Journal of Experimental Psychology: Human Learning and Memory
1976, Vol. 2, No. 6, 680–687

Depth of Processing in Recall and Recognition Memory: Differential Effects of Stimulus Meaningfulness and Serial Position

John G. Seamon and Pauline Murray
Wesleyan University

Structural and semantic levels of processing were distinguished in two experiments that varied stimulus meaningfulness in an incidental learning paradigm. Meaningfulness had a significant positive effect on recall and recognition memory for subjects attending to word meaning but no effect for subjects orienting to the position of their lips during word vocalization. An analysis of the primary and secondary memory components of recall found levels of processing to influence only the secondary component. A signal detection analysis of the recognition data showed large differences in memory sensitivity, d', between groups; accuracy and confidence were higher for people orienting to semantic factors. Lastly, an intentional learning control group showed meaningfulness effects comparable to those of the semantic group in both experiments. It was suggested that orienting tasks affected the depth of processing, while meaningfulness affected the spread of encoding. Subjects may be able to determine the depth of processing by varying their perceptual analysis; spread of encoding, however, would seem to be determined by the structure of semantic memory.

Craik and Lockhart (1972) hold that memory is a by-product of perceptual analysis. Retention is assumed to reflect the level or depth of information processing, from structural analyses of the physical aspects of an event to more complex analyses that abstract meaning. Perception and memory are viewed as indivisible, with retention the natural and automatic consequence of specific perceptual analyses.

There is much recent research on the memorial outcomes of different perceptual analyses. Virtually all of this work has been conducted within the context of an incidental learning paradigm which, according to Craik and Lockhart (1972), provides a relatively pure measure of retention for a given processing level. Craik (1973, Experiments 4 & 5) found that recall and recognition memory were greater when semantic decisions were made about words than when letter case or acoustic decisions were made. These results have been found in both free recall (e.g., Craik & Tulving, 1975; Hyde & Jenkins, 1969) and recognition (e.g., Craik & Tulving, 1975; Elias & Perfetti, 1973).

The present research was conducted to test several hypotheses from the levels-of-processing approach concerning the effect of stimulus meaningfulness and serial position on recall and recognition memory. In addition, this research sought to determine if levels of processing influenced primary or secondary memory in recall and memory sensitivity in recognition.

Experiment 1

This experiment measured the effects of orienting task, stimulus meaningfulness, and serial position on free recall. The variable of central interest was stimulus meaningfulness. If meaningfulness differentially affects recall as a function of the orienting task employed, such that it has no effect with a structural orienting task but has a major effect with a semantic orienting task, then the distinction be-

This research was supported by a Grant-in-Aid of Research awarded to the first author from Sigma Xi, The Scientific Research Society of North America. Appreciation is expressed to the reviewers for their thoughful comments. The second author is now at the University of Connecticut.

Requests for reprints should be sent to John G. Seamon, Department of Psychology, Wesleyan University, Middletown, Connecticut 06457.

tween structural and semantic levels of processing is supported.

It is well documented that meaningfulness is important to memory and should, therefore, influence recall under intentional instructions. This variable should not, however, affect the recall of subjects attending only to structural aspects of the stimuli if these aspects are abstracted independently of meaning. The effect of stimulus meaningfulness on the performance of subjects orienting to semantic features is less obvious. If depth of processing is optional, recall of high meaningful words may be better than that of low meaningful words under standard free recall conditions because with a fixed rate of presentation, high meaningful words can be processed "deeper" than low meaningful words. Use of a self-paced task and the same orienting rule for both high and low meaningful words would appear to minimize the effectiveness of this variable and its potential for differentiating processing levels. Alternatively, depth of processing may be obligatory in that high meaningful stimuli will always be processed to a deeper level than low meaningful stimuli when a semantic orienting task is employed (Craik & Lockhart, 1972, p. 676). If this is the case, high meaningful words should be recalled better than low meaningful words with a semantic orienting task even though the orienting rule applied to both types of stimuli is the same. The interaction of meaningfulness with orienting tasks would suggest that processing levels could be differentiated and that a level may not be defined solely in terms of the orienting task employed. Rather, level of processing would be defined by both the orienting task and the object oriented to. A similar argument was made by Jenkins (1974), noting the parallel of his work to Gibson's (1966) view of perception.

A second variable of interest is serial position. Intentional learning controls should show the usual bow-shaped curve. This follows from either a dual-storage model (Glanzer & Cunitz, 1966) or a levels-of-processing approach (Craik & Lockhart, 1972). In the latter case, the primacy effect is said to result from the use of more effective or deeper coding strategies on initial items than on subsequent items; the recency effect reflects the output of items that are still in short-term or primary memory at the time of recall.

Control of the processing activities with orienting tasks should modify the serial position function. Studies of intentional learning in which each stimulus is rehearsed only while in view (e.g., Fischler, Rundus, & Atkinson, 1970) have shown little primacy in recall. A similar lack of primacy should be found with orienting tasks since the coding conditions are equated across all serial positions. Type of orienting task may, however, interact with serial position. Recall of terminal items should not be affected by type of orienting task because these items are still in primary memory at the time of recall or are coded in forms that are equally effective at generating recall if it is immediate. Recall of nonterminal list items, on the other hand, may be better with a semantic orienting task. The availability or accessibility of different codes may change over time with semantic codes forgotten less rapidly than codes based on structural features (Craik, 1973); alternatively, information transfer from short- to long-term storage may vary with coding strategy (Modigliani & Seamon, 1974). This finding would imply that a level of processing influences only the secondary memory component of free recall.

Method

Subjects. The subjects were 36 Wesleyan undergraduates, 12 in each condition, who served as paid volunteers.

Apparatus. The stimuli were presented on a rear projection screen by a Gerbrands Projection Tachistoscope with manual response times obtained from a Durgin and Brown millisecond timer. The timer was activated by the opening of the stimulus shutter on each trial and terminated with the closure of a microswitch located under either response key. The subject was comfortably seated at a desk and shielded from the tachistoscope and timer by a wall partition.

Stimulus materials. Two lists of 60 English nouns were obtained from the Paivio, Yuille, and

TABLE 1

Mean Decision Time and Recall as a Function of Experimental Task and Stimulus Meaningfulness in Experiment 1

Condition	Decision time	Recall
Control		
High M	—	10.33
Low M	—	7.16
Total	—	17.49
Structural		
High M	1.72	3.09
Low M	1.80	3.08
Total	1.76	6.17
Semantic		
High M	1.34	7.84
Low M	1.53	5.09
Total	1.44	12.93

Note. Recall is expressed in terms of number of words and decision time in seconds, with total decision time the mean of all decision times. M = meaningfulness.

Madigan (1968) word norms. Each list was composed of 30 words, rated high in meaningfulness, imagery, and concreteness (mean ratings of 7.11, 6.02, and 6.19, respectively) and 30 words rated significantly lower on these dimensions (mean ratings of 4.70, 4.01, and 3.72, respectively). Only meaningfulness will be discussed in this paper, as a pilot study found that varying imagery and concreteness over these ranges had no effect upon recall when meaningfulness was held approximately constant.

Meaningfulness was varied within lists by alternating high and low meaningful words across adjacent serial positions within a list. One list started with a high meaningful word and the other list began with a low meaningful item. Each list was presented to half of the subjects in each condition.

Procedure. Subjects were alternatively assigned to one of the two incidental learning groups or the control condition in order of their appearance in the laboratory. Incidental learning subjects thought they were participating in an experiment on decision making. They were told that they were going to be shown a list of words, with the time to make a binary decision about each word recorded. Decisions were made by pressing either a left or a right response lever with the left or right index finger. Subjects in the semantic group were instructed to orient to the meaning of each word and to decide as quickly and as correctly as possible if that word was a general (e.g., TOOL) or a specific (e.g., HAMMER) instance of a semantic category by pressing either the left (general) or the right (specific) response key. None of the lists contained a general and specific instance of the same semantic category. People in the structural group were given the same experimental set but were told to attend to the position of their lips while subvocally repeating each word during list presentation. They were told to press the left lever if their lips were touching at either the beginning (e.g., MANE) or end (e.g., TOMB) of a word, or the right lever if their lips were touching at both the beginning and the end (e.g., BOMB) of a word or neither the beginning nor the end (e.g., CLOCK).

A trial in the semantic and structural conditions took approximately 10 sec. A word typed in uppercase letters was shown for 2 sec, and the subject made a decision according to the appropriate orienting rule. After the lever response, the subject repeated the word orally and then gave a verbal rationale for the decision. The experimenter recorded the response time while the subject explained the decision, then asked if the subject was ready, and initiated the next trial. Immediately after the last trial, the subject was given pencil and paper and asked to recall in any order as many words as possible. Approximately 3 min. was given for free recall.

Control subjects were told that this was an experiment on memory and were given standard free recall instructions. The same word lists were used in this condition, but the items were shown at a fixed 2-sec rate with a blank 2-sec interstimulus interval. Again, 3 min. were allowed for free recall.

Results and Discussion

The mean number of words recalled by subjects under all conditions is shown in Table 1 along with the mean decision times where appropriate. It may readily be seen that the semantic group recalled more words than the structural group,

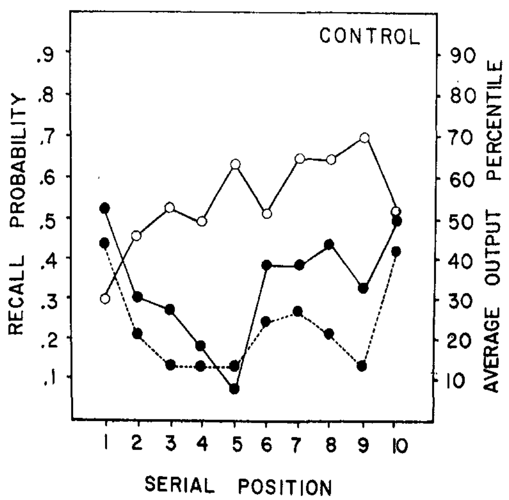

FIGURE 1. Mean probability of recall of high (solid line) and low (dashed line) meaningful words over serial positions and mean output percentile (open circles) over serial positions for the control group. (Each point on the abscissa represents a block of six serial positions.)

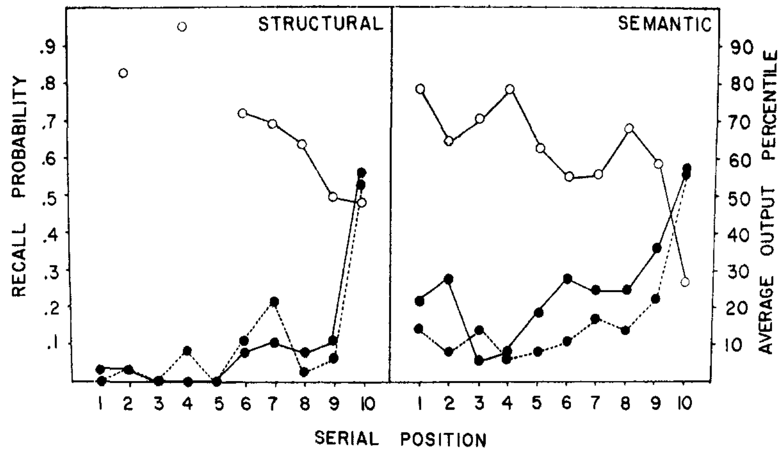

FIGURE 2. Mean probability of recall of high (solid line) and low (dashed line) meaningful words over serial positions and mean output percentile (open circles) over serial positions for the structural and semantic orienting groups. (Each point on the abscissa represents a block of six serial positions. Nonrecall or very low recall resulted in several missing output percentile points for the structural group.)

although decision time was less in the former condition than in the latter. Meaningfulness affected performance of both control and semantic groups; subjects in these conditions recalled more high meaningful than low meaningful stimuli. Meaningfulness, however, did not affect recall in the structural group.

Since presentation rate was fixed for control subjects and self-paced for subjects with the orienting tasks, no direct comparisons can be made between experimental groups and the control group. An analysis of variance on the number of words recalled showed a significant difference between the semantic and structural groups, $F(1, 22) = 24.59$, $MS_e = 5.56$, and a significant interaction of Groups $\times$ Meaningfulness, $F(1, 22) = 24.80$, $MS_e = .91$, indicative of an effect of meaningfulness on the semantic group but not on the structural group. Like the semantic group, there was a significant effect of meaningfulness for the control group, $t(11) = 3.30$, with more high meaningful items recalled than low. A $p < .05$ rejection region was used in all analyses.

The decision-time data for the two incidental learning groups were also analyzed by an analysis of variance. As suggested in Table 1, there was a significant difference between groups, $F(1, 22) = 10.60$, $MS_e = .12$, with the semantic group responding faster than the structural group, and a significant effect of meaningfulness on decision time, $F(1, 22) = 25.77$, $MS_e = .009$. Only a marginal interaction of Groups $\times$ Meaningfulness was observed, $F(1, 22) = 4.05$, $p < .10$, $MS_e = .009$. An examination of decision time across serial position found the difference between conditions to remain relatively constant over the entire list.

The above results are consistent with previous research in showing an effect of levels of processing on free recall. The interaction of meaningfulness and processing levels supports the distinction between structural and semantic levels of analysis. Further, the results are in agreement with Craik and Lockhart's (1972) assertion that processing time per se is unimportant and that high meaningful items will be processed faster and retained better than low meaningful items.

Mean recall as a function of groups, meaningfulness, and serial position within a list is shown in Figures 1 and 2. Also shown is the average output percentile, which is a measure of relative recall output position as a function of serial list input position (see Bjork & Whitten, 1974, for

TABLE 2

MEAN NUMBER OF WORDS RECALLED FROM PRIMARY AND SECONDARY MEMORY

Orienting group	Memory component: Primary	Secondary
Structural		
High M	1.67	1.42
Low M	1.58	1.50
Total	3.25	2.92
Semantic		
High M	1.92	5.92
Low M	1.67	3.42
Total	3.59	9.34

Note. M = meaningfulness.

details). As expected, a typical bow-shaped serial position function was obtained for the control group, with the recall of high meaningful items fairly consistently higher than low meaningful items across most positions. Consistent with Bjork and Whitten (1974), the order of recall mirrors the probability of recall; subjects recalled items from the ends of the list both sooner and with higher probability than items from the middle of the list.

The serial position functions for the two incidental learning groups shown in Figure 2 differ from those of the control group in several ways. First, there is not a strong primacy effect in probability of recall for either incidental group. This is consistent with the view that the primacy effect results from unequal processing across serial positions and, when equated as in the present case, the effect is removed. Second, while the output percentile functions again mirror the recall probability functions, the effect is very different. Subjects in both groups showed a strong tendency to recall the most recently presented items first and to work backward through the list. Moreover, meaningfulness has a fairly constant effect across most serial positions for the semantic group and has no effect across serial position for the structural group.

While both groups showed the same degree of recency as measured by recall probability in Figure 2, the functions fall to different asymptotic levels. Since output functions were similar for both groups, levels of processing, and not retrieval processes, must influence asymptotic performance levels in free recall. Either more information is transferred into long-term storage or more retrievable codes are present when items are encoded by a semantic rather than a structural orienting rule. Tulving and Patterson's (1968) method of distinguishing between primary and secondary memory components in free recall was employed to examine this issue. Primary memory recall was defined as recall of an item from one of the last six serial positions, instead of the last four serial positions (Tulving & Patterson, 1968). This provided a more conservative test of the effect of orienting task and meaningfulness on primary memory. The results of this analysis are shown in Table 2, where the number of words recalled has been parsed into primary and secondary memory components. The semantic group recalled more words from secondary memory than did the structural group, yet recall from primary memory was equal. An effect of meaningfulness was evident only in the secondary memory component of the semantic group.

The above results were supported by an analysis of variance. There was a significant interaction of Groups × Memory Component, $F(1, 22) = 19.31$, $MS_e = 2.87$, with no reliable difference between the groups on the primary memory component. There was also a significant three-way interaction of Groups × Memory Component × Meaningfulness, $F(1, 22) = 5.93$, $MS_e = 1.48$, which is attributable to the recall difference between high and low meaningful stimuli from secondary memory for the semantic group, as none of the other meaningfulness comparisons in Table 2 approached significance. Thus, the effect of meaningfulness, which was used to distinguish structural from semantic analyses, operates only on the secondary memory component—the same component that is influenced by levels of processing.

EXPERIMENT 2

A second experiment was conducted to determine if the results of the first experiment would be obtained in a recognition memory paradigm. A failure to find a difference between the semantic and structural groups in recognition would imply that, since only the retrieval condition varied between experiments, the recall difference in Experiment 1 reflects a difference in retrieval between the two orienting groups, rather than a difference in the amount of information stored. Alternatively, if the semantic group shows greater recognition than the structural group, the interpretation of the result would vary with the theory of recognition memory employed (see Watkins & Tulving, 1975), and the determination of whether levels of processing affects storage or retrieval may not be resolvable. Recognition, however, does permit measurement of the effect of levels of processing on memory sensitivity.

Predictions about the effect of meaningfulness upon recognition memory are complicated by inconsistent results in the literature (e.g., Martin & Melton, 1970; McNulty, 1965). The levels-of-processing assertion that high meaningful stimuli are processed to a deeper level than low meaningful stimuli (Craik & Lockhart, 1972) suggests that meaningfulness should have the same effect upon performance whether tested by recall or recognition. As such, meaningfulness may not be a stimulus factor that will differentiate recall and recognition memory.

Lastly, since the list items will be spread evenly over all positions on the recognition test, the item retention intervals will be more similar in the present experiment than they were in the former. As a result, recognition serial position effects should be attenuated.

Method

Subjects. The subjects were 30 Wesleyan undergraduates, 10 in each condition, who served as paid volunteers.

Procedure. Subjects in each incidental learning group were treated in the same fashion as those of Experiment 1 up to the time of memory testing. Subjects in the control group were also treated identically except that they were told that a recognition test would follow. The apparatus and all stimulus materials were the same.

After viewing 60 stimulus words, subjects were given a mimeographed list of 120 words. They were to examine each item in sequential order and assign a recognition confidence rating of 1 (presented–positive), 2 (presented–probable), 3 (presented–guess), 4 (not presented–guess), 5 (not presented–probable), or 6 (not presented–positive). The confidence ratings were explained to all subjects before the recognition test. Subjects were given as much time as necessary to complete the rating task, although none took more than 5 or 6 min.

A recognition list of 120 words was constructed from the two input lists of 60 words each. Each input list was shown to half of the subjects in each group. The recognition list thus contained 60 words from the list that the subject had just seen and 60 distractor words from an alternative list that had not been presented. The serial position of the words from the input lists was randomized over the recognition list such that the mean input serial positions for items in the beginning, middle, and end of the recognition test were very close to the chance expected value of 30.50.

Results and Discussion

The decision time and recognition data are shown in Table 3. It may readily be seen that performance was influenced in the same fashion in both experiments. Decision time for the semantic group was virtually identical with that of Experiment 1; decision time for the structural group was faster than previously observed due to the presence of two exceptionally fast subjects. An analysis of variance showed only a significant effect of meaningfulness upon decision time, $F(1, 18) = 6.88$, $MS_e = .02$, suggesting again that high meaningful words are encoded faster than low meaningful items. Input serial position had little effect on decision time.

Hit and false-alarm rates are shown in columns 2 and 3. The semantic group had a high hit rate and a low false-alarm rate, while the reverse was true for the structural group. Meaningfulness had the same effect as that observed in recall; the control and semantic groups yielded a higher hit rate and lower false-alarm rate for high meaningful than low meaningful items, while the structural group had comparable values for both levels of meaningfulness.

Analysis of the recognition data was performed using signal detection theory.

TABLE 3

MEAN DECISION TIME AND RECOGNITION DATA AS A FUNCTION OF EXPERIMENTAL CONDITION AND STIMULUS MEANINGFULNESS IN EXPERIMENT 2

Condition	Decision time (sec)	Hit rate (%)	False-alarm rate (%)	d'	Confidence ratings: Targets	Confidence ratings: Distractors
Control						
High M	—	89.66	7.01	3.00	1.54	5.25
Low M	—	83.73	11.63	2.37	1.81	4.89
Total	—	86.73	9.29	2.59	1.65	5.07
Structural						
High M	1.47	68.11	21.04	1.35	2.71	4.60
Low M	1.51	72.38	26.97	1.26	2.66	4.35
Total	1.49	70.19	24.02	1.28	2.69	4.47
Semantic						
High M	1.38	96.66	1.00	4.13	1.17	5.87
Low M	1.58	87.96	2.33	3.37	1.60	5.73
Total	1.48	92.32	1.66	3.66	1.38	5.80

Note. The data are means over individual subjects in each condition. M = meaningfulness.

The measure of memory sensitivity, d', was obtained for each subject under each meaningfulness condition from Elliott (1964) and averaged over groups, with the results shown in column 4 of Table 3. The d' value for the semantic group with high meaningful stimuli is actually a conservative estimate. Four of the subjects in this condition performed perfectly, thereby making it necessary to use the .99 and .01 hit and false-alarm rates, respectively, to estimate their sensitivity. An analysis of variance on the d' values found the difference between the semantic and structural groups, $F(1, 18) = 181.65$, $MS_e = .33$, and the Groups × Meaningfulness interaction, $F(1, 18) = 8.03$, $MS_e = .14$, to be significant. The difference in meaningfulness d' values was significant for the semantic group but not for the structural group. As before, the control group showed the same effect as the semantic group, with a reliable difference in d' for high and low meaningful stimuli, $t(9) = 2.58$.

Very clearly, both recognition and recall show the same effect of levels of processing and stimulus meaningfulness. While the difference between groups in recognition could reflect either storage or retrieval differences, the d' analysis indicates that levels of processing has a major effect on memory sensitivity. A direct comparison of the response criteria is not possible, as the range of possible criterion values varies directly with d' (Banks, 1970). However, the present results cannot be obtained by assuming equal sensitivity and different decision criteria between groups.

Mean confidence ratings are shown in columns 5 and 6 of Table 3 for the target and distractor words, respectively. The results show confidence ratings to vary in the same fashion as recognition performance. An examination of confidence ratings for targets across input serial positions indicated that the ratings were relatively stable across all serial positions and the effect of meaningfulness was present over all portions of the input list for the control and semantic groups.

CONCLUSIONS

The present research distinguished between structural and semantic levels of processing by the use of meaningfulness. This variable had a significant positive effect on both recall and recognition for the semantic group but no effect on the structural group. These results suggest that an account of performance in terms of only the perceptual orienting task is insufficient. A similar conclusion was reached by Craik and Tulving (1975), who argued that retention is a function of both the depth of processing and the spread of encoding. Use of different orienting tasks in the present experiments may have varied the depth of processing, while manipulating meaningfulness may have varied the spread. During encoding, the semantic group was forced to consider each stimulus in relation to other items in semantic memory. Greater contact or activation may have occurred for high than low meaningful items. No comparisons in semantic memory were required, however, for the structural group. In the present scheme, subjects may be able to determine the depth of processing by varying their perceptual analysis; spread of encoding, however, would seem to be determined by the structure of semantic memory. If depth of processing and spread of encoding affect memory performance in terms of the availability of retrieval cues, the present results could reflect differences in the availability of effective cues over the experimental conditions. Memory may still be viewed as a by-product of perceptual analysis, but the analysis determines cue effectiveness, and retrieval cues mediate performance.

REFERENCES

Banks, W. P. Signal detection theory and human memory. *Psychological Bulletin*, 1970, *74*, 81–99.

Bjork, R. A., & Whitten, W. B. Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, 1974, *6*, 173–189.

Craik, F. I. M. A "levels of analysis" view of memory. In P. Pliner, L. Krames, & T. M. Alloway (Eds.), *Communication and affect: Language and thought.* New York: Academic Press, 1973.

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.

Elias, C. S., & Perfetti, C. A. Encoding task and recognition memory: The importance of semantic encoding. *Journal of Experimental Psychology*, 1973, *99*, 151–156.

Elliott, P. B. Tables of *d'*. In J. A. Swets (Ed.), *Signal detection and recognition by human observers.* New York: Wiley, 1964.

Fischler, I., Rundus, D., & Atkinson, R. C. Effects of overt rehearsal procedures on free recall. *Psychonomic Science*, 1970, *19*, 249–250.

Gibson, J. J. *The senses considered as perceptual systems.* Boston: Houghton Mifflin, 1966.

Glanzer, M., & Cunitz, A. R. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, *5*, 351–360.

Hyde, T. S., & Jenkins, J. J. Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, 1969, *82*, 472–481.

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.: Erlbaum, 1974.

Martin, E., & Melton, A. W. Meaningfulness and trigram recognition. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 126–135.

McNulty, J. A. Short-term retention as a function of method of measurement, recording time, and meaningfulness of the material. *Canadian Journal of Psychology*, 1965, *19*, 188–195.

Modigliani, V., & Seamon, J. G. Transfer of information from short- to long-term memory. *Journal of Experimental Psychology*, 1974, *102*, 768–772.

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

Tulving, E., & Patterson, R. D. Functional units and retrieval processes in free recall. *Journal of Experimental Psychology*, 1968, *77*, 239–248.

Watkins, M. J., & Tulving, E. Episodic memory: When recognition fails. *Journal of Experimental Psychology: General*, 1975, *104*, 5–29.

(Received February 11, 1976; revision received April 16, 1976)