INSTITUTE FOR THE
STUDY OF
INTELLECTUAL
BEHAVIOR

Remembering the Levels of Information in Words

Mark A. McDaniel, Alinda Friedman, and Lyle E. Bourne, Jr.
University of Colorado

Program* on Cognitive Factors in Human Learning and Memory Report No. 60
Institute for the Study of Intellectual Behavior
University of Colorado
Boulder, Colorado 80309
March, 1977

ABSTRACT

The qualitative nature of the memory trace was examined as a function of incidental perceptual and conceptual processing. Performance on recall and auditory recognition tests replicated the general finding that semantic processing leads to better retention of words than nonsemantic processing. This pattern of results was reversed on a visual recognition test designed to measure the amount of perceptual information remembered. These data suggest that different types of processing result in different aspects of the stimulus being encoded, with conceptual processing resulting in semantic information being encoded and perceptual processing resulting in physical information being encoded. Thus, the effectiveness of a particular kind of processing for good memory performance depends on the kind of information being tested. In addition, evidence is presented that supports simultaneous information extraction.

---

*Supported by NSF Grant BNS 72-02084

The levels of processing notion proposed by Craik and Lockhart (1972; see also Lockhart, Craik, & Jacoby, 1976) has proven to be a productive framework for thinking about the way we remember information. This approach views the processing of events as proceeding through a series of qualitatively distinct domains, beginning with superficial information about color, lines, angles, and the like, and ending in the "deeper," semantic domains of conceptual knowledge. The most distinctive feature of the approach is that memory is seen as a byproduct of the processing which a subject is required to perform on a stimulus. Properties of the memorial trace are therefore a consequence of the type of processing executed.

While a levels of processing framework raises many interesting issues, the literature in general has focused on only one of these, namely, the notion that trace durability is a function of depth of analysis, with deeper levels associated with longer lasting traces (Craik & Tulving, 1976; Jenkins, 1974; McDaniel & Masson, 1977). At least as important, however, is the issue of how processing differences affect the qualitative nature of memory. According to a levels of processing conceptualization, the memory trace resulting from perceptual processing should contain primarily perceptual information, and, likewise, the memory trace resulting from semantic processing should contain primarily meaningful or conceptual information.[3] Incidental learning studies which demonstrate a memory superiority of semantic as compared to nonsemantic processing often assume that different types of processing do yield qualitatively different memory traces (Cermak, Youtz, & Onifer, 1976). These data are subject to another interpretation, however. It may well be that both semantic and nonsemantic processing produce traces which are more or less <u>durable,</u> but which are otherwise qualitatively identical.

Arbuckle and Katz (1976) were aware of this problem, and attempted to resolve the issue by using orienting tasks that involved "yes-no" judgments about the acoustic or semantic similarity of word pairs. Since acoustic judgments produced good recognition of acoustically related pairs, and semantic judgments produced good recognition of semantically related pairs, they concluded that their orienting tasks had indeed produced qualitatively different memory traces. As they note, however, their data could be reinterpreted as meaning that pairs for which the orienting response was "yes" were better remembered than pairs for which the orienting response was "no", Clearly, what is needed is a study which unambiguously tests whether the qualitative nature of the memorial trace is, in fact, a function of the type of processing performed.

There is a second major question implicit in the levels of processing framework, which arises from the idea that a stimulus or event contains different levels of information, any or all of which may be extracted for subsequent processing. Two alternative models have been proposed to describe the manner in which the different levels of information are extracted. Craik and Lockhart (1972) suggest a sequential model in which physical information is processed initially, then structural information, and then deeper, more conceptual information. This view can be contrasted with the notion that the different levels of information are extracted simultaneously though possibily with different time characteristics (Firedman & Bourne, 1976; Goldman & Pelligrino, 1977).

We address these two issues in the following experiments: First, we use a new methodology to investigate whether qualitatively different information is encoded and retained as a function of the type of processing performed on a stimulus. Our theory implies that when semantic information is tested (as in recall), items which were semantically processed will fare better than those which were processed for surface features. The corrolary, which has received little empirical support thus far, is that when perceptual information is properly tested (i.e., when we do not, for example, inadvertently require semantic information for tests of perceptual memory), then perceptually processed (elaborated) items will fare at least as well as items for which processing was unelaboratively carried to deeper levels.

Secondly, we address the issue of sequential-parallel processes. One implication of the sequentiality notion is that upon presentation of a stimulus, physical information should be available for use sooner than conceptual information. Although a parallel model does not necessarily imply faster extraction of physical information, such an assumption could easily be incorporated by proposing that physical information requires generally more efficient encoding procedures than does semantic information. Friedman and Bourne (1976) found that physical information was indeed generally available sooner than conceptual information when subjects were required to make speeded judgments about the common (perceptual or conceptual) features of two stimuli. Importantly, they also found that for this task, subjects processed each stimulus only to the extent needed to perform the task. That is, subjects extracted the two features from each stimulus which were pertinent to the required judgments, and then responded without further processing. Serial and parallel models of information extraction make divergent predictions about what the subject's memory traces would contain following this kind of task. In particular, suppose we required subjects to make speeded inferences on the basis of superficial, physical information (e.g., the case or typefont of printed words). The sequential model implies that only the physical information required for the inferences would be retained. On the other hand, the parallel model allows conceptual as well as physical information to be extracted simultaneously, with processing terminating when the physical information required to perform the task is available. Since, according to the parallel model, some conceptual information, however impoverished, would have been extracted, some trace of it ought to be present for superficially processed items. By administering appropriate memory tests, we may be able to see whether both conceptual and perceptual information is available after perceptual processing of the stimulus items.

In the following experiments, we had subjects make speeded inferences about either perceptual or conceptual aspects of word stimuli. We used this task because it meets the requirement that subjects do only the amount of processing necessary to obtain the relevant information, and no more. In Experiment 1, after completion of a set of speeded inference problems requiring either perceptual or conceptual solutions, subjects were given a battery of memory tests designed to measure the amount of semantic and surface (perceptual) information they had retained. We hoped to find out whether the different levels of processing for speeded inferences influenced the qualitative nature of the memorial trace, and in addition, we hoped to be able to distinguish between the serial and parallel models. We expected that qualitatively different memory traces would manifest themselves in an interaction between the type of speeded inference (perceptual or conceptual) and the particular memory tests we used. Conceptual inferences should yield good memory on tests of conceptual information (recall and auditory recognition), whereas, perceptually processed instances should be recognized at least as well as conceptually processed instances on a memory test of visual feature information. Whether or not visual feature recognition of perceptually processed words will exceed that of conceptually processed words will depend on how much perceptual information was elaborated when the inferences were conceptual. Semantic processing could result in reasonable perceptual recognition regardless of whether information is extracted in serial or in parallel. Assuming that processing ceases when the necessary information is extracted, a serial model does not allow for conceptual recognition of items processed during perceptual inferences.

Experiment 1

Method

Subjects and design. Subjects were 16 undergraduate students, participating to fulfill an introductory course requirement. All subjects solved two types of problems, conceptual and perceptual. One group of 8 subjects solved perceptual problems first and conceptual problems second, and the other group of 8 solved conceptual problems first and perceptual problems second.

Stimulus materials. Eight different 2 x 2 matrices of words, in which each cell of a matrix contained the name of something that was either LARGE or SMALL on the size dimension, and which was a member of either the ANIMAL or WEAPON category were constructed. Each particular instance in a matrix was typed in either UPPER or LOWER case, using either

PLAIN (IBM Letter Gothic) or FANCY (IBM italic) typefonts. Thus, a given set of four instances yielded both a conceptual and a perceptual matrix (see Table 1), and each of the 32 instances used in the 8 matrices exemplified one value from each of the size, category, case, and typefont dimensions. The instances were photographed, and the problems were presented with slides.

A particular matrix yields 8 different trial types; 4 trial types for each dimension with 2 trial types for each value of the dimension. For example, in the conceptual matrix illustrated in the upper left corner of Table 1, HIPPO-tank and tank-HIPPO are two different trial types which require the response "large"; for the perceptual matrix using the same two instances, these two trial types require the response "plain". It should be apparent from Table 1 that using the same four instances for both perceptual and conceptual problems requires only that they be "rearranged" in the matrix. That is, the instances comprising a particular trial type (cell A - cell B) were different, depending upon whether they were in the perceptual or conceptual version of the matrix, but since subjects saw all 8 trial types for every matrix they processed, all subjects viewed each of the 32 instances on 4 separate occasions. Furthermore, although corresponding perceptual and conceptual matrices always had dimensions which were redundant (e.g., size and typefont for Matrix 1 in Table 1; category and case for Matrix 2), the particular dimensions that were redundant were different for each matrix, and were completely counterbalanced across the matrices. Thus, a subject could not learn to use these redundancies, and did in fact have to process the instances along either entirely perceptual or entirely conceptual dimensions in order to perform correctly. A particular subject processed 16 instances (comprising 4 different matrices) for their case and typefont values, and 16 different instances for their size and category values.

In addition to the matrices above, one matrix was constructed with geometric stimuli, and was used for generating practice problems. This matrix was constructed from color (YELLOW or RED) and shape (SQUARE or TRIANGLE) dimensions.

Memory tests. All subjects were given three tests of their memory for the instances, always in the same order. First, subjects were asked to recall as many of the instances as they could. Then, they received a verbally presented recognition test (auditory recognition) which was intended to measure the amount of conceptual or name code information that the subjects had retained. We chose distractor items which were conceptually similar to the instances that the subjects had seen, so that the test consisted of the 32 nouns that subjects had actually seen, randomly interspersed with 32 distractors which exemplified the conceptual dimensions used in the conceptual problems. That is, eight distractors were large animals, eight were small animals, eight were large weapons, and eight were small weapons. Finally, subjects were given a visual recognition test, consisting of only the 32 nouns actually seen during the problems, printed in all four possible styles (e.g., HIPPO, hippo, _HIPPO, hippo_). This test was designed to measure how much subjects remembered about the physical characteristics of the words, when they had been given the name codes.

Procedure. The subjects were seated in front of a rear projection screen. The nature of the problems and solutions were described, and the subjects were led to believe that the variable of interest was how fast they could name the common value for each pair of instances. Nothing in the instructions gave any indication that memory tests would be administered upon completion of the problems.

The subjects were given 16 practice problems with the geometric stimuli, followed by 64 wording problems, consisting of the 8 trial types for each matrix. The conceptual and perceptual problems were presented in blocks of four matrices each; instructions were given at the beginning of each block which described the solution values that were appropriate (e.g., LARGE, SMALL, ANIMAL, WEAPON, or UPPER, LOWER, PLAIN, FANCY), and which alluded to the fact that although the instances might change every once in a while, the solution values would be the same. Each subject processed all 32 instances 4 times each; 16 of the instances were processed for perceptual values, and 16 for conceptual values.

A trial consisted of the following sequence of events: The experimenter pushed a start button; after a delay of about .5 sec, the first slide came on and was replaced by the second instance .8 sec later. The appearance of the second slide activated a millisecond timer, which was stopped when the subject spoke his or her response into a microphone connected to a voice-activated relay. The experimenter recorded the time and initiated the next trial.

Immediately upon completing the last trial, subjects were asked to recall as many of the instances as they could. A maximum of eight minutes was given for this recall test. After the recall test, the experimenter read the 64 words comprising the auditory recognition test; subjects were instructed to say "yes" if they remembered seeing the word and "no" if they did not remember it. Subjects were then given the visual recognition test, and told that it consisted of all the words they had actually seen while solving the problems. They were instructed to circle the exact form of the word as it had been presented in the problems, and were informed that each word had appeared in only one form.

Results and Discussion

The rejection level for all of the following analyses was set at $p < .05$. There was a lower error rate for conceptual problems (2.3%) than for perceptual problems (6.8%), $F(1,14) = 14.08$, $MS_e = 1.17$, reflecting the fact that case and typefont information was more difficult to determine accurately than the size and category information. The mean reaction time for the last block of eight trials on the perceptual and conceptual problem sets was calculated for every subject, with error data excluded. An analysis of variance indicated that there were no differences between the times to solve the two types of problems (the mean time for perceptual problems was 880 msec and the mean for conceptual problems was 942 msec). This outcome is fortuitous, because it precludes an explanation of the memory effects to be described based on different processing times.

The mean number of words recalled and the mean number of words recognized (both in auditory and visual recognition) as a function of type of processing (perceptual or conceptual) and processing order (perceptual or conceptual problems first) are shown in Table 2. These data were examined as follows: First, separate analyses of variance were performed for each of the memory tests, with processing order between-subjects and type of processing within-subjects. Because we were interested in comparing performance on the auditory and visual recognition tests, d' scores (see Table 3) for each subject were used in a 3-way analysis of variance, in which processing order was again between-subjects, and type of processing and type of test was within-subjects.[4]

How well a word was remembered depended on both type of processing and kind of information required by the memory test. That is, memory performance was characterized by an interaction between the type of processing during speeded inferences and the type of test administered. For example, the recall test examined both the retrievability and the durability of items which were processed for either physical or semantic features. In this test, there was a large and reliable advantage (7.38 items) for conceptual (deeper) processing, F (1, 14) = 115.48, MS_e = 3.77. Similarly, the auditory recognition test examined the durability of conceptual or name code information, given physical or semantic processing, and in the absence of specifically visual cues. Again, there was a reliable advantage (8.31 items) for words which were semantically processed, F (1, 14) = 97.12, MS_e = 5.69. In contrast, the visual recognition test examined the durability of perceptual information, given either physical or semantic processing; although performance on this test was generally poor, a reversal of the pattern found for the other two tests was observed. The visual features of 6.31 perceptually processed items were correctly recognized, while the visual features of only 5.06 conceptually processed items were recognized. Although this particular main effect failed to reach significance, F (1, 14) = 3.40, $.05 < p < .10$, it clearly represents a case in which conceptual, or deeper processing of an item did not yield a memory advantage for certain features of that item.

The analysis of variance on the auditory and visual d' scores confirmed the impression that type of processing interacted with type of memory test. The main effect of type of processing, F (1, 14) = 32.60, MS_e = .31, confirmed the notion that deeper processing generally yields better memory performance (d' conceptual = 1.89 and d' perceptual = 1.10). The main effect of type of test, F(1, 14) = 537.65, MS_e = .16, indicates that, in general, the conceptual, name code information tested by auditory recognition was more durable than the physical feature information tested by visual recognition (d' auditory = 2.67 and d' visual = .34). However, both of these conclusions are tempered by the reliable interaction between type of processing and type of test, F (1, 14) = 111.97, MS_e = .15. On the auditory recognition test, words that were conceptually processed for speeded inferences were more often recognized than words that were perceptually processed, while on the visual recognition test, perceptual processing on the speeded inference task yielded more correct recognitions than conceptual processing did.

Thus, semantic processing of items during speeded inference problems appears to yield memorial traces consisting primarily of conceptual, or name code information, which conferred an advantage on these items for both recall and auditory recognition. While semantically processed items were generally remembered better than perceptually processed items, the latter type of processing did yield memorial traces consisting primarily of visual-feature information, which proved advantageous on the visual recognition test.

In addition to the effects discussed so far, there were interactions between type of processing and processing order in recall, F (1,14) = 17.55, MS_e = 3.77, in auditory recognition, F (1, 14) = 5.28, MS_e = 5.69, in visual recognition, F (1, 14) = 4.89, MS_e = 3.68, and in the d' analysis that included both auditory and visual recognition, F (1, 14) = 6.68, MS_e = .31. In general, memory for words that were conceptually processed during the speeded inference problems was better when the conceptual problems were last. Similarly, memory for words that were perceptually processed during speeded inferences was better when the perceptual problems were last. Since the interaction was reliable for all three memory tests, we are dealing with more than just a simple recency effect. Closer examination of the interactions with Newman-Keuls tests suggests that for recall and auditory recognition, both of which rely more heavily on semantic information, the recency effect was larger for perceptually processed items. For example, the perceptual item recall was reliably boosted by an average of 3.62 items when perceptual processing was most recent, while the conceptual item recall was boosted by 2.13 items, which was significant, but also reliably smaller. In auditory recognition, processing order made no difference in the numbers of conceptually processed items that were recognized (14.38 vs. 14.88 for first vs. last items, respectively). As in recall, however, semantic (auditory/lexical) recognition of perceptually processed items was greatly enhanced if those items were processed during the last half of the speeded inference trials; auditory recognition of perceptually processed items increased reliably from an average of 4.63 items to an average of 8.00 items. For the visual recognition test, which required recognition of primarily visual feature information, the variable which seemed to affect performance most was how recently the items were processed; only those items which were processed most recently, regardless of type of initial processing, yielded above chance visual recognition performance $t(7) = 3.33$, $SE_{dm} = .53$ for conceptual processing, and $t(7) = 3.10$, $SE_{dm} = 1.01$ for perceptual processing.

The results described above suggest, first, that there was some name code information available after perceptual processing, but that it was much less durable than the name code information available after conceptual processing. According to a strict interpretation of a serial model of information extraction, only perceptual information should have been available for the perceptually processed items. On the other hand, a parallel model allows all levels of information to be extracted simultaneously, although the different levels may be elaborated differently, depending on the particularly type of information required by the task. It appears that, in the present task, some conceptual information was extracted from words that were perceptually processed. Its lack of stability might reflect the fact that name code information from perceptually processed words, because it is not required for perceptual inferences, is not elaborated.

There is at least one other explanation for our data. It could be argued that name code information is reconstructed, at the time of testing, from the physical information stored during the perceptual speeded inference problems. A comparison of visual and auditory recognition data makes this alternative suspect, however. The d' scores for these two tests showed that auditory recognition was better than visual recognition for <u>perceptually</u> processed words, $t(15) = 10.24$, $SE_{dm} = .13$. If name code information were being reconstructed from perceptual information, and if perceptual information were all that was available for perceptually processed items, then there should have been no difference between these two tests for the perceptually processed items. Since perceptually processed items yielded name code recognition above and beyond the level expected if subjects were merely reconstructing name codes from fragments of perceptual traces, we have some evidence here for the simultaneous extraction of both types of information.

Our most important finding is that the qualitative nature of the memorial trace is a consequence of the processing performed on a stimulus. This is reflected by the fact that semantic processing produced better performance on tests requiring semantic type information (recall and auditory recognition) than did perceptual processing. However, semantic processing did not confer an advantage over perceptual processing on a test (visual recognition) requiring perceptual information about a word.

Still, the results from the visual recognition test were not as clear cut as they could have been, because perceptually processed items did not show a clear and large advantage over the conceptually processed items. The recency of the processing (conceptual or perceptual) seemed far more important for this test; the items which were perceptually processed first were not recognized above chance on the visual recognition test, although the difference was large enough to be encouraging, t (7) = 1.87, SE_{dm} = .80. Yet, the fact remains that for this group, there was no reliable indication of physical information being "stored" as a result of perceptual processing. One possible explanation derives from the notion that physical information deteriorates rapidly and may be susceptible to interference (Cermak, et al. 1976). The visual recognition test was given 15-18 min after completion of the interference problems, and with two other tests completed during that interval. It might be that such a delay, filled with other tasks, would be enough to erase any visual traces resulting from perceptual processing. Experiment 2 eliminated this problem by testing visual recognition immediately upon completion of the speeded inference trials.

Experiment 2

Method

Subjects and design. The subjects were 16 undergraduates, participating to fulfill an introductory course requirement. One subject was replaced after volunteering the information that she was dyslexic. As in Experiment 1, one group of 8 subjects solved perceptual problems first, then conceptual problems, and the other group of 8 subjects solved conceptual problems first, then perceptual problems.

Procedure. The procedure was exactly the same as in Experiment 1, except that recall and auditory recognition tests were not administered.

Results and Discussion

The error rates were 6.5% and 8.9% for conceptual and perceptual problems, respectively. An analysis of variance of the error data failed to yield any significant effects. There was again no difference between the mean reaction times for solving conceptual (891 msec) and perceptual (900 msec) problems for the last block of eight trials on each problem type.

Analyses of the hit rates for the visual recognition test indicated that performance for all cells was significantly higher than expected by chance: conceptual processing first (5.12), t (7) = 2.34, SE_{dm} = .48; conceptual processing second (6.38), t (7) = 2.97, SE_{dm} = .80; perceptual processing first (6.38), t (7) = 2.00, SE_{dm} = 1.19; and perceptual processing second (6.25), t (7) = 3.00, SE_{dm} = .75 (all one-tailed tests). Analysis of variance on these data failed to yield any reliable effects. Thus, with time delay and/or intervening tests eliminated, there is evidence of a perceptual trace even for the first words processed, regardless of the type of processing.

These results offer additional and more robust support for our prediction that when perceptual information is tested, items that are perceptually processed will fare at least as well as semantically processed items. This prediction was derived from the notion that different types of processing result in qualitatively distinct memory byproducts. Specifically, we argue that a memorial trace of physical information is a consequence of processing the physical characteristics of a word. The results from this experiment, in particular, strongly support this contention since every cell showed a detectable memory for the physical information. The fact that physical information was remembered for words that were processed conceptually is not inconsistent with this interpretation; both a serial and a parallel levels of processing model would say that conceptual information processing would result in some perceptual trace.

The corollary to the above prediction is that when semantic information is tested, semantic processed items will fare better than those items which were processed for surface features. In Experiment 1 we found evidence consistent with this corollary: semantic processing provided an advantage over nonsemantic processing for recall and auditory recognition. Yet rather than supporting a levels of processing theory, this outcome might have been a result of the functional stimulus for the two processing conditions being different. Specifically, to solve the conceptual problems, the entire word had to be processed; perceptual problems, however, could have been solved by processing only one letter or some fragment of each word. The subject's goal was to solve each problem as rapidly as possible, and a problem solving strategy that involved glancing at only, say, the first letter in each word would be consistent with this goal for the perceptual problems. If subjects did indeed use such a strategy, then the fact that conceptual processing produced better recall and auditory recognition than perceptual processing is less interesting. Although perceptual processing did produce equal or better performance than conceptual processing in visual recognition, suggesting that subjects processed the entire word during the perceptual inference problems, we felt that eliminating this possible confound would be desirable. Experiment 3 was designed to demonstrate that even when the perceptual problems required processing the entire lexical item, the name codes for those items would not be as well remembered as the name codes for items which were conceptually processed.

Experiment 3

To insure that perceptual problems required processing the entire words, we changed the speeded inference task in the present experiment. Specifically, the typefont dimension was converted to a WORD-NONWORD dimension, so that, in effect, a random half of the perceptual inferences now required a lexical decision on the part of the subject.

Method

Subjects and design. The subjects were 16 undergraduates participating to fulfill an introductory course requirement. As in the other experiments, half of the subjects solved conceptual problems first and half solved perceptual problems first.

Materials. To keep all three experiments as comparable as possible, the 32 instances used in the problems for Experiments 1 and 2 were also used in the present experiment. The conceptual dimensions were again size (LARGE or SMALL) and category (ANIMAL or WEAPON), and the perceptual dimensions were case (UPPER or LOWER) and "wordness" (WORD or NONWORD). Since each perceptual matrix now consisted of 2 words and 2 nonsense stimuli, we needed 16 matrices, rather than 8, in order to present all 32 words. We found 32 additional instances for the conceptual matrices to be used as filler items; 8 items were found for each combination of LARGE or SMALL and ANIMAL or WEAPON. The nonsense stimuli used for the perceptual matrices were then constructed from these filler items by replacing all of their vowels, such that the resulting letter combinations were pronounceable, but meaningless. Table 4 shows an example of the matrices for this experiment. We once again counterbalanced the perceptual and conceptual dimension redundancies across the 16 matrices.

Although each subject saw 16 matrices (64 different instances, 4 times each), we were interested in performance on the 32 critical items only, since these items were the same across perceptual and conceptual conditions. Consequently, the auditory recognition test was the same as that used in the first experiment (i.e., it consisted of the 32 critical items in addition to 32 distractors), but some of the distractors were changed because they had been used as fillers in the additional matrices. The visual recognition test also consisted of only the 32 critical items, and was changed to reflect the fact that there were only two (instead of 4) possible ways in which a word could be typed (i.e., in upper or lower case).

Procedure. The procedure was identical to that of Experiment 1, except that subjects received twice as many problems (8 matrices yielding 64 conceptual problems and 8 matrices yielding 64 perceptual problems).

Results and Discussion

The error rates for conceptual and perceptual problems were 2.9% and 3.9%, respectively. An analysis of variance determined that this difference was not significant. The reaction time difference between the two processing conditions on the last block of trials was larger than in the previous experiments (991 msec for conceptual problems and 875 msec for perceptual problems), but was still not reliable, $F(1, 14) = 3.94$, $.05 < p < .10$. Subjects were apparently not as efficient at extracting the physical features (case) of words, or at making lexical decisions, as they were at extracting aspects of meaning (i.e., size and category information about the words' referents). This result is consistent with the notion of parallel extraction of all levels of information, and that lexical information per se does not necessarily have to precede recovery of deeper, more semantic information.

Table 5 shows the recall and recognition data as a function of type of processing and processing order. As in Experiment 1, we performed separate

Type of Processing X Processing Order analyses on the data from each memory test, and a 3-way analysis (processing order X type of processing X type of test) on the d' scores (see Table 3) for auditory and visual recognition. The separate analyses showed that, once again, conceptual processing procedured more recall, $\underline{F}$ (1, 14) = 33.58, MS_e = 4.05, and better auditory recognition, $\underline{F}$ (1, 14) = 37.69, MS_e = 5.17, than did perceptual processing[5], and there was no difference in the amount of visual recognition as a function of processing type, $\underline{F}$ (1, 14) $<$ 1. Thus, conceptual processing was only advantageous on tests of "conceptual" memory. This conclusion was again confirmed with the d' analysis; conceptual processing (d' = 1.97) produced better overall recognition than perceptual processing (d' = 1.48) , $\underline{F}$ (1, 14) = 16.30, MS_e = .24, and name code information (d' auditory = 2.90), was generally remembered better than case information (d' visual = .54), $\underline{F}$ (1, 14) = 250.79, MS_e = .35, but there was once again an interaction between these two variables, $\underline{F}$ (1, 14) = 31.44, MS_e = .26. There were more conceptually processed words recognized on the auditory recognition test and more words from the perceptual matrices recognized on the visual recognition test. In fact, additional analyses on the data from the visual recognition test showed that performance was above chance regardless of whether the perceptual items were processed first, t (7) = 4.31, SE_{dm} = .67, or last, t (7) = 3.52, SE_{dm} = .71. However, the perceptual information of conceptually processed words was well remembered only if these items were the most recently processed, t (7) = 3.62, SE_{dm} = .76 (all t-tests one-tailed).

The effects of Experiment 1 were essentially replicated in the present experiment. Processing type interacted with test type, and these results were not confounded by the possibility of functional stimulus differences between perceptual and conceptual problems. There were, however, some interesting differences between the results obtained in the first experiment and the results obtained here. While all the interactions between processing time and processing order for each memory test (i.e., the "recency" effects) were reliable in Experiment 1, none were reliable here. The reason appears to be that the scores for the perceptually processed items varied considerably less as a function of processing order in this experiment than they did in Experiment 1. In other words, the trace resulting from having to make lexical decisions was more durable than the trace resulting from case and typefont decisions. We would account for this fact with the argument that the perceptual problems used in Experiment 3 involved "deeper", more elaborate processing of the stimuli (identification of an item as a word) than did perceptual problems in Experiment 1 (identification of typefont and case). This difference between experiments has other implications. In Experiment 3, visual recognition was above chance in both perceptual conditions, while, in Experiment 1, only the most recently processed items were recognized for their perceptual features. The visual recognition test required more than merely identifying typefonts or case. It involved, in addition, establishing the typefont and case associated with a particular word. Thus, the WORD-NONWORD decisions required in the present experiment allowed more elaboration of the "perceptual" information specifically tested in the visual recognition test.

General Discussion

Our approach, like that of others (Craik & Lockhart, 1972), views a stimulus as an event whose inherent significance (how it is understood and remembered) will depend primarily on the use to which it is put in a given situation. In our dealings with the world, especially the "linguistic" world, meaning is far more important than specific visual details; it is no surprise, therefore, that (a) we are normally engaged in "conceptual" processing, and (b) conceptual or semantic processing generally results in good memory. The expectation that "deeper" processing leads to better memory clearly has much intuitive, as well as empirical, support. Beyond this, however, our theory asserts that, in situations which call for "perceptual" elaboration of stimulus events, there should be durable memorial evidence of that processing. Our prediction is that not only how much, but also what is remembered about a stimulus will depend on the type of processing accorded to it, which in turn will be a function of task demands. While this feature of the theory also seems intuitively reasonable, empirical data supporting it have been difficult to come by.

We believe that the three experiments presented here provide the necessary evidence. In these experiments we have investigated the qualitative nature of the memory trace formed as a consequence of different types of processing performed on a stimulus object. Interactions between type of processing and type of test imply that different information about the stimulus items was extracted as a consequence of the processing required by the prior speeded inference task. Conceptually processed words were well remembered on the tests requiring the memory of name-code information, but on the test requiring memory of the perceptual features of these words, performance was often not above chance. On the other hand, perceptually processed words were not remembered as well as conceptually processed words on the tests of name-code information. Yet on the test of the physical information about the word itself, perceptually processed words were remembered as well if not better than conceptually processed words. In fact, in Experiment 3, when the perceptual problems induced subjects to extract specifically the kind of information asked for in the visual recognition, perceptual processing produced a more durable trace than conceptual processing. Thus conceptual processing resulted primarily in semantic information about the stimulus being encoded, while perceptual processing resulted primarily in surface information about the stimulus being encoded.

This research suggests a reexamination of the many studies in recent years (Craik & Tulving, 1975; Jenkins, 1974) demonstrating the superiority of semantic processing for memory performance. Our view implies that, when considering the effectiveness of one or another kind of processing for memory, one must consider the question "What *kind* of memory is this particular type of processing good for?" Tests of conceptual memory will clearly benefit from conceptual processing; tests of perceptual memory will not. Conceptual processing is therefore not categorically "better" than perceptual processing. Different aspects of a stimulus or event are encoded as a function of the task at hand, and are thus remembered as a function of the type of processing which that task involves. That different aspects of a stimulus are remembered becomes apparent when memory tests are designed to be sensitive to different information.

In addition to having the evidence to assert that different aspects of the stimulus are encoded and remembered as a consequence of the processing induced by the task at hand, we can also specify more clearly the temporal dynamics of that information extraction process. The availability of any kind of conceptual information after perceptual processing is clearly inconsistent with a serial model of information extraction. The fact that the conceptual information tested by auditory recognition was recognized when words had been superficially processed supports the view that both perceptual and conceptual information can be extracted simultaneously from a stimulus.

References

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

Cermak, S., Youtz, C. P., & Onifer, W. Retention of semantic and phonemic features of words: differential decay or differential interference? Paper presented at the Psychonomic Society Meetings, St. Louis, Missouri, November, 1976.

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.

Friedman, A., & Bourne, L. E., Jr. Encoding the levels of information in pictures and words. *Journal of Experimental Psychology: General*, 1976, *105*, 169-190.

Goldman, S. R., & Pellegrino, J. W. Processing domain, encoding elaboration and memory trace strength. *Journal of Verbal Learning and Verbal Behavior*, in press.

Jenkins, J. J. Can we have a theory of meaningful memory? In R. L. Solso (Ed.), *Theories in cognitive psychology: the Loyola Symposium*. Potomac, Maryland: Lawrence Erlbaum, 1974.

Lockhart, R. S., Craik, F. I. M., & Jacoby L. Depth of processing, recognition and recall. In J. Brown (Ed.), *Recall and Recognition*. New York: John Wiley & Sons, 1976.

McDaniel, M. A., & Masson, M. E. Long term retention: when incidental semantic processing fails. *Journal of Experimental Psychology: Human Learning and Memory*, in press.

Footnotes

[1] This research was conducted within the Institute for the Study of Intellectual Behavior, University of Colorado, and is Publication No.   of the Institute. The work was supported by Research Grant GB 340-77X from the National Science Foundation. We would like to thank Cathy Brown for collecting the data of Experiment 3.

[2] Now at National Jewish Hospital and Research Center, 3800 East Colfax Avenue, Denver, CO, 80206.

[3] Note that the distinction among levels of information is not intended as an endorsement of any dual or multicode hypothesis. Those hypotheses address the format of memorial representations whereas we are particularly concerned with matters of content.

[4] Across subjects, there were very few false alarms in auditory recognition (6 out of 335 possibilities); in calculating the d' values, the Z score corresponding to an area of 1/2n, where n is equal to the total number of false alarms possible for a given recognition test, was used for a perfect hit (+z) or false alarm (-z) rate.

[5] Recall scores for conceptually processed words appear to be low when compared with the results from Experiment 1. However, in the present experiment, half of the conceptually processed words (the fillers) were not tabulated in the recall protocols. The recall scores reflect only the words recalled that were one of the 32 instances used in Experiments 1 and 2. If all of the words are counted, recall is higher on the average (13.00 items) than it was in Experiment 1 (10.07 items).

Table 1

Examples of Stimulus Materials for Experiments
1 and 2

Conceptual Problems

	Matrix 1		Matrix 2	
	Animal	Weapon	Animal	Weapon
Large	^{A}HIPPO	^{C}tank	^{A}dinosaur	^{C}CANNON
Small	^{B}mouse	^{D}PISTOL	^{B}RABBIT	^{D}hatchet

Perceptual Problems

	Matrix 1		Matrix 2	
	Upper	Lower	Upper	Lower
Plain	^{A}HIPPO	^{C}tank	^{A}RABBIT	^{C}dinosaur
Fancy	^{B}PISTOL	^{D}mouse	^{B}CANNON	Dhatchet

Table 2

Mean Number of Items Remembered as a Function of

Type of Processing, Processing Order, and Type of Test (Experiment 1)

	Test					
	Recall		Auditory Recognition		Visual Recognition	
Processing Order	Conceptual	Perceptual	Conceptual	Perceptual	Conceptual	Perceptual
First	9.00	.88	14.38	4.63	4.38	5.50
Second	11.13	4.50	14.88	8.00	5.75	7.12
Mean	10.07	2.69	14.63	6.32	5.06	6.31

Table 3

Recognition (d') of Lexical/Semantic Versus Perceptual Information

As a Function of Type of Processing and Experiment *

	Test			
	Auditory Recognition		Visual Recognition	
Experiment	Conceptual	Perceptual	Conceptual	Perceptual
1	3.577	1.747	.209	.459
2			.357	.470
3	3.505	2.295	.432	.657

*Experiment 2 did not contain an auditory recognition test.

Table 4

Examples of Stimulus Materials for Experiment 3

Original Matrix

HIPPO	tank
mouse	PISTOL

Matrices Derived for Experiment 3

Conceptual Problems

	Animal	Weapon	Animal	Weapon
Large	HIPPO	howitzer	LION	tank
Small	mouse	whip	turtle	PISTOL

Perceptual Problems

	Word	Nonword	Word	Nonword
Upper	HIPPO	HAWETZIR	PISTOL	LAEN
Lower	mouse	whep	tank	tartlo

Table 5

Mean Number of Critical Items Remembered as a Function of Type of Processing, Processing Order, and Type of Test (Experiment 3)

	Test					
	Recall		Auditory Recognition		Visual Recognition	
Processing Order	Conceptual	Perceptual	Conceptual	Perceptual	Conceptual	Perceptual
First	5.88	2.12	14.25	10.12	8.75	10.88
Second	7.62	3.12	15.12	9.38	10.75	10.50
Mean	6.75	2.62	14.68	9.75	9.75	10.69