Depth of Processing, Retrieval Cues, and Uniqueness of Encoding as Factors in Recall

MORRIS MOSCOVITCH AND FERGUS I. M. CRAIK

University of Toronto

Three experiments were conducted to examine the relations between retrieval and encoding in a levels of processing framework. Retrieval conditions were manipulated in the following ways: (a) Retrieval cues were either present or absent, (b) Cues were either unique for each target word or were shared among targets, and (c) The cue and target either did or did not form a congruent, meaningful unit. It was found that the beneficial effects of cueing, uniqueness and congruence were greater at deeper levels of encoding. It is proposed that deeper encoding establishes a higher ceiling on potential memory performance. The extent to which the potential is realized depends on the degree to which retrieval conditions recreate the encoding context that uniquely specifies the target item.

Craik and Lockhart (1972) proposed that memory traces are by-products of the perceptual and cognitive operations performed on stimuli. They argued that the nature and durability of the memory trace is determined by the level at which a stimulus is processed: Deep, semantic processing is associated with higher retention than is processing at a shallower, perceptual level. As Craik and Lockhart noted, this view of memory focuses primarily, if not exclusively, on the encoding processes operating at the time of input. Although there are many reasons to believe that encoding processes determine what is stored in memory, the only way of assessing the contents of that store at present is by testing the subject's performance in recall and recognition. Since it is known that memory, as measured by performance at testing, is influenced by retrieval factors as well as by encoding operations (Tulving & Thomson, 1973), the levels of processing framework, as presently stated, can provide only a partial explanation of memory processes. In short, the Craik and Lockhart formulation lacks any clear hypotheses 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.

Craik (1973) described a study in which subjects were shown a series of words. As each word was presented, the subject's task was to answer one of the following questions: (a) Is the word written in capital letters? (b) Does the word rhyme with ____? (c) Is it a member of category ____? A superficial, visual analysis of the stimulus is sufficient to answer question (a), while a deeper more semantic analysis of the word is necessary to answer (c). In a later unexpected memory test, subjects performed in accordance with the levels of processing notion. That is, deeper levels of encoding were associated with superior memory performance. One result was
puzzling, however: Fewer words receiving negative responses in the initial decision task were recalled and recognized than words receiving positive responses. Since both negative and positive words presumably must be processed to the same level before a decision is reached, the difference in retention of these two kinds of words poses a problem for the levels-of-processing view.

An answer to the problem was provided by Craik and Tulving (1975). They suggested that positive words are compatible with the encoding question and therefore the word and question can combine to form a coherent, integrated memory trace. In turn, if the encoding question is remembered during the memory test, the question can then plausibly act as a cue to aid in retrieval of the word. For example, if the original question was “Is the word an animal name?” then BEAR will be retrieved more readily than SPEECH in the later retention test, since remembering there was a question about animals is a cue for BEAR but not SPEECH. This line of argument can be extended to the more general suggestion that some encoding questions function as more potent retrieval cues than others. The differential retention of words after case, rhyme, and category questions might thus reflect differential effectiveness of the question to act as cues, rather than differences in durability of the trace.

Two notions, suggested by this approach, were examined in the present studies. One possibility is that deeper encoding questions are themselves more memorable. That is, subjects may remember the question “Is the word an animal name?” but not “Does the word rhyme with fat?” If differential accessibility to retrieval cues, in this sense, is a factor in the levels of processing paradigm, the levels effect should be attenuated or eliminated when the original encoding questions are provided at retrieval. That is, the levels effect should be attenuated under cued recall conditions; this proposition was examined in Experiment I. A second possibility is that deeper questions lead to a more distinctive or unique encoding in memory. In that case, shallow questions, such as those concerning typescript and rhyme, will be relatively poor retrieval cues, since they do not specify the target encoding uniquely. This factor was explored in Experiments II and III by manipulating the number of words encoded with each question. If uniqueness of the question-word combination is a factor in determining the question’s cue effectiveness, then using the same question for several words should reduce its potency as a cue, in much the same way as the cueing power of a category name declines as the number of category instances to be remembered increases (Tulving & Pearlstone, 1966; Watkins & Watkins, 1975).

EXPERIMENT I

As described above, the point of this study was to eliminate any differences in memorability of the encoding questions by reproviding them at retrieval. If strong differences in retention between words encoded at different levels still remain under these conditions, this aspect of a retrieval explanation of depth of processing can be ruled out.

Method

Thirty-six undergraduates of both sexes were each paid $2.00 to act as subjects in the experiment. Typically, they were tested in groups of two or three. Subjects were informed that the experiment investigated the speed and accuracy with which they could answer questions about words. They were not told about the subsequent memory test.

On each trial a word was exposed on a TV screen for 200 msec. Before the word was exposed the subject was asked a question about it. The questions, which were printed on cards that the subject read, took one of these three forms:

(a) Rhyme: Does it rhyme with ___?
(b) Category: Is it in the category ___?
The purpose of these questions was to induce the subject to process the words to different levels of analysis. The rhyme question presumably required a relatively superficial phonemic analysis, while the category and sentence questions required deeper, semantic analyses. There is little a priori justification for classifying sentence questions as deeper than category questions; empirically, however, Craik (1973) found higher recognition levels associated with sentence questions. The questions were printed on a deck of cards which was placed before the subject. On each trial, he turned over the top card, read the question, and responded "yes" or "no" as quickly and as accurately as he could after word presentation by pressing one of two response keys. Each trial took about 10 sec to complete.

After 60 such trials and a brief rest, subjects were asked, unexpectedly, to write down all the words they could remember seeing on the screen. Half of the subjects were given the encoding questions as retrieval cues during recall (cued condition) while the remaining subjects were asked to recall the words in any order, with no retrieval cues as aids.

The stimuli were common four, five, and six letter words. On each set of 60 trials, one-third of the words were associated with rhyme questions, one-third with category questions, and one-third with sentence questions. Within each type, half of the questions required positive answers and half negative answers. Negative and positive questions, as well as the type of question, were presented in random order throughout the 60 trials. To avoid artifacts due to particular word–question combinations, each word was associated with each type of encoding question an equal number of times, so that on one set of 60 trials a word would be associated with a rhyme question while for another subject it would be associated with a "sentence" question, and so on.

Results

The results are shown in Fig. 1. For free recall, the data are similar to those reported by Craik and Tulving (1975). That is, recall levels were higher for semantically encoded words than for phonemically encoded ones; recall was also higher for words associated with positive responses than for those associated with negative responses (the apparently anomalous recall level for the sentence-positive group will be examined in the General Discussion). For the cued recall group the same general pattern was obtained. Critically, the levels effect was not attenuated under cued recall conditions; on the contrary, the effect was amplified. Cueing raised all recall levels, but helped positive responses more than negative responses, and semantic encodings more than phonemic encodings. These impressions were confirmed by the results of a three-factor analysis of variance. Significant main effects of encoding question \(F(2, 68) = 81.1, p < .001\), cueing \(F(1, 34) = 35.5, p < .001\), and response type \(F(1, 34) = 237, p < .001\) were found. In addition, the inter-
actions between encoding question and response type \( F(2, 68) = 16.1, p < .001 \), between encoding question and cueing \( F(2, 68) = 16.7, p < .001 \), and between response type and cueing \( F(1, 34) = 108, p < .001 \) were statistically reliable. These interactions are interpreted as showing a greater superiority of positive responses at deeper encoding levels, a greater benefit of cueing at deeper levels, and a greater benefit of cueing for positive responses, respectively. Finally, the three-way interaction between question type, response type, and cueing was statistically reliable, \( F(2, 68) = 5.25, p < .01 \). This interaction shows that the effect of encoding questions on the difference between positive and negative responses was greater in the cued than in the noncued condition.

**Discussion**

The results show that making all the encoding questions equally accessible by re-presenting them as retrieval cues at the time of recall magnifies, rather than diminishes the differences in retention associated with different levels of processing. Clearly, the findings do not support the hypothesis that differences in the accessibility of retrieval cues underlie the effects of levels of processing on retention.

However, the results are in good agreement with the viewpoint that retention is to be understood as an interaction of encoding and retrieval factors (Tulving, 1974). At a general level of description, the present pattern of results could be interpreted as showing that recall performance depends (a) on the quality of the trace, (b) on the presence of retrieval cues, and (c) on the degree of integration of the item with its encoding context. The quality of the trace, in turn, is a function of depth of processing; the compatibility of item and context will depend on their congruity (Schulman, 1974) from past experience.

Put slightly differently, the qualitative nature of the trace (as determined by the level of processing) may set an upper limit on recall and recognition; how near the subject's performance approaches the upper limit for a given level of processing will depend on the effectiveness of the retrieval environment. If retrieval is conceptualized as a process of reconstructing the original encoded percept (involving item and context) then the degree of congruity or integration between each word and its encoding question will determine how easily the word can be reconstructed when the question is accessed. Thus, even in free recall, words associated with positive responses will show an advantage, since they are more congruous with such encoding questions as are spontaneously remembered. In cued recall, the advantage of positive over negative words will be further amplified, since re-presenting the encoding question as a retrieval cue will facilitate regeneration of the encoding context and lead to higher recall levels only if the encoding question and its associated word formed a well integrated unit. This idea is supported by the finding that cueing primarily enhances the recall of positive rather than negative words.

The proposed relation between levels of processing and retrieval operations has much in common with Tulving and Pearlstone's (1966) proposal that recall is a function of the availability and accessibility of to-be-remembered items. In terms of the present study, the availability of the item is determined by the level to which it is processed while its accessibility is determined by the efficacy of the retrieval environment. If this is a valid comparison, memory performance in a levels of processing paradigm should be influenced by the same retrieval factors that Tulving and Pearlstone found affected recall of categorized lists of words. One major finding in the Tulving and Pearlstone study was that the probability of an item's retrieval declined as the number of items presented per category increased. In the present paradigm, the cueing effectiveness of the encoding question may decline as more items are associated with each
question. This notion was tested in Experiment II.

**EXPERIMENT II**

The point of this study was to compare the situation in which each word in a levels of processing paradigm is associated with its own, unique encoding question with the situation in which groups of 10 words shared the same question. On the basis of Tulving and Pearlstone's (1966) results, it was expected that the sharing manipulation would reduce the question's cue effectiveness and that recall would be decreased. However, in light of the argument that the level of encoding determines an upper limit for recall, an increase in recall from phonemic to semantic encoding was also expected under shared-cue conditions. Since 18 subjects had been tested under conditions of cued recall, with unique encoding questions, in Experiment I, their results formed the basis of comparison in the present study. Eighteen further subjects were tested under similar conditions to those in the first experiment, except that now the 60 words were divided into six sets of 10 words, with each set of 10 words (five positive responses and five negative responses) sharing the same encoding question. Again, memory performance was assessed by cued recall.

**Method**

Two groups of subjects were compared in the present study. The first group of 18 subjects (from Experiment I) was tested under unique cue conditions, while the second group of 18 was tested under shared cue conditions. Eighteen undergraduate subjects were each paid $2.00 to participate in the shared-cue condition. The procedure for the shared-cue condition was similar to that of the unique cue condition except that now the 60 words were divided into six sets of 10 words, with each set of 10 words (five positive responses and five negative responses) sharing the same encoding question. In order to reduce artifactual effects associated with specific questions, three formats of the shared-cue materials were constructed; each format had a different set of words and questions. Six subjects were assigned to each format.

The testing procedure was identical to that used in Experiment I. Subjects were shown 60 words, one at a time, on a TV screen and were required to answer a question about each word. The encoding questions were of three types: rhyme, category, and sentence. Subjects understood that their task was to answer the questions rapidly by pressing one of two response keys; they were not informed of the later memory test. Following a brief rest after the presentation of all the words, the subjects were unexpectedly asked to recall as many of the words as they could. They were given the encoding questions as retrieval cues.

**Results**

Figure 2 shows that recall levels were lower for the shared-cue group. However, the detrimental effects of cue-sharing were not constant for all experimental conditions; the effects were greater for positive than for negative responses, and within positive responses...
the detrimental effects of sharing were confined to category and sentence questions. As in Experiment I, the recall level for the "sentence-positive" group is apparently low in the shared condition; this finding will be examined in the General Discussion. As in previous experiments, the main effects of type of encoding question \( F(2, 68) = 102, p < .001 \) and of response type \( F(1, 34) = 372, p < .001 \) were statistically reliable. More important, recall levels for the shared-cue group were reliably lower, \( F(1, 34) = 25.3, p < .001 \). The Cue-sharing \( \times \) Question type interaction was significant \( F(1, 34) = 25.1, p < .001 \), showing that sharing was more detrimental to semantic levels of encoding. Also, the Cue-sharing \( \times \) Response type interaction was reliable, \( F(1, 34) = 10.7, p < .01 \), showing that sharing was more detrimental to positive responses. Finally, the three-way interaction of Sharing \( \times \) Question type \( \times \) Response type was significant, \( F(2, 68) = 8.88, p < .001 \). This last interaction demonstrates that the interaction between Question type (level of processing) and Response type (positive or negative) was greater under unique-cue conditions.

**Discussion**

The present study showed that fewer words are recalled in the shared than in the unique, condition\(^1\). This result is consistent with Tulving and Pearlstone's (1966) finding that the probability of recalling any one word associated with a retrieval cue diminishes as the number of words associated with that cue increases. It is also consistent with cue-overload theory (Watkins & Watkins, 1975).

Although they were not predicted, the most interesting results were the interactions between cue-sharing and levels of processing, and between cue-sharing and response type. The provision of unique cues was of greater benefit to semantically-encoded words and to words associated with positive responses. This pattern of results is somewhat similar to that obtained in Experiment I, where the provision of cues was again of greater benefit to deeper encodings and to positive responses. Both sets of results are in line with the notion that whereas the level of encoding determines an event's potential memorability, the quality of the retrieval information determines the extent to which that potential is realized.

This point is discussed more fully in the final discussion; one aspect is amplified here, however. Assume that the uniqueness or distinctiveness of an encoded trace is one major determinant of its memorability; distinctive traces can be retrieved relatively easily, analogously, perhaps, to the relatively easy perception of distinctive events in the visual field. In this sense, words encoded primarily in terms of their physical or phonemic features will not be particularly unique, since many other encoded word episodes share the same features. For words encoded physically or phonemically, there are relatively small alphabets of features, used in various combinations, to encode the events. For semantically encoded words, on the other hand, the forms of encoding are virtually limitless and, speculatively, these semantic encodings are less overlapping in their content than are physical and phonemic encodings. It follows from this analysis that items encoded to shallow levels would suffer relatively small decrements in memorability from the cue-sharing manipulation; such encodings are already similar to many others in memory. On the other hand, the beneficial uniqueness of semantically-encoded words should be affected to a larger extent, and this is the result obtained (see Restle, 1974, for a similar viewpoint).

A further point suggested by the foregoing analysis is the distinction between uniqueness of form and uniqueness of substance. To illustrate, rhyme-encoded words were very little affected by the sharing manipulation in Experiment II, although each word had a
different rhyme question in the unique condition. This finding suggests that formal or surface uniqueness is less important than is substantive uniqueness; if uniqueness at a substantive, semantic level is lacking, surface or formal uniqueness will not promote memory. In the previous experiment, no attempt was made to separate formal from substantive uniqueness; the encoding and retrieval questions were either unique in form and substance or they shared both formal and substantive properties. It is possible to separate the two, however. For example, the following semantic encoding questions are unique with respect to their formal or surface properties but similar with regard to substance: Is it a garment? Is it a form of apparel? Is it clothing? Can you wear it? Examples of phonemic encoding questions of this type are: Does it rhyme with lore? with door? with roar? If uniqueness is effective only at the level of meaning, then words associated with questions that differ in form, but mean the same thing, should be remembered no better than words associated with identical questions. Experiment III was designed to test this prediction.

**EXPERIMENT III**

As in the previous experiments, the subjects were shown a series of words about which they had to answer either a phonemic or a semantic encoding question. These questions were re-presented as cues at retrieval. The subjects were divided into three groups of 16, each of which participated in only one of the following conditions: (a) *Unique condition*: Each target word was associated with questions that were unique in form and substance. (b) *Similar condition*: Each target word was associated with questions that were unique in form, but, for each set of six target words, the questions were similar in substance, that is, they all meant the same thing or they all rhymed with each other. (c) *Shared condition*: Sets of target words were composed such that all the words in that set shared the same question. If uniqueness operates primarily at the level of meaning then memory performance should not differ between the shared and similar condition. Moreover, the effects of uniqueness should be observed primarily for those words associated with semantic questions. Words processed phonemically should benefit little, if at all, from their association with unique questions.

**Method**

Forty-eight male and female undergraduates were paid to participate in the experiment. The general procedure was the same as in the other experiments. Groups of two to eight subjects were tested together. This time, however, they were informed that they were participating in a memory experiment, since it has been shown that performance varies only slightly between intentional and incidental learning (Craik & Tulving, 1975). The subjects were shown 60 words, one at a time, on a TV screen for 200 msec each, and asked to answer a question about each word. The questions were typed on cards. Half of the questions were rhyme questions (e.g., Does it rhyme with pool?) and half were questions about the word’s semantic category (e.g., Is it an animal?). Half of the answers for each type of question were positive and half were negative. Each trial took about 10 sec. Following the presentation of the words, the subjects were given a brief rest and then they were asked to recall as many words as possible. The encoding questions were given as retrieval cues and the subjects were asked to write the word that had been presented with each cue.

Sixteen subjects were assigned to each of the three conditions of the experiment. In the *unique condition*, each word was paired with a question that was unique in form and substance. For the *similar* and *shared* condition, the sixty words were divided into 12 sets with five words in each set. Six of the sets were associated with phonemic questions, six with semantic questions. Within each group of six, three sets were positive and three were negative.
In the similar condition, each word in a given set was associated with a question that was unique in form, but similar in substance. Examples of rhyme questions are: Does it rhyme with lore? Does it rhyme with door? Does it rhyme with roar? Examples of category questions are: type of apparel, something to wear, type of garment. In the shared condition, all the words in a set were associated with a single question.

Two separate pools of 60 words were used, one for the unique condition and one for the similar and shared conditions; the nature of the experimental manipulation precluded the use of one common pool. In the unique condition, four formats were constructed such that each word was rotated through the combinations of semantic and phonemic encoding with positive and negative responses. In the similar and shared conditions, two different orders of question presentation were employed. For all conditions, the order in which the cues were presented at retrieval differed from their order of presentation at encoding.

Results

Two criteria were applied in scoring the subjects’ cued-recall performance: a strict criterion in which responses were counted correct only when they were given to questions with which they were paired at encoding, and a liberal criterion, in which responses were counted correct regardless of the eliciting cue. The main conditions affected differently by the scoring criterion were negative responses and all words in the similar condition. More negative words were scored correct under the liberal criterion because spontaneously retrieved words were often paired with inappropriate cues. Also, many words in the similar condition were often recalled after presentation of a cue that was correct in substance but not in form. Results of the two scoring methods are shown in Table 1. Since the pattern of results is broadly similar under the two scoring methods, all further analyses were based on the liberal scoring criterion. Since the results of an analysis of variance conformed to the general pattern obtained in Experiment II, they will not be reported in detail. It should merely be noted that all main effects were significant (p < .01), and that the recall superiority of positive over negative words was highest in the unique condition (p < .01) and when the words were semantically encoded (p < .001).

The main result of interest is that recall levels were no higher under similar than under shared cue conditions, even when the results are scored liberally. In fact, recall performance for rhymes was significantly worse in the similar condition (Newman–Keuls, p < .05), probably because the presence of many cues that rhymed with each other, as well as with

<table>
<thead>
<tr>
<th>Condition</th>
<th>Response</th>
<th>Shared</th>
<th>Similar</th>
<th>Unique</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rhyme</td>
<td>Category</td>
<td>Rhyme</td>
</tr>
<tr>
<td>Strict scoring criterion</td>
<td>Positive</td>
<td>.29</td>
<td>.64</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>.03</td>
<td>.08</td>
<td>.01</td>
</tr>
<tr>
<td>Liberal scoring criterion</td>
<td>Positive</td>
<td>.32</td>
<td>.66</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>.22</td>
<td>.29</td>
<td>.10</td>
</tr>
</tbody>
</table>

TABLE 1

PROPORTIONS OF WORDS RECALLED IN EXPERIMENT III
the target, made it difficult for subjects to differentiate targets from cues. There is thus no evidence that formal uniqueness by itself is beneficial to memory performance. In the semantic conditions, recall is superior only when the words and questions form encodings which are unique in substance. In the rhyme conditions, even substantive uniqueness confers little benefit, suggesting that the distinctiveness of present events, both from previous events and from accumulated knowledge, is important for retention. Distinctiveness of current episodes from each other may not be sufficient if they are similar in substance to many previous events.

Otherwise the present results bear out previous findings. Subjects recalled more semantically-encoded words than rhyme-encoded words; also, more positive responses than negative responses. As in Experiment II, the provision of unique cues increased recall levels for positive responses but not for negative responses; within positive responses, the beneficial effect of uniqueness was largely restricted to category encodings.

GENERAL DISCUSSION

The experiments in this study were designed to examine retrieval aspects of Craik and Lockhart’s levels-of-processing approach to memory. As has been noted, the findings of this study do not challenge the basic premise of that approach, which is that retention of words depends on the depth to which they are processed. At the same time, however, the study has shown that other factors may be as important as levels of processing in determining memory performance. Thus, for a given level of processing, performance depended on whether retrieval cues were present or absent, on whether the encoding question received a positive or negative answer, and on whether retrieval and encoding processes were unique or shared for a given word.

Experiment I tested the notion that the levels-of-encoding effects reported by Craik (1973) and by Craik and Tulving (1975) were associated with differential accessibility of encoding contexts at retrieval. Thus, if semantic contexts were generally more accessible, this would lead to enhanced retrieval of semantically encoded events. This analysis leads to the notion that if all encoding contexts were made equally accessible at retrieval, by re-providing them as retrieval cues, the typical levels effect should be reduced. Contrary to this prediction, Experiment I showed that cued recall amplified the differences in retrievability associated with different encoding contexts. Thus the initial results reported by Craik and Tulving (1975) cannot be attributed in any simple fashion to differential accessibility of encoding contexts.

However, the results obtained in Experiment I can be taken as illustrating the position that overall retention must be viewed as a joint function of both trace information and retrieval information (Tulving, 1974). If deeper levels of encoding are associated with the formation of richer, more elaborate traces, then such encodings will plausibly be retrieved more often, even under free recall conditions. With the provision of more adequate retrieval information in cued recall, the greater potential of deeper-level encodings can be more fully realized. This last statement leads to the notion that encoding operations establish a ceiling on potential memory performance, and retrieval cues determine the extent to which that potential is utilized.

What factors determine the effectiveness of a retrieval cue? One major factor has been enunciated as the encoding specificity principle by Tulving and Thomson, (1971). In essence, this principle states that a retrieval cue is effective to the extent that it formed part of the item’s initial encoding; the more completely the initial encoding is reinstated by the information provided at retrieval, the greater the probability that remaining aspects of the original encoding will be remembered. The present results provide further support for the encoding specificity principle. However, they also point to the necessity of incorpora-
ting other factors in any final account of retrieval. One such factor is the uniqueness of the link between the retrieval cue and the encoded event; a second is the degree to which the event and its encoding context (which later serves as the retrieval cue) form an integrated encoded unit.

One way of describing the difference between free recall and cued recall in Experiment I is to say that in cued recall the retrieval information was unique and specific to each encoded word, whereas in free recall the retrieval information is much more general and is, in effect, shared by all encoded events. Experiments II and III directly explored the effects of the uniqueness of the link between each event and its encoding context. Experiment II showed that unique cues were associated with enhanced recall, especially of semantically encoded positive responses. Experiment III showed that formal or surface uniqueness did not benefit recall; events must be distinctive semantically before recall is enhanced. Both experiments illustrated the importance for recall of a unique link between the encoded event and the retrieval cue, and that this factor of uniqueness is largely restricted to cases where the event and the cue form an integrated unit (that is, for positive responses) and to cases where the relationship between the event and the cue is semantic in nature. The relative ineffectiveness of rhyme cues was also reported by Jacoby (1974) and by Nelson, Wheeler, Borden, and Brooks (1974). They postulated that semantically encoded word pairs form highly specific, integrated codes, whereas rhyme encoded words yield codes which are less specific and less well integrated.

The necessity for integration between an event and its cue was also shown in a study by Craik and Tulving (1975, Experiment 7). In this experiment, the elaborateness of a word's encoding was manipulated by presenting each word in the context of a sentence frame which was more or less complex; also, the word either did or did not fit meaningfully into the sentence frame. It was found that representation of the sentence frame as a cue in a later retention test was ineffective for cases in which the word did not fit the sentence. The cues did enhance recall of positive response words, however, and this effect was greatest for the most complex sentence frame. Thus, arguably, retention was best (a) when retrieval information was optimized by a re-presentation of the encoding context as a cue, (b) when the context and the encoded event were well integrated, and (c) when the event was made distinctive by being qualified by an elaborate sentence frame.

A similar explanatory framework may help account for the usefulness of imagery as a mnemonic aid in verbal recall. As Lesgold and Goldman (1973) have shown, the effectiveness of a visual mnemonic depends on the degree to which it can evoke the encoding context that uniquely specifies the target item. In the present series, one unexpected feature of the results may also be described speculatively in these terms. In Experiments I and II, sentence-encoded words are less well recalled than category-encoded words in condition "free recall-positive responses" (Exp. I) and condition "shared cue-positive responses" (Exp. II). However, under "cued-recall-positive responses" (Exp. I) and "unique cue-positive responses" (Exp. II) more sentence-encoded words than category-encoded words were recalled (see Figs. 1 and 2). It seems at least possible that the sentence encodings conferred the greatest uniqueness and thus potential recall level, but this specificity of encoding could only be utilized to aid recall when the specific context was reinstated under cued recall conditions.

In conclusion two points will be re-examined: Why are the effects of cuing and cue-sharing greater with semantic than with phonemic encodings? What do the present results contribute to our understanding of the concept of depth of encoding? With regard to the first point, it was suggested earlier that semantic encodings are inherently more
distinctive than are phonemic encodings. Words are somewhat similar to each other phonemically; since there is a relatively small pool of phonemes, there is necessarily much overlap of these features in any sample of words. On the other hand, the domain of possible meanings is essentially limitless. If this analysis is correct, it follows that semantic encodings have much more to lose under shared cue conditions. Phonemic encodings are already similar to each other, so sharing has comparatively little further effect. The retrieval of semantic encodings is particularly effective under unique cue conditions since the cue reinstates a highly specific encoding. When the cue is made general, however, performance is lowered, both because the encodings are now more similar to each other (since the cues also form the encoding context) and because the cue no longer specifies one particular encoding with any precision.

Finally, how does the concept of depth stand in the light of these findings? Craik and Lockhart (1972) suggested that when an event is perceived, it is analysed progressively from its shallow structural aspects to its deep semantic aspects. They also suggested that the memory trace reflects those analyses performed, and that deeper analyses are associated with more durable traces. The present results say nothing about the organization of the analyzing system, so no further comments will be made about that aspect of the Craik and Lockhart position. However, with regard to the structure of the memory trace and its subsequent retrievability, the present studies suggest the possibility that the concepts of depth and uniqueness have much in common. Does uniqueness of encoding provide a preferable description of well-remembered events? Although in some ways this is an attractive simplification—for example, it removes the necessity to specify how shallow and deep analyses are interrelated and suggests ways of quantifying depth objectively—nevertheless, there are problems with this description. For instance, nonmeaningful visual patterns are poorly recognized even although each pattern is unique in the experiment (Goldstein & Chance, 1970); similarly, Experiment 8 by Craik and Tulving (1975) showed that the recognition of case-encoded words was not enhanced when the number of such encodings was reduced from 40 to four out of 60 items presented. Apparently uniqueness is not enough; the fact that an item is unique in a set of items to be memorized, or even unique in the person’s experience, is not sufficient to guarantee its memorability. The event must be discriminable and unique semantically, before retention is enhanced. As suggested by Ausubel (1962), well-remembered events must be compatible with the subject’s organized store of past experiences, but they must also be discriminable from such past experiences. This analysis and the present results suggest that the principles of depth and uniqueness are both necessary, in some form, for any final theory of memory.

REFERENCES


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


(Received February 20, 1976)