6, 502-508.

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

Postman, L. The effects of language habits on the acquisition and retention of verbal associations, Journal of Experimental Psychology, 1962, 64, 7-19.

Postman, L. Methodology of human learning. In W.K. Estes (Ed.), Handbook of learning and cognitive processes Vol. 3, Hillsdale, N.J.: Erlbaum, 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, 16, 353-369.

Postman, L., Thompkins, B.A., & Gray, W.D. The interpretation of encoding effects in retention. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 681-705.

Ross, B.H. The more, the better?: Number of decisions as a determinant of memorability. Memory & Cognition, 1981, 9, 23-33.

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

Schwartz, M. Subject-generated versus experimenter-supplied mediators in paired-associate learning. Journal of Experimental Psychology, 1971, 87, 389-395.

Shaughnessy, J.J. Subjective rating scales and the control of encoding in incidental learning. Bulletin of the Psychonomic Society, 1979, 14, 205-208.

Shulman, A.I. Memory for words recently classified. Memory & Cognition, 1974, 2, 47-52.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Slamecka, N.J., & Graf, P. The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 592-604.

Sumby, W.H. Word frequency and serial position effects. Journal of Verbal Learning and Verbal Behavior, 1963, 1, 443-450.

Toglia, M.P., & Battig, W.F. Handbook of semantic word norms. Hillsdale, N.J.: Erlbaum, 1978.

Tulving, E., & Thomson, D.M. Retrieval processes in recognition memory: Effects of associative context. Journal of Experimental Psychology, 1971, 87, 116-124.

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

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

Winograd, E., & Smith, A.D. When do semantic orienting tasks hinder recall? Bulletin of the Psychonomic Society, 1978, 11, 165-167.

Footnote

[1]. Nelson (1979), Nelson and McEvoy (1979), and Nelson and Friedrich (1980) have pointed out that recall performance is directly affected by the interactive nature of the cue and processing during study. They found that in the absence of contextual cues, rhyme cues are more effective than synonyms but the reverse appears to be true when contextual cues are present during encoding and encoding time is relatively long. Thus, semantic superiority is attributed to the context provided by the cue itself. These investigators demonstrated that when retrieval is equated by controlling on an a priori basis the strength of the cue-target relationship, semantic superiority is not observed. A similar finding has also been obtained by Postman et al. (1978).

Table 1

Mean number of words recalled in Experiment 2 as a function of task, rehearsal condition, pleasantness value, and frequency value of the words (Maximum M=9.0).

		Pleasantness value				
		Low		High		
		Frequency value		Frequency value		
Task	Rehearsal condition	Low	High	Low	High	Total
Pleasantness	Free	1.75 (1.14)	3.58 (1.16)	3.08 (1.51)	3.92 (1.93)	12.33
	Forced I	1.67 (1.50)	4.42 (1.38)	3.17 (1.53)	3.33 (1.56)	12.58
	Forced II	1.17 (1.03)	2.67 (1.30)	2.08 (1.31)	2.42 (1.56)	8.33
	M	1.53	3.56	2.78	3.22	11.08
Frequency	Free	1.25 (0.97)	2.50 (1.38)	2.17 (1.40)	3.08 (1.93)	9.17
	Forced I	1.50 (1.17)	3.67 (1.61)	2.17 (1.27)	3.25 (1.66)	10.58
	Forced II	0.42 (0.67)	1.83 (1.03)	1.50 (1.17)	1.83 (1.27)	5.58
	M	1.06	2.67	1.95	2.72	8.44
Overall M		1.29	3.11	2.36	2.97	

Note: Numbers in parentheses are standard deviations.

Table 2

Proportion of words recalled in Experiment 2 as a function of rehearsal condition, target item value, task, and scale value of preceding words.

Rehearsal condition	Target item value	Task: Pleasantness — Preceding word: Pleasantness value: Low	Task: Pleasantness — Preceding word: Pleasantness value: High	Task: Frequency — Preceding word: Frequency value: Low	Task: Frequency — Preceding word: Frequency value: High
Free	Low	0.24	0.34	0.22	0.16
	High	0.41	0.36	0.36	0.27
Forced I	Low	0.33	0.34	0.19	0.16
	High	0.28	0.45	0.38	0.34

Note: In the pleasantness task, target items had low and high pleasantness value.
In the frequency task, target items had low and high frequency value.

Table 3

Mean number of errors in Experiment 2 as a function of task, rehearsal condition, and source of error.

Rehearsal condition	Task	Errors: Words not presented in the list	Errors: Extralist comparison words
Free	Pleasantness	1.00	
	Frequency	2.17	
	M	1.59	
ɔrced I	Pleasantness	0.84	
	Frequency	1.33	
	M	1.09	
Forced II	Pleasantness	1.17	0.58
	Frequency	1.33	0.58
	M	1.25	0.58
Overall M	Pleasantness	1.00	
	Frequency	1.62	

Figure 1. Mean recall in Experiment 2 of low and high frequency value of words as a function of the pleasantness value of words.

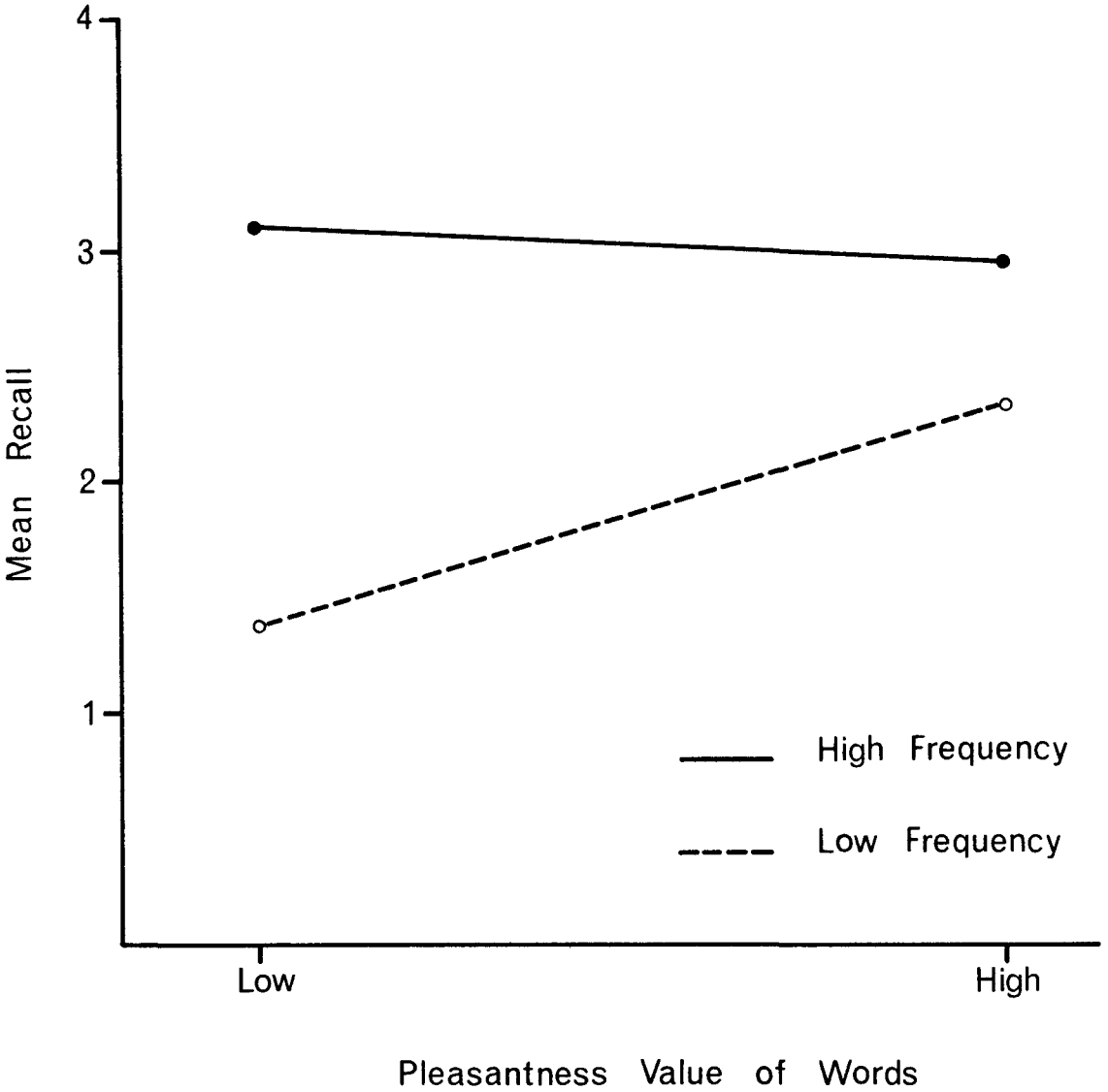
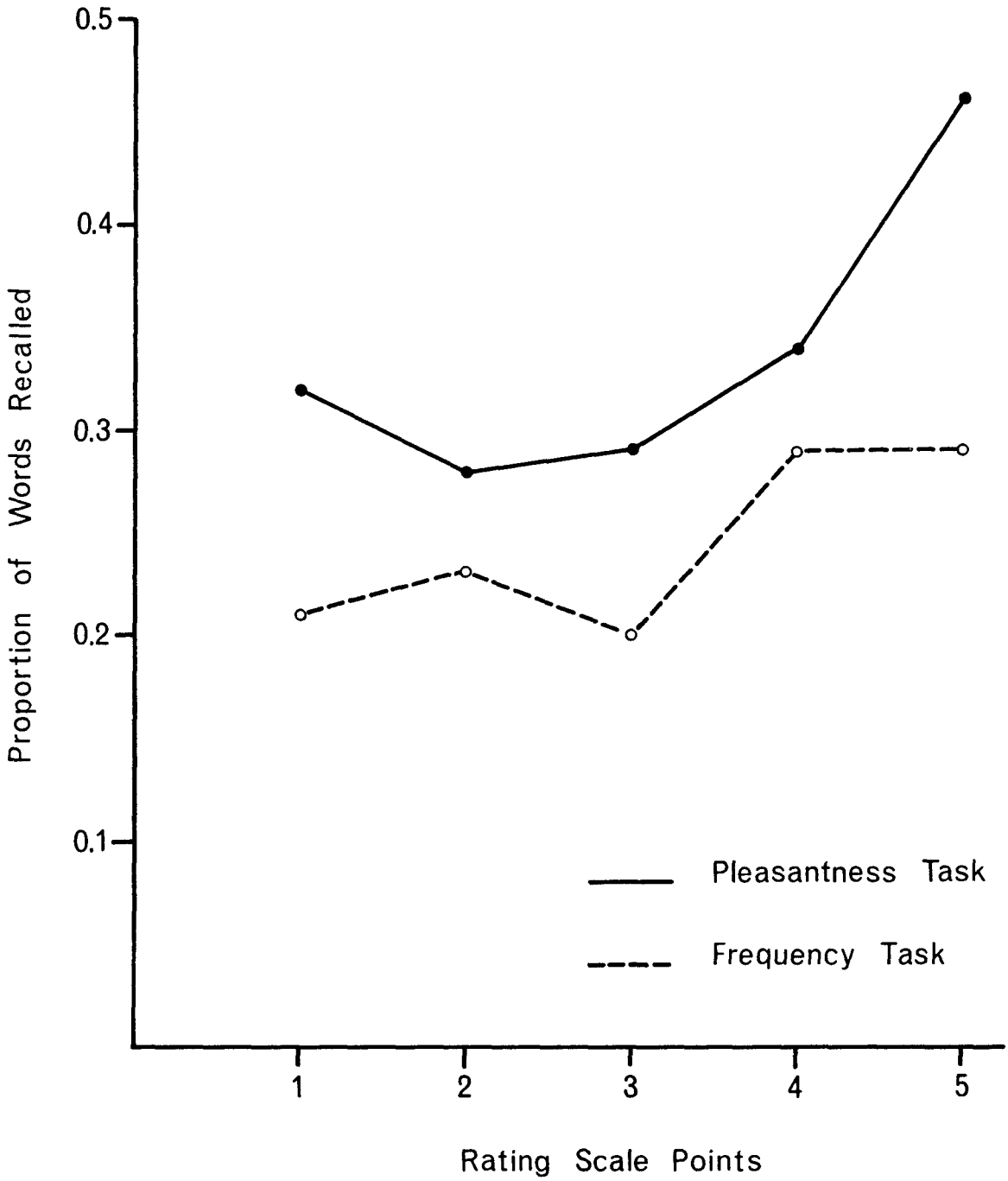

4
3
2
1
Mean Recall
High Frequency
Low Frequency
Low
High
Pleasantness Value of Words

Figure 2. Proportion of words recalled in Experiment 2 as a function of the subjects' ratings of the target words.


0.5
0.4
0.3
0.2
0.1
Proportion of Words Recalled
Pleasantness Task
Frequency Task
1
2
3
4
5
Rating Scale Points

APPENDIX A

Instructions for free rehearsal conditions

This experiment is a series of studies in which we hope to learn how people process everyday English words.

In front of you is a booklet containing a series of pages. On each page is a single word. Immediately below each word is a scale of the numbers 1 to 5 (show example page). Your task in this experiment is to rate on the 1-5 scale each word according to how pleasant the meaning of this word (how frequent the usage of this word) seems to you.

On the scale,

a 1 indicates very unpleasant (very infrequent);

a 2 indicates unpleasant (infrequent);

a 3 indicates neutral (average);

a 4 indicates pleasant (frequent);

a 5 indicates very pleasant (very frequent).

(show it)

Please put a check mark (√) on the small line above the point which you consider best for a particular word.

Read each word silently to yourself before rating it. You will see each word for a very brief duration of only 5 seconds during which you must perform the rating of the word.

At the end of the 5 second there will be a sound of click from this automatic timer (demonstrate). When you

hear the click, turn immediately to the next page and perform the rating on the next word. Please do not turn the page before you hear the click.

A couple of other points are important in this experiment.

1. Give your first impression when you read the word.
2. Be careful to give only one answer to each word.
3. Be careful that you do not skip any pages in the booklet.

Are there any questions about the task? (wait)

Is everybody ready?

Please begin now.

Instructions for one minute filler task

Please fill in the information asked for on this sheet. The information is required so that we may keep in touch with you and inform you about the results of the experiment.

While you are asked for your name, among other things, please be assured that this is only used for the purpose of contacting you. It will not be used in any other way.

Recall Instructions

For the final part of this experiment, I would like you to write down as many of the words as possible of those which you saw earlier. You may write down the words in whatever order you think of them.

Instructions for forced rehearsal conditions I and II

This experiment is a series of studies in which we hope to learn how people process everyday English words.

In front of you is a booklet containing a series of pages. On each page is a word printed capital letters on the left side of the page. To the right of the capitalized word is a pair of words, both of which are printed entirely in small letters.

Your task will be to compare the capitalized word on the left (show it) with the two words to the right of it. The words differ in their capacity to elicit a feeling of pleasantness. Some words induce a feeling of pleasantness in us, whereas other words evoke an unpleasant feeling. (Words differ in their usage. Some words are used frequently, whereas other words are used infrequently.) From the two words printed on the right, circle the word that is closest to the left hand word in terms of pleasantness (frequency of usage) of this word (show an example).

You will see words printed on each page for a brief duration of 5 seconds only. After every 5 seconds you will hear a sound from this automatic timer (demonstrate). On hearing this sound, immediately turn to the next page and compare the next word with the words presented along with it.

Remember a few points:

1. Read the words carefully as you have limited time and

answer as accurately as possible.

2. Be sure that you circle one word and only one word from the two words printed on the right side of the page. It is very important that you select one of these words from each pair.
3. Be careful while turning the pages that do not skip any pages.

(In the forced rehearsal condition I:) Remember on the first two pages of the booklet you will find single words only. You are not required to compare these words, but read them carefully as they will be used later on. I shall tell you when to start comparing the left hand capitalized word with the other two words.

Are the instructions clear to everyone? (wait)

Should we start?

Start.

Instructions for recall test for forced rehearsal condition I

This is the final part of this experiment. Now I would like you to write down as many words as you can remember. Write them in any order. Write the words only once even if you have seen them more than one time.

Instructions for recall test for forced rehearsal condition II

This is the final part of this experiment. Now I would like you to write down as many words as you can remember of

those that you saw on the left side of the page. Write them in any order.

APPENDIX B

Analysis of Variance Summary Table for Mean Recall Scores

Source	SS	df	MS	F
Task	31.34	1	31.34	11.43**
Rehearsal condition	72.92	2	36.46	13.30**
Task x Rehearsal condition	1.05	2	0.52	.19
Error	180.94	66	2.74	
Pleasantness value	13.78	1	13.78	8.25**
Pleasantness value x Task	.03	1	.03	.02
Pleasantness value x Rehearsal condition	3.52	2	1.76	1.05
Pleasantness value x Task x Rehearsal condition	.65	2	.32	.19
Error	110.27	66	1.67	
Frequency value	108.78	1	108.78	86.39**
Frequency value x Task	.00	1	.00	.00
Frequency value x Rehearsal condition	5.02	2	2.51	1.99
Frequency value x Task x Rehearsal condition	.34	2	.17	.14
Error	83.10	66	1.26	
Pleasantness value x Frequency value	28.75	1	28.75	15.33**
Pleasantness value x Frequency value x Task	1.84	1	1.84	.98
Pleasantness value x Frequency value x Rehearsal condition	3.17	2	1.59	.85
Pleasantness value x Frequency value x Task x Rehearsal condition	1.72	2	.86	.46
Error	123.77	66	1.88	

** $p < .01$.

APPENDIX C

Analysis of Variance Summary Table for Recall as a Function of Preceding Word Value

Source	SS	df	MS	F
Task	.26	1	.26	7.31**
Rehearsal condition	.03	1	.03	.85
Task x Rehearsal condition	.01	1	.01	.15
Error	1.59	44	.04	
Target word value	.46	1	.46	21.32**
Target word value x Task	.07	1	.07	3.07
Target word value x Rehearsal condition	.01	1	.01	.24
Target word value x Task x Rehearsal condition	.02	1	.02	.91
Error	.94	44	.02	
Preceding word value	.01	1	.01	.41
Preceding word value x Task	.09	1	.09	5.24*
Preceding word value x Rehearsal condition	.06	1	.06	3.46
Preceding word value x Task x Rehearsal condition	.00	1	.00	.05
Error	.79	44	.02	
Target word value x Preceding word value	.01	1	.01	.26
Target word value x Preceding word value x Task	.01	1	.01	.35
Target word value x Preceding word value x Rehearsal condition	.05	1	.05	2.23
Target word value x Preceding word value x Task x Rehearsal condition	.10	1	.10	4.76*
Error	.92	44	.02	

* $p < .05$.

** $p < .01$.

APPENDIX D

Analysis of Variance Summary Table for Errors

Source	SS	df	MS	F
Task	6.72	1	6.72	1.63
Rehearsal condition	3.11	2	1.55	.38
Task x Rehearsal condition	3.11	2	1.55	.38
Error	272.33	66	4.13	
Total	285.28	71	4.02	