

Remembering the levels of information in words

MARK A. McDANIEL, ALINDA FRIEDMAN, and LYLE E. BOURNE, JR.
University of Colorado, Boulder, Colorado 80309

The qualitative nature of the incidental memory trace produced by perceptual and conceptual processing within a speeded inference task was examined. Performance on recall and auditory recognition tests replicated the general finding that semantic processing leads to better retention of words than does 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 primarily in the encoding of semantic information and perceptual processing resulting primarily in the encoding of physical information. Thus, the effectiveness of a particular kind of processing for good memory performance depends on the kind of information being tested.

The levels of processing notion proposed by Craik and Lockhart (1972; see also Lockhart, Craik, & Jacoby, 1976) has proved to be a productive framework for thinking about the way people 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 by-product of the processing that 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, 1975; 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.¹

This research was conducted within the Institute for the Study of Intellectual Behavior, University of Colorado, and is Publication 75 of the institute. The work was supported by Research Grants BNS 72-02084 and BNS 76-81416 from the National Science Foundation. We would like to thank Cathy Brown for collecting the data of Experiment 3. Address reprint requests to Mark McDaniel, Department of Psychology, University of Colorado, Boulder, Colorado 80309. Alinda Friedman is now at National Jewish Hospital and Research Center, 3800 East Colfax Avenue, Denver, Colorado 80206.

This prediction is interesting in itself and raises an important theoretical issue. If different types of processing do result in the storage of different kinds of information, then the type of processing performed on a stimulus *defines* what is in memory. For example, suppose Process A on Stimulus X yields Trace a, which contains relatively "superficial" information, and Process B on Stimulus X yields Trace b, which contains semantic information. Then it stands to reason that the memorial effectiveness of these different encoding processes will be maximized when the type of information tested about X is congruent with the resulting trace (e.g., knowledge of trace a should be better after Process A than Process B and, similarly, knowledge of b should be better after Process B than A). In other words, if memory is tested for knowledge of semantic information, then the processing that yields a trace containing more semantic information will be more effective in producing good performance. We claim that this is precisely what most previous levels of processing studies (e.g., Craik & Tulving, 1975) have concluded. But what about a situation in which memory is tested for knowledge of superficial information? Using the reasoning above, it is logical to assume that nonsemantic processing will benefit performance more than semantic processing on a test for nonsemantic information. Such an observation would suggest that less emphasis be placed on the possibility that deeper processing produces better memory and that researchers focus on an empirical analysis of the relation between different encodings and degree and quality of memory. What is needed is a study that unambiguously tests whether the qualitative nature of the memorial trace is in fact a function of the type of processing performed, and if so, whether there are differences in the durability of these qualitatively different traces.

Several investigators have attempted to show that structural vs. semantic processing of stimuli results in

memorial traces that contain qualitatively different information. For example, Arbuckle and Katz (1976) used 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."

Fisher and Craik (in press) had subjects make one of three types of decisions for each word of a list. The decisions involved saying whether a target word rhymed with another word, or whether it was from the same category as another word, or whether it fit into a sentence frame. The subjects were then asked to recall the targets, either in the presence of the original encoding context, or in a different, but valid context (i.e., if the original pair was bail-HAIL, the cue could be either bail or snow). Better recall was obtained when both the encoding and retrieval contexts were the same, which Fisher and Craik interpreted as an encoding specificity (Tulving & Thomson, 1973) effect. Obtaining such an effect, they reasoned, is evidence for the notion that different encoding conditions produced qualitatively different traces. Importantly, they also found that the semantic encoding/semantic cue condition produced more recall than the rhyme encoding/rhyme cue condition, so that a levels of processing effect was obtained under comparable encoding specificity conditions. However, since recall was their only dependent measure, and since recall of a word appears to involve the ability to reconstruct semantic or name-code information from the memory trace of the word, their data only indirectly test for the presence and durability of perceptual information. That is, in Fisher and Craik's paradigm, subjects were not given a chance to demonstrate the extent to which perceptual or structural information was in fact available to them; they could only demonstrate the extent to which an entire lexical item could be retrieved under different encoding/cue conditions. Their paradigm may therefore be construed as one in which memory for a is demonstrated only if a can be used to generate b.

In the following experiments, we tested directly whether qualitatively different information is encoded and retained as a function of the type of processing performed on a stimulus. We had subjects make speeded inferences (Friedman & Bourne, 1976) about either perceptual, lexical, or conceptual aspects of word stimuli. 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. The particular pattern of results we obtained reveals the manner in which different levels of processing required for speeded inferences influence the qualitative nature of the memorial trace.

If different types of processing in fact do not lead to qualitatively different encodings in memory, that is, if semantic (conceptual) processing simply provides a more durable or stronger encoding of the item than does nonsemantic (perceptual) processing, then we would expect to find for all types of tests that semantic processing results in better performance than nonsemantic processing. On the other hand, if, as the levels of processing framework implies, different types of processing do result in qualitatively different encodings, then we predict one of two possible results. If we assume that semantic processing provides an encoding that contains both semantic and nonsemantic information but that nonsemantic processing provides only, or predominantly, nonsemantic information, then, on a test of semantic information, items that were semantically processed should fare better than those that were processed for surface features. On a test for perceptual information, however, no difference should be found between semantically and nonsemantically processed items. Alternatively, we could assume that semantic processing results in a primarily semantic trace, while nonsemantic processing results in a primarily nonsemantic trace. If this were the case, we would still expect that semantic information should be better remembered after semantic processing. For tests of perceptual information, however, instead of predicting that semantic and nonsemantic processing should produce equal memorial performance, we would then expect that nonsemantic processing should produce better performance than semantic processing.

Regardless of the formulation adopted, it is of interest to determine how well the perceptual information is remembered. Although perceptual processing supposedly results in an encoding of a stimulus that is not very durable with regard to tests for semantic information (as in recall), this does not necessarily mean that perceptual information will not be well remembered if tested directly.

Finally, it should be noted that, while the auditory and visual recognition tests employed in these experiments compare memory performance under "identical" cuing conditions (i.e., the subject is given the entire lexical item in both cases), these tests tap different aspects of one's memory representation. Thus, we are not simply investigating an interaction between encoding specificity and levels of processing; rather, we are testing for the availability in memory of qualitatively different information.

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, which we label conceptual and perceptual. One group of eight subjects solved perceptual problems first and conceptual problems second, and the other group of eight solved conceptual problems first and perceptual problems second. The subjects were tested individually.

Stimulus materials. Eight different 2 by 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 that was a member of either the "animal" or "weapon" category, were constructed. Each particular instance in a matrix was typed in either upper- or lowercase, using either plain (IBM Letter Gothic) or fancy (IBM Italic) typefaces. 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 eight matrices exemplified one value from each of the size, category, case, and typeface dimensions. The instances were photographed, and the resulting slides were used for presenting the problems.

A problem consisted of two instances (shown sequentially) that shared a common value (e.g., animal, weapon, uppercase, etc.), the task being to name the common value. The problems were derived from the eight matrices, with each value of a dimension yielding two separate problems. For example, in the conceptual matrix illustrated in the upper left corner of Table 1, HIPPO-tank and tank-HIPPO are the two problems that can be constructed from the large value of the size dimension; for the perceptual matrix using the same two instances, these two trial types represent the plain value of the typeface dimension. Since each matrix represents two values for each of two dimensions, there are eight different problems that can be constructed from a particular matrix. Subjects were presented with all eight problems for every matrix they processed, so that they viewed each of the 32 instances on four separate occasions.

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. Thus, the eight matrices produce 64 conceptual problems using 32 instances and 64 perceptual problems using the same 32 instances. Each subject processed half of the instances during perceptual problems (i.e., for their case and typeface values) and half during conceptual problems (i.e., for their size and category values). The particular instances in each condition were counterbalanced across subjects. Thus, a particular subject processed half the words for only one type of information but was tested for retention of both perceptual and conceptual information in every instance. Furthermore, although corresponding perceptual and conceptual matrices always had dimensions that were redundant (e.g., size and typeface 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.

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 write down as many of the instances as they could remember. Then, they received a verbally presented recognition test (auditory recognition) that was intended to measure the amount of conceptual or name-code information that the subjects had retained. We chose distractor items that 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 that 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 four-alternative forced-choice visual recognition test, consisting of only the 32 nouns actually seen during the problem, 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 word problems, consisting of the eight possible problems for each of the eight matrices. 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 and which alluded to the fact that, although the instances might change every once in a while, the solution values would be the same. In particular, at the beginning of the block of conceptual problems, subjects were instructed that the solution values were large, small, animal, or weapon. Similarly, to begin the block of perceptual problems, subjects were instructed that the solution values were upper, lower, plain, or fancy. Thus, the subjects were primed for the nature of the solution value so that they would need only to encode both words at one particular level. Since each instance was contained in four out of the eight problems per matrix, each subject processed all 32 instances four 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 for .8 sec. About .75 sec after the offset of this first instance, the second instance was presented. 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.

Table 1
Examples of Stimulus Materials for Experiments 1 and 2

	Conceptual Problems				Perceptual Problems				
	Matrix 1		Matrix 2		Matrix 1		Matrix 2		
	Animal	Weapon	Animal	Weapon	Upper	Lower	Upper	Lower	
Large	HIPPO	tank	dinosaur	<i>CANNON</i>	Plain	HIPPO	tank	RABBIT	dinosaur
Small	<i>mouse</i>	<i>PISTOL</i>	RABBIT	<i>hatchet</i>	Fancy	<i>PISTOL</i>	<i>mouse</i>	<i>CANNON</i>	<i>hatchet</i>

Table 2
Mean Number of Items Remembered as a Function of Type of Processing, Processing Order, and Type of Test (Experiment 1)

Processing Order	Test					
	Recall		Auditory Recognition		Visual Recognition	
	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

Immediately upon completing the last trial, subjects were asked to write down as many of the instances as they could remember. A maximum of 8 min 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%) [$F(1,14) = 14.08$, $MSe = 1.17$], reflecting the fact that case and typefont information was more difficult to determine accurately than the size and category information. The mean reaction time on the last block of eight trials within both the perceptual and the conceptual problem sets was calculated for every subject, with error data excluded. An analysis of variance indicated that there was no reliable difference between the times to solve the two types of problems (the mean time for perceptual problems was 880 msec and for conceptual problems, 942 msec). This outcome is fortuitous, because it weakens an explanation of the memory effects to be described that is 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 three-way analysis of variance, in which processing order was again between subjects, and type of processing and type of test was within subjects.²

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 that 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 [$F(1,14) = 115.48$, $MSe = 3.77$]. 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 that were semantically processed [$F(1,14) = 97.12$, $MSe = 5.69$]. In contrast, the visual recognition test examined the durability of visual 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 only approached 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 [$F(1,14) = 32.60$, $MSe = .31$] supports the notion that deeper processing generally yields better memory performance (d' concep-

Table 3
Recognition (d') of Lexical/Semantic vs. Perceptual Information as a Function of Type of Processing and Experiment

	Recognition Test			
	Auditory		Visual	
	C	P	C	P
Experiment 1	3.58	1.75	.21	.46
Experiment 2			.36	.46
Experiment 3	3.50	2.30	.43	.66

Note—Experiment 2 did not contain an auditory recognition test. C = conceptual; P = perceptual.

tual = 1.89; d' perceptual = 1.10). The main effect of type of test [$F(1,14) = 537.65$, $MSe = .16$] 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 = 2.66; d' visual = .33). However, both of these conclusions are tempered by the reliable interaction between type of processing and type of test [$F(1,14) = 11.97$, $MSe = .15$]. 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 of visual features 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 that contained 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 [$F(1,14) = 17.55$, $MSe = 3.77$] in auditory recognition [$F(1,14) = 5.28$, $MSe = 5.69$], in visual recognition [$F(1,14) = 4.89$, $MSe = 3.68$], and in the d' analysis that included both auditory and visual recognition [$F(1,14) = 6.68$, $MSe = .31$]. 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, this is more than just a simple recency effect. Closer examination of the interactions with Newman-Keuls tests shows 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 that seemed to affect performance most was how recently the items were processed; only those items that were processed most recently, regardless of type of initial processing, yielded above chance visual recognition performance (four items recognized was chance performance) [$t(7) = 3.33$, $SE_{dm} = .53$ for conceptual processing; $t(7) = 3.10$, $SE_{dm} = 1.01$ for perceptual processing].

One might be inclined to interpret the foregoing results as showing that subjects learned how to more efficiently process and store the information needed to perform the inferences so that material processed in the second block of problems was more richly encoded and hence better remembered. If there were indeed a learning-to-learn effect, such that the second block of problems was more efficiently solved (i.e., better information processing and storage), then one would expect this to show up in a comparison of error and latency data between Block 1 and Block 2 problems. An analysis of the latencies and errors with problem block as the variable of interest, however, indicated that there were no differences between blocks of problems. Thus, our interpretation of the interactions in terms of differential durability of traces due to processing type, rather than processing order, seems justified.

Therefore, 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. 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 perceptually processed words [$t(15) = 10.24$, $SE_{dm} = .13$]. If name-code information was being reconstructed from perceptual information, and if perceptual information was all that was available for perceptually processed items, then there should have been no difference between these two tests for the perceptually processed items.

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 or lexical information (recall and auditory recognition) than did perceptual processing. Yet semantic processing did not confer an advantage over perceptual processing on a test (visual recognition) requiring perceptual information about a word.

The results from the visual recognition test were not as clear cut as they could have been, however. The items that were perceptually processed first were not recognized above chance (again, chance is four items) on the visual recognition test. Although the amount of recognition was large enough to be encouraging [$t(7) = 1.87$, $SE_{dm} = .80$], the fact remains that for this group there was no reliable indication of physical information being stored as a result of perceptual processing. Thus, the question remains as to whether conceptual processing yields a trace consisting of the same or less surface information than the trace produced by perceptual processing. One possibility for further exploration derives from the notion that physical information deteriorates rapidly and may be susceptible to interference (Cermak, Youtz, & Onifer, Note 1). The visual recognition test was given 15 to 18 min after completion of the inference 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 additional subject was replaced after volunteering the information that she was dyslexic. As in Experiment 1, eight subjects solved perceptual problems first, then conceptual problems, and the other eight 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 no reliable group differences; performance for all cells was significantly higher than expected by

chance (four items expected by chance) [conceptual processing first: 5.12 items, $t(7) = 2.34$, $SE_{dm} = .48$; conceptual processing second: 6.38 items, $t(7) = 2.97$, $SE_{dm} = .80$; perceptual processing first: 6.38 items, $t(7) = 2.00$, $SE_{dm} = 1.19$; perceptual processing second: 6.25 items, $t(7) = 3.00$, $SE_{dm} = .75$ (all one-tailed tests)]. Thus, with time delay and intervening tests eliminated, there is evidence of a perceptual trace even for the first block of words processed, regardless of the type of processing.

These results offer additional and more robust support for the idea 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 by-products. Specifically, we argue that a memorial trace of physical information is a consequence of processing the physical characteristics of a word. The results from Experiment 2, 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; as outlined in the Introduction, one approach would claim that conceptual information processing results in some perceptual trace.

The corollary to the above prediction is that when semantic information is tested, semantically processed items will fare better than those items that are 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 due to the fact that the functional stimuli for the two processing conditions were 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

Table 4
Examples of Stimulus Materials for Experiment 3

Original Matrix	Matrices Derived for Experiment 3									
	Conceptual Problems						Perceptual Problems			
	Animal	Weapon	Animal	Weapon	Word	Nonword	Word	Nonword		
HIPPO tank mouse PISTOL	Large Small	HIPPO mouse	HOWITZER whip	LION turtle	tank PISTOL	Upper Lower	HIPPO mouse	HAWETZIR whep	PISTOL tank	LAEN tartlo

codes for those items would not be as well remembered as the name codes for items that 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 required a lexical decision on the part of the subject. We acknowledge the fact that these inferences are probably no longer totally perceptually based. We continue to label them "perceptual" only to indicate that on a continuum they are more perceptual than the conceptual inferences.

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 consisted of two words and two 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; eight 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, four 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 Experiment 1 (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 four) possible ways in which a word could be typed (i.e., in upper- or lowercase).

Procedure. The procedure was identical to that of Experiment 1, except that subjects received twice as many problems (eight matrices yielding 64 conceptual problems and eight 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 [$F(1,14) = 3.94, .05 < p < .10$].

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 by processing order analyses of variance on the data from each memory test, and a three-way analysis of variance (processing order by type of processing by type of test) on the d' scores (see Table 3) for auditory and visual recognition. The separate analyses showed that, once again, conceptual processing produced better recall [$F(1,14) = 33.58, MSe = 4.05$] and better auditory recognition [$F(1,14) = 37.69, MSe = 5.17$] than did perceptual processing,³ and there was no difference in the amount of visual recognition as a function of processing type [$F(1,14) = 1.96$]. Thus,

Table 5
Mean Number of Critical Items Remembered as a Function of Type of Processing, Processing Order, and Type of Test (Experiment 3)

Processing Order	Test					
	Recall		Auditory Recognition		Visual Recognition	
	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

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 ($d' = 1.48$) [$F(1,14) = 16.30$, $MSe = .24$], and name-code information (d' auditory = 2.90) was generally remembered better than case information (d' visual = .54) [$F(1,14) = 250.79$, $MSe = .35$], but there was once again an interaction between these two variables [$F(1,14) = 31.44$, $MSe = .26$].

Although there was no significant difference due to processing type on the visual recognition test, subsequent analyses suggested that perceptual processing did provide subjects with a longer lasting visual trace than did conceptual processing. Performance was above chance (eight items represents chance performance) for perceptually processed words regardless of whether the 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 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 Experiment 3. 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 Experiment 1 and the results obtained here. While all the interactions between processing type and processing order for each memory test (i.e., the "recency" effects) were reliable in Experiment 1, none were reliable here. In other words, the trace resulting from having to make the lexical decisions in Experiment 3 was more durable than the trace resulting from case and typefont decisions made during Experiment 1. We would account for this fact with the argument that, as suggested in the Introduction to the present experiment, the lexical decisions involved "deeper," more elaborate processing of the stimuli (identification of an item as a word) than did the perceptual decisions (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.

The results from the present experiments, taken alone, still leave unanswered the question of whether perceptual processing is ever superior to conceptual processing. The three experiments together, however,

show a small but persistent advantage for perceptual processing on the visual recognition test. In fact, when we performed an experiment by processing order by processing type ANOVA on the visual recognition data from all three experiments, the main effect of processing type [$F(1,42) = 4.86$, $MSe = .18$] was reliable, with perceptual processing ($d' = .53$) producing more recognition of perceptual information than conceptual processing ($d' = .33$). In addition, the interaction between processing order and processing type was reliable [$F(1,42) = 5.92$, $MSe = .18$]. The visual features of perceptually processed items were recognized with equal efficiency regardless of order (d' first = .48; d' last = .57), while the visual features of conceptually processed items were only recognized somewhat above chance when those items were the last to be processed (d' first = .16; d' last = .50). These data are some indication that more superficial processing can result in better memory performance than deeper processing when memory for superficial information is tested.

GENERAL DISCUSSION

Our approach, like that of others (Craik & Lockart, 1972), views a stimulus as an event whose inherent significance, that is, 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 generally more important than specific visual details. It is no surprise, therefore, that (1) we are normally engaged in "conceptual" processing, and (2) conceptual or semantic processing generally results in durable memory. The expectation that "deeper" processing leads to better memory clearly has much intuitive, as well as empirical, support. Beyond this, however, we believe that in situations that call for perceptual elaboration of stimulus events, there should be durable memorial evidence of that form of processing. Thus, not only how much, but also what, in particular, 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.

The three experiments presented here 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 and ostensibly unrelated 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 typically 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 somewhat better than conceptually processed words. This was evidenced by the fact that when the results of all three experiments were taken together, we found that perceptual processing produced reliably better performance than conceptual processing for visual recognition. Additionally, in Experiment 3, when the perceptual problems induced subjects to extract specifically the kind of information asked for in 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 reevaluation of the interpretation given to many studies in recent years (Craik & Tulving, 1975; Jenkins, 1974) that have demonstrated 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 that the task involves. That different aspects of a stimulus are remembered becomes apparent when memory tests are designed to be sensitive to different information.

REFERENCE NOTE

1. 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.

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.
- 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.
- 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*, in press.
- FRIEDMAN, A., & BOURNE, L. E., JR. Encoding the levels of information in pictures and words. *Journal of Experimental Psychology: General*, 1976, 105, 169-190.
- JENKINS, J. J. Can we have a theory of meaningful memory? In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium*. Potomac, Md: 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, 1976.
- McDANIEL, M. A., & MASSON, M. E. Long term retention: When incidental semantic processing fails. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 270-281.
- TULVING, E., & THOMSON, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.

NOTES

1. 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 representation, whereas we are particularly concerned with matters of content.

2. 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.

3. 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 1 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).

(Received for publication July 28, 1977;
accepted October 31, 1977.)