JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 21, 165–174 (1982)

Elaborative Encoding as an Explanation of Levels of Processing

GARY L. BRADSHAW AND JOHN R. ANDERSON

Carnegie–Mellon University

Three experiments were performed to study how elaboration of memory structures affects recall accuracy and response latency. The experiments introduce a methodology which can independently manipulate the amount and type of elaborations given to subjects. Using this methodology, it was shown that integrated, highly elaborated memory traces were better recalled than either small unelaborated traces or large, poorly integrated traces. The results have implications for current analyses of levels of processing phenomena, and were found to support the encoding elaboration model (J. R. Anderson & L. M. Reder, *Levels of processing in human memory*. Hillsdale, N.J.: Erlbaum, 1979).

In their original conceptualization, Craik and Lockhart (1972) argued that memory traces were a record of the processing done on a stimulus. Qualitatively different traces were said to result from processing a stimulus in different ways. Researchers characterized processing tasks in terms of the level, or depth, of processing. Early studies (Hyde & Jenkins, 1973; Schulman, 1970) reported meaningful, semantic processing of materials led to superior performance on a subsequent recall or recognition task than performance after shallow, nonsemantic processing. This superiority was explained in terms of greater durability of semantic traces (Kintsch, 1975), or less interference between semantic traces (Wickelgren, 1973) over other types of traces. However, this simple analysis was inadequate to account for differences in recall for material processed at the same, deep level (Bobrow & Bower, 1969). Later accounts expanded upon the original formulation by proposing that processing within a level could differ in terms of its degree of elaboration (Craik & Tulving, 1975), or the discriminability of the memory trace (Lockhart, Craik, & Jacoby, 1976).

Craik and Tulving (1975) performed a series of experiments which they interpreted as showing both effects of depth of processing and degree of elaboration. Replicating earlier research, they found better performance for semantically processed material than for shallow, physically processed material. This result was obtained even when the shallow task took longer to perform than the semantic task. Craik and Tulving felt that differences in elaboration could not account for these results. Such a conclusion was natural from a levels of processing framework, which regards memory as simply a record of perception. Since the shallow analysis took longer to perform, they argued that the memory trace must have been more elaborate than the semantic analysis. They concluded that the memory traces were qualitatively different between these tasks, due to processing items at different levels.

Craik and Tulving observed differences in recall for semantically processed items, contingent upon a yes or no response in the orienting task. They also found better performance when subjects had to judge words in the context of more elaborate sentence frames. These results they thought reflected the effect of degree of elaboration at a particular level. Craik and Tulving concluded that two mechanisms were at work

This research was supported by Grant BNS78-17463 from the National Science Foundation. Requests for reprints should be sent to John R. Anderson, Department of Psychology, Carnegie–Mellon University, Pittsburgh, Pa. 15213. We would like to thank Lynne Reder for her comments on an earlier draft of this manuscript.

0022-5371/82/020165-10$02.00/0

to produce the results observed. Processing tasks with different levels affected the qualitative nature of the memory trace, while differences which occurred at the same level of processing were due to differences in elaboration of the memory trace.

Recently, a major series of criticisms have been leveled at the original levels of processing formulation. Craik and Lockhart (1972) claimed that subjects only encode attributes of a stimulus item required by the orienting task. A question about the sound of a word was said to lead to a trace which encoded that attribute, but not information about the orthographic features or meaning of that word. Numerous experiments have been done (e.g., Bransford, Franks, Morris, & Stein, 1979; Eysenck, 1979; Nelson, 1979) that indicate a memory trace records multiple levels of information. For instance, Nelson, Reed, and McEvoy (1977) showed that phonetic or semantic interference will lower recall independent of the original orienting task. Nelson (1979) argued that orienting tasks have their effect by emphasizing the attributes that are involved by them. Subjects automatically process items at many levels, but may direct further controlled processing to elaborate upon a particular attribute.

Most current models of levels of processing phenomena (Jacoby & Craik, 1979; Nelson, 1979; Eysenck, 1979) accept the general formulation which has been developed here: Subjects will automatically encode most attributes of stimulus materials regardless of the orienting task. Manipulations in the orienting tasks are effective because they direct further controlled processing to different levels. Differences in current models do not center around the initial encoding of information, but rather how the encoded information leads to differences in retrieval. There are three general types of explanations: models that emphasize the match between study and test (Bransford et al., 1979; Tulving, 1979), distinctiveness models (Jacoby & Craik, 1979; Lockhart et al., 1976; Eysenck, 1979), and elaboration models (Anderson & Reder, 1979).

The first class of models (Bransford et al., 1979; Tulving, 1979) account for reduced performance for shallow processing by noting the information extracted in the semantic processing test is more consistent with the retrieval requirements of the surprise recall task than information extracted in the other tests. Bransford et al. refer to this theory as transfer-appropriate processing while Tulving sees this as an extension of his encoding specificity principle. To be somewhat neutral we will refer to this type of account of depth of processing as *encoding-appropriateness*. Our major criticism of this approach is that it is simply a summary of empirical data, rather than a model of the processes involved. As such, it is unable to predict the effects of many study manipulations including the ones to be reported here.

The distinctiveness models (Bransford et al., 1979; Jacoby & Craik, 1979; Nelson, 1979; Eysenck, 1979; Wickelgren, 1973) rest on the assertion that semantically processed traces are more distinct from one another than are traces processed with respect to surface properties like phonetic structure. Calling on past research from the interference literature, these models predict that more distinctive traces will be better recalled. Anderson and Reder (1979) argued that there are many situations where recall for highly distinctive events can be further facilitated by elaborative processing. For instance, Bower, Karlin, and Dueck (1975) demonstrated better memory for distinctive drawings when the drawings were given meaningful interpretations. The experiments to be reported here will contain a direct contrast between the predictions of distinctiveness and elaboration.

Anderson and Reder (1979) proposed an elaboration model which could explain depth-of-processing effects. According to this theory the to-be-remembered material and any elaborations are encoded into a propositional network. Any particular encoded proposition is fragile. There is a significant chance that the subject will not be able to activate that proposition at test. However, if the subject generated a memory episode which encoded a set of multiple propositions that were partially redundant with the to-be-remembered information, he or she would have a much better chance of recalling it at the time of test. There are two basic ways that a more elaborate encoding can lead to better memory. One, which we call *network redundancy,* involves using alternate retrieval paths created by the elaborations. The second possibility, which is more appropriate to sentential material than single word material, is what we call *inferential redundancy.* This refers to the fact that the subject may be able to infer the material he studied from remaining elaborations.

A critical assumption in the theory is that subjects find certain types of elaborations easier to make at study, and are better able to interpret certain types of elaborations at test. A person is best at elaborating that type of information with which he or she has had the most experience. This is because one's elaborative abilities are a function of what one knows about that domain. Semantic processing tends to produce better memory than surface processing because one tends to have a richer knowledge base about the semantic properties of concepts than the surface properties of words.

The present paper explores the hypothesis that the redundancy, or interconnectedness, of propositions is the important factor in improving recall. Previous studies of elaboration effects (e.g., Reder, 1979; Bobrow & Bower, 1969) used instructional manipulations to vary the elaborative process. By allowing each subject to generate his or her own elaborations, the experimenter cannot separate redundancy effects from the number of propositions in a memory episode.

Table 1 illustrates the material that we had our subjects study in the various experiments. In all cases we were interested in recall for the main fact like *Newton became emotionally unstable and insecure as a child.* In the caused-by and resulted-in conditions subjects were provided with two other highly related facts. We assume that these facts are like the kinds of elaborations subjects would generate themselves. The unrelated condition also included two additional facts, but these facts had little connection to the target central fact. Finally, in the single condition, subjects studied just the main fact.

The elaboration model predicts best recall in the related condition. This is because this condition creates the kind of network and inferential redundancy that is postulated to promote good recall. Memory performance should be worst in the unrelated conditions, according to the ACT theory, because these facts only cause interference to the target fact.

The ACT theory contrasts both with the encoding-appropriate models and the distinctiveness models in that it predicts performance should be better in the related conditions than in the unrelated condition. As noted earlier, it is hard to derive any predictions from the encoding-appropriateness models, but it seems particularly unlikely that they should predict a difference between the related and unrelated conditions, because the relationship between study and test is identical for the two conditions. The distinctiveness model seems committed to predicting that performance under the two conditions will be worse than that under the single condition, since the additional facts should lower the distinctiveness of the central fact. If anything, the similarity between the facts in the related conditions should reduce the distinctiveness of the central fact in these conditions, leading to poorest performance in these conditions.

Latency measurements were collected in all three experiments. ACT predictions for such measurements are not clear. We predicted that the unrelated condition should be worse than the single or related condi-

TABLE 1
EXAMPLES OF STIMULI SEEN BY SUBJECTS

Caused-by condition
Newton became emotionally unstable and insecure as a child
This fact was caused by:
Newton's father died when he was born
Newton's mother remarried and left him with his grandfather
Resulted-in condition
Newton became emotionally unstable and insecure as a child
This fact resulted in:
Newton became irrationally paranoid when challenged by colleagues
Newton had a nervous breakdown when his mother died
Unrelated condition
Newton became emotionally unstable and insecure as a child
This fact is unrelated to:
Newton was appointed Warden of the London mint
Newton went to Trinity College in Cambridge
Single condition
Newton became emotionally unstable and insecure as a child

tions, but the ordering of related and single items is uncertain. It should take the subject longer to utilize the network of inferential redundancies in the related condition. On the other hand, subjects in the related condition can still use direct retrieval if that is faster. So, their memory judgments will be some mixture of slow and fast decisions, with the judgment often determined by the fastest retrieval. Without unwarranted assumptions about the relative temporal distributions of the two types of judgments, ACT cannot make a more specific prediction.

A second difficulty in the interpretation of the reaction time data is due to a confounding in the experimental design. Reaction times for the second session were always collected after a cued recall task, which might focus additional elaboration on some of the items. The reaction time results are shown in Appendix 1 for all three experiments. In general, reaction times were faster for single items than for unrelated items. This replicates a frequently-reported result which is known as the fan effect (Anderson, 1974, 1976; Hayes-Roth, 1977; Lewis & Anderson, 1976; Moeser, 1979; Smith, Adams, & Schorr, 1978). The data do not reveal systematic trends in the relative reaction times between the single a[nd] related conditions, and so will not be d[is]cussed further.

EXPERIMENT 1

Method

Subjects. The subjects were 19 Ca[r]negie–Mellon undergraduates who eith[er] received course credit or were paid $3.0[0 an] hour for their participation.

Apparatus. The study materials we[re] presented to each subject individually via [a] cathode ray tube (CRT) terminal controlle[d] by a PDP 11/34 computer. Cued recall wa[s] done as a paper and pencil test. Reactio[n] time trials were again displayed on the CR[T] and responses were recorded via the term[i]nal keyboard.

Materials. The materials for the experi[ment] ment were a set of seven biographical fact[s] for each of 28 famous historical figures. Th[e] historical people were chosen so that mos[t] subjects should recognize each figure, bu[t] would have little actual knowledge abou[t] the lives these people led. We decided t[o] use real people rather than making up arbi[trary] trary names because we wanted our sub[jects] jects to treat the materials as naturally as possible. The memorization task should be easier since learning the names of people in the experiment is not necessary. In addition, we felt this design would lead to fewer confusions among responses during the cued recall task. Most of the facts were drawn from biographies of the people, and were true. Occasionally, a fact was altered or fabricated to fit the constraints of the experimental design. We felt the resulting plausibility might encourage subjects to treat the facts as real, instead of fictional curiosities of no import outside the laboratory.

Each subject saw only a subset of the facts about each historical figure. One of the facts for each person, called the *central* fact, was common to all four subsets. The relationship between this fact and the others shown in the subset, called the *supportive* facts, determined the condition of the set. The first condition, the caused-by condition, included two supportive facts which were plausible reasons for the central fact. The two causal facts were designed to be as unrelated to each other as possible, and independently imply the central fact. Thus, if subjects were able to recall either causal fact alone, they should have been able to generate inferences from that fact which might contain the central fact. If they could recall both causal facts, the intersection of the plausible inferences should be a very restricted set which included the central fact. In the second condition, the resulted-in condition, two supportive facts were each plausible outcomes of the central fact. Again they were chosen to be as independent of each other as possible. The third condition, the unrelated condition, included two supportive facts about each person which were unrelated to the central facts and to each other. In the final condition, the central fact was shown alone. Historical figures were assigned to one of the four conditions (single, caused-by, resulted-in, unrelated) on a random basis for each subject, with the constraint that each subject saw an equal number of figures in each condition. Whenever multiple facts were shown about a figure, subjects were explicitly told the relationship between the central fact and the other facts. Historical figures were randomly assigned to conditions such that seven figures appeared in each condition. Every subject received an independent randomization of figures to conditions.

Procedure. The experiment was conducted in two sessions, the second session was held exactly 1 week after the first. The first session consisted of four tasks, performed in a fixed order: A memorization phase for the materials; a dropout phase, where subjects received additional practice on the materials; a cued recall phase, and a reaction-time verification phase. In the memorization phase, each set of facts about a figure was presented one at a time on the screen. Single fact sets were displayed for 18 seconds, while all other sets were displayed for 36 seconds. The subject was asked to memorize the facts presented about each person.

In the next part of the experiment, the dropout phase, the sentence was presented with the subject of the sentence replaced by the word Who or Whose. The subject responded to each query by typing in a name. Feedback was given, then the subject was asked to specify how that fact was related to others in its set. Again feedback was given. If the subject correctly identified both the name and the condition of a fact, it was eliminated from the completion set. After all items had been tested, subjects were retested on items which remained in the completion set. The process was continued until every fact was correctly identified once.

On completing the study phases, subjects were given a sheet of paper containing the names of each historical figure, and were asked to write down all the facts they could remember about each person. In addition, they were asked to identify how the facts were related. Subjects were allowed to complete this phase at their own pace. The final task in the first session was a reaction time task. In this task, a sentence appeared on the terminal screen, and subjects were asked to decide whether or not the fact was one they had studied earlier. Foils were constructed by randomly mis-pairing names and sentence stems. At the end of the reaction time trials, subjects were reminded of the second session, but were not given any information about the tasks in the second session.

The second session consisted of two tasks, a cued-recall task and a reaction time task. Each of these tasks was similar to the corresponding task in the first session. In the reaction time task, the same foils were used, but all items were presented in a different random order.

Results and Discussion

In order to analyze the initial ease of learning particular items, we calculated the average number of trials for items in each condition to be correctly completed in the dropout phase. The mean number of trials to dropout is 1.43, 1.37, 1.66, and 1.72 for the caused-by, resulted-in, unrelated, and single conditions, respectively. An analysis of variance revealed a significant main effect of item condition ($p < .005$). This analysis was performed by averaging together scores for all facts which were presented in a given condition. The experimental design required supportive facts to vary by conditions. Only the central facts were constant for each possible condition (see Table 1). The results could have been influenced by spurious differences in supporting items in the various conditions. Therefore, a second analysis was performed on the average dropout rates of central items only. Dropout rates in the conditions were 1.36, 1.39, 1.55, and 1.72 for the caused-by, resulted-in, unrelated, and single conditions, respectively. An analysis of variance again showed a significant effect of item condition ($p < .01$). Orthogonal comparisons showed no difference in dropout rates for related conditions. The difference between unrelated and single items was not significant ($t = 0.96$). However, the mean of these items took significantly longer to learn than related items ($p < .01$). Because of differences in supportive items across conditions, future analyses will only be reported for central items. Unless otherwise mentioned, analyses of all items showed a similar pattern to analyses of central items.

In the cued-recall task, items were scored for gist. Table 2 shows recall rates for both sessions by condition. An analysis of variance showed a significant effect of condition type ($p < .001$). The effect of delay approached significance ($p < .086$), while the interaction was not significant. Separate orthogonal comparisons for each session showed recall rates for unrelated items to be lower than that of single items, which did not differ from related items. Recall rates in the related and single conditions were quite high. We had hoped to show that initial learning was equivalent in all conditions. However, trying to remember multiple, unrelated items about a single historical figure was so difficult that, even though subjects received more practice on these items than on related items, immediate recall was somewhat reduced.

The results of this experiment were generally supportive of the ACT elaboration model. Clear differences were found between the related and unrelated condition in terms of recall of central facts. However, this experiment failed to find a significant advantage of the related conditions over the single condition in percentage recall. We speculated that this might be due to high levels of recall displayed in these conditions. Therefore, the second experiment was designed to achieve lower levels of recall.

Experiment 2

Experiment 2 is basically a repetition of Experiment 1 with a slight change in the design. We decided to reduce initial learning of items, to see if the performance on single and related sets was similar due to

TABLE 2
PERCENTAGE RECALL RATES FOR CENTRAL ITEMS BY CONDITIONS AND DELAY

Experiment	Condition			
	Caused-by	Resulted-in	Unrelated	Single
Experiment 1				
Immediate Recall	93	96	80	93
Week Recall	89	92	74	87
Experiment 2				
Week Recall	70	75	45	61
Experiment 3				
Week Recall	52	61	32	38

ceiling effects. We chose to omit the initial cued recall task. Since initial recall levels were so high, these scores did not reveal much about the effects of the experimental manipulation.

Method

The method of Experiment 2 was similar to that of Experiment 1, with differences noted below.

Subjects. Subjects were 41 Carnegie–Mellon University undergraduates who either received course credit or $3.00/hour for their participation.

Procedure. Experiment 2 took place in two sessions, like Experiment 1, 1 week apart. The only difference in procedure took place in the first session, which did not have an initial cued recall task. Thus, the first session had three phases: the study phase, the dropout phase, and the reaction time verification phase. The procedure in the second session was identical to the first; a cued-recall task followed by a reaction time task.

Results and Discussion

The average number of trials for central items to be completed in the dropout phase are: 1.48, 1.46, 1.64, and 1.67 for the caused-by, resulted-in, unrelated, and single items, respectively. An analysis of variance confirmed item condition to have a significant main effect ($p < .015$). Dropout performance was similar to that obtained in Experiment 1. Again they reflect difficulty in learning single and unrelated items.

Table 2 shows cued recall rates for items tested in the second session. An analysis of variance on recall rates revealed a significant effect of item condition ($p < .001$). As in Experiment 1, recall rates in the unrelated condition were significantly poorer than in other conditions. Unlike Experiment 1, recall rates for single items were somewhat below the mean of related items. Although this difference was not statistically significant ($t = 1.33$), the overall pattern more closely conforms to predictions made from the encoding elaboration model. Also not significant, but consistent with Experiment 1, the resulted-in condition had slightly higher performance than the caused-by condition.

Experiment 2 thus replicated most of the findings of the first experiment and is entirely consistent with the elaboration model. Again we observe the powerful effect of having to remember multiple unrelated propositions in lowering the level of recall. Although recall of central items in related conditions was not significantly higher than recall of single items, the trend in Experiment 2 was clearly in this direction. Accordingly, we decided to reduce the level of initial learning still further in Experiment 3, to see if a significant separation of these conditions could be achieved.

Experiment 3

Experiment 3 used the basic design developed in Experiments 1 and 2. Two changes to this design were made, however. The level of practice with materials in the first session was reduced still further, and in the reaction time task administered in the second session, a new set of foils was used. These foils were constructed to be thematically related to targets. In general, the use of such foils prevents subjects from making reaction time decisions on the basis of thematic information (Reder & Anderson, 1980), and so the fan effect should be present for both related and unrelated items.

Method

The method of Experiment 3 was similar to that of Experiment 2, with differences noted below.

Subjects. Subjects were 28 Carnegie–Mellon University undergraduates who either received course credit or $3.00/hour.

Procedure. Experiment 3 again took place in two sessions, 1 week apart. The reaction time task was eliminated from the first session. It consisted of two parts, the study phase and the dropout phase. The second session was identical to that of the first experiments, except that the reaction time task used thematically related foils instead of unrelated foils.

Results and Discussion

The average number of trials for central items to be completed in the dropout phase are: 1.53, 1.52, 1.70, and 1.72, for the caused-by, resulted-in, unrelated, and single conditions, respectively. An analysis of variance did not reveal a significant effect for item condition, although the overall pattern was similar to that obtained in the first two experiments.

Table 2 shows cued recall rates for items tested in the second session. An analysis of variance performed on recall rates revealed a significant effect of item condition ($p <$ .001). Orthogonal comparisons revealed th[e] single item recall rate was significantl[y] lower than the mean of the related items (p $<$ .001). The trend toward a separation o[f] these conditions in Experiment 2 wa[s] strong enough to reach significance in th[e] present experiment. The data are thus completely in accord with the pattern predicted by the encoding elaboration model.

General Discussion

The experiments reported here show a consistent and powerful effect of item organization on recall of materials. In every experiment, we found recall performance for the central fact to be best when it was supported by related facts, next best when alone, and worst when studied along with unrelated facts. This pattern of data supports the ACT model. As mentioned in the introduction, both the encoding-appropriateness model and the distinctiveness model do not make clear predictions about the ordering of related and unrelated conditions. In addition, the distinctiveness model does not predict the superiority of the related condition to the single condition in terms of percent recall. Thus we feel the elaboration model to be the most useful one to explain depth of processing effects.

An alternative formulation of the distinctiveness model might be that *sets of traces* form integrated units, which are more or less distinctive. If this interpretation is accepted, it is unclear why we should expect differences between the unrelated and related sets of items. In fact, unrelated sets of items should be highly unusual, and therefore very memorable. It has been suggested that this problem could be circumvented by postulating that unrelated items are not collected into sets, while related items are. However, this just points out the vagueness of current distinctiveness model formulations.

The experiment found a consistent superiority of the resulted-in condition over the caused-by condition. Although the differences never approached significance, they represent an anomaly. Note that this comparison is based on the same central facts; the conditions only differ in terms of the supporting facts. We would like to think that it reflected a weaker relationship among the facts in the caused-by condition. However, our efforts to establish that this was the case have not been successful. We had a separate group of 20 subjects rate the mean interrelatedness of the triples of facts in the caused-by and resulted-in condition. Resulted-in triples had an average relationship of 5.12 on a 7-point scale (7 was highly related), whereas caused-by facts had a mean of 5.37 on the same scale. Also, we found little correlation between these ratings for individual historical figures and recall performance on those figures ($r = -.01$). Thus, our rating measures were failing to tap important factors influencing recall. It is perhaps not surprising that subjects have a poor sensitivity to the kind of interconnectedness that makes for good recall.

The results of these experiments are somewhat at odds with the results reported by Reder and Anderson (1980). These authors contrasted memory of text passages with memory of summaries of those passages. In their *text* condition subjects studied the original text while in the *summary* condition the subjects studied summaries that just contained the main facts. These summaries were between one-fourth and one-fifth as long as the original texts. Reder and Anderson have consistently found better recall in the summary condition.

There are clear analogies between the summary condition and our single condition and the text and our related condition. We speculate that the reason why Reder and Anderson found the opposite result was that their textbook facts were simply less related than our related facts. However, there are some potentially significant methodological differences between the Reder and Anderson situation and our own. In our experiment, subjects had more study time in the related condition than in the single condition whereas this tended not to be the case in the original Reder and Anderson experiments. More recently, Reder and Anderson (in press) have performed an experiment that showed study time is not the relevant factor.

APPENDIX

Reaction Times (sec) and Percent Error Rates for Central Items by Conditions and Delay

Experiment	Condition			
	Caused-by	Resulted-in	Unrelated	Single
Experiment 1				
Immediate test	1.80 (3.7)	1.62 (7.5)	1.77 (7.5)	1.55 (2.2)
Week test	1.77 (4.5)	1.67 (8.3)	1.68 (9.0)	1.54 (3.7)
Experiment 2				
Immediate test	1.60 (6.6)	1.52 (4.5)	1.77 (5.9)	1.60 (4.9)
Week test	1.70 (6.3)	1.67 (7.3)	1.79 (9.1)	1.76 (5.2)
Experiment 3				
Week test (correct items)	2.08 (17.9)	2.10 (1.43)	2.15 (9.7)	2.11 (12.8)
Week test (foil items)	2.20 (13.2)	2.34 (10.7)	2.17 (10.7)	2.14 (11.2)

REFERENCES

ANDERSON, J. R. Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 1974, **6**, 451–474.

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

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

BOBROW, S., & BOWER, G. H. Comprehension and recall of sentences. *Journal of Experimental Psychology*, 1969, **80**, 455–461.

BOWER, G. H., KARLIN, M. B., & DUECK, A. Comprehension and memory for pictures. *Memory and Cognition*, 1975, **3**, 216–220.

BRANSFORD, J. D., FRANKS, J. J., MORRIS, C. D., & STEIN, B. S. Some general constraints on learning and memory research. In L. Cermak & F. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J.: Erlbaum, 1979.

CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 671–684.

CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, **104**, 268–294.

EYSENCK, M. W. Depth, elaboration, and distinctiveness. In L. Cermak & F. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J.: Erlbaum, 1979.

HAYES-ROTH, B. Evolution of cognitive structures and processes. *Psychological Review*, 1977, **84**, 260–278.

HYDE, T. S., & JENKINS, J. J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 471–480.

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

KINTSCH, W. Memory representations of text. R. L. Solso (Ed.), *Information processing a cognition*. Hillsdale, N.J.: Erlbaum, 1975.

LEWIS, C. H., & ANDERSON, J. R. Interference w real world knowledge. *Cognitive Psycholog* 1976, **8**, 311–335.

LOCKHART, R. S., CRAIK, F. I. M., & JACOBY, L. Depth of processing, recognition and recall: So aspects of a general memory system. In J. Bro (Ed.), *Recall and Recognition*. London: Wil 1976.

MOESER, S. D. The role of experimental design in vestigations of the fan effect. *Journal of Expe mental Psychology: Human Learning and Me ory*, 1979, **5**, 125–134.

NELSON, D. L. Remembering pictures and words: A pearance, significance, and name. In L. Cerm & F. Craik (Eds.), *Levels of processing in hum memory*. Hillsdale, N.J.: Erlbaum, 1979.

NELSON, D. L., REED, V. S., & MCEVOY, C. L. E coding strategy and sensory and semantic inte ference. *Memory and Cognition*, 1977, **5**, 462–46

REDER, L. M. The role of elaborations in memory f prose. *Cognitive Psychology*, 1979, **11**, 221–23

REDER, L. M., & ANDERSON, J. R. A partial resol tion of the paradox of interference: The role integrating knowledge. *Cognitive Psycholog* 1980, **12**, 447–472.

REDER, L. M., & ANDERSON, J. R. Effects of spaci and embellishment on memory for the main poin of a text. *Memory and Cognition*, in press.

SCHULMAN, H. G. Encoding and retention of semant and phonemic information in short-term memor *Journal of Verbal Learning and Verbal Behavio* 1970, **9**, 499–508.

SMITH, E. E., ADAMS, N., & SCHORR, D. Fact r trieval and the paradox of interference. *Cognitiv Psychology*, 1978, **10**, 438–464.

TULVING, E. Relation between encoding specificit and levels of processing. In L. Cermak & F. Crai (Eds.), *Levels of processing in human memory* Hillsdale, N.J.: Erlbaum, 1979.

WICKELGREN, W. A. The long and the short of mem ory. *Psychological Bulletin*, 1973, **80**, 425–438.

(Received August 14, 1981)