

Michigan State University) Ann Arbor, University Microfilms, 1969, No. 69-16.

WINOKUR, G. Conditioning versus in the galvanic skin response. *Comparative and Physiological Psychology*, 1969, 1, 169-170.

WATSON, P. E. Interpretation of the positive and negative of the skin potential response. *Record*, 1966, 16, 357-360.

MILLER, R. D., & STRAHAN, R. F. Ties of skin conductance and skin psychophysiology, 1968, 5, 253-268.

G., & JOHNSON, L. C. A re-analysis of conditioning. *Psychophysiology*, 1965, 1, 1-10.

WATSON, P. E., & EBEL, H. C. Three distinct of the classically conditioned GSR subjects. *Journal of Experimental Psychology*, 1967, 73, 247-256.

KOTSES, H., & BEVER, J. Autonomic orienting and defensive reflexes. *Experimental Psychology*, 1969, 80, 1-10.

Pavlovian conditioning and its proper procedures. *Psychological Review*, 1967, 74, 1-10.

WATSON, P. E., MILLER, L. H., & COHEN, S. I. Among age and sex groups in electroconditioning. *Psychophysiology*, 1968, 5, 1-10.

A., STERN, J. A., WINKOUR, G., & WATSON, P. E. An analysis of GSR conditioning. *Psychological Review*, 1961, 68, 60-67.

WATANABE, T., & NIIMI, Y. Orientants of the skin potential reflex in conditioning. *Japanese Journal of Psychology*, 1969, 40, 98-102.

Received March 29, 1973)

## LEVELS OF PROCESSING IN WORD RECOGNITION AND SUBSEQUENT FREE RECALL<sup>1</sup>

JOHN M. GARDINER<sup>2</sup>

*Birkbeck College, University of London, London, England*

The present experiment explored a situation in which Ss were unexpectedly required to recall the target words from a perceptual decision-making task. The targets were defined with respect either to their phonemic or semantic attributes, and Ss held these attributes in "working memory" for varying time intervals prior to target presentation. Semantically defined targets were better recalled subsequently than were phonemically defined targets, although the latter gave rise to longer decision latencies in the initial task. Also, subsequent target recall was not affected by the length of time the target-defining attributes had been held in working memory. These results were discussed within the context of Craik and Lockhart's "levels-of-processing" approach.

Craik and Lockhart (1972) described a framework for memory research in which the memory trace is viewed essentially as the by-product of perceptual analyses. Central to their argument is the notion that the stability of the memory trace is a positive function of the type and depth of processing involved in the encoding of perceptual events. Greater depth of processing is defined in terms of the degree of cognitive involvement in carrying out stimulus analyses. Memory is assumed to be tied to a continuum of levels of processing which range, for example, from sensory analyses to the activation of associative semantic attributes.

As Craik and Lockhart (1972) suggested, this formulation implies that research should be directed toward determining the memorial consequences of various types of perceptual operations. Craik (in press, Experiments IV and V), for example, reported 2 studies in which S was given an initial perceptual decision-making task followed by an unexpected memory test. The purpose of the initial decision-making task was to lead S to

encode different words at different levels of analysis. For each of a series of words presented briefly in a tachistoscope, S was asked a question such as (a) Is there a word present? (b) Does the word rhyme with \_\_\_\_? (c) Is the word a member of the following category? or (d) Does the word fit into the following sentence? In general, Craik's results showed that "deeper" decisions about words gave rise both to longer decision latencies in the initial task, and to better memory performance subsequently.

These results are in good agreement with the levels-of-processing view. However, it could be argued that the differences obtained in memory performance were due to the corresponding differences in initial processing time, rather than due to increasing "depth" of analysis. Although words processed at "sentence level" were better remembered than words processed at "rhyme level," for example, S also took longer in making sentence level decisions.

The general aim of the present experiment was to obtain further evidence in support of the levels-of-processing approach. Again, the strategy was to examine the effects of depth of analysis in a perceptual decision-making task upon subsequent memory for "target" words. The present study was also designed to provide more information on the effects of initial processing time in such a situation. In this context, it is important to note that processing time refers, poten-

<sup>1</sup>This research was supported by a Science Research Council Postgraduate Award. The author is grateful to Fergus I. M. Craik and Peter Herriot for their helpful comments on an earlier draft of this article.

<sup>2</sup>Requests for reprints should be sent to John M. Gardiner, who is now at the Department of Social Science and Humanities, The City University, St. John Street, London EC1V 4PB, England.

tially, both to the time *S* takes to process the target word as such and to the time *S* spends in processing those attributes that serve to define the target. The present paradigm permits the effects of these 2 processing times to be evaluated separately.

In essence, *Ss* were presented with a series of word sets, each containing a unique target word. Prior to the presentation of each word set, *S* was given the "defining attributes" of the target word. For half of the *Ss*, the target words were defined by their containing either 1 or 2 phonemes. For the remaining *Ss*, the targets were defined as members of either a single category or a nested category. The target word POLAND, for example, might be defined as containing the phonemes /l/ or the phonemes /l/ and /n/, on the one hand, or as a country or a European country, on the other hand. Thus, one group of *Ss* was required to identify targets at a "structural" level of analysis while the other group of *Ss* had to identify targets at a deeper, semantic level. The *Ss*'s decision latencies in response to targets provided a measure of target-word processing time. In order to examine the effects of processing time for target-defining attributes, the number of nontarget words preceding the target in each set varied across trials. Thus, the length of time the target-defining attributes were held in "working memory" was indexed by the position of the target in any given word set. Following this perceptual decision-making task, *Ss* were unexpectedly asked to write down all the target words they could remember.

The experiment had, therefore, the following 2 main objectives: (a) to compare recall performance following either phonemic or semantic processing of target words, and (b) to examine the relationship between initial processing times and subsequent recall.

#### METHOD

*Subjects.* The *Ss* were 32 volunteers, mostly undergraduate students, who were either paid for their services or were fulfilling a course requirement. The *Ss* were tested individually.

*Design.* The experiment was conducted using a mixed design. Level of processing (phonemic or semantic) was the between-*Ss* variable, called here for convenience "attribute type." Sixteen *Ss* were assigned at random to each between-*Ss* condition. Within *Ss*, the design comprised a factorial combination involving 2 levels of "attribute number" and 4 levels of "attribute lag." Attribute number refers to target-word definition by either 1 or 2 phonemes, on the one hand, or by a single or nested category name, on the other hand. Attribute lag refers to the position of the target within each word set.

The *Ss* were presented with 38 word sets in the target-recognition session. Prior to the presentation of the first word in each set, the target for that set was defined either as a member of a single or nested category for *Ss* in the semantic condition, or by the word containing either 1 or 2 given phonemes for *Ss* in the phonemic condition. The target word was always the final word in the set. On any given trial, it was preceded by 1-6 nontarget words. The critical positions of the target for scoring purposes were 3, 4, 5, and 6. Positions 2 and 7 were included to reduce target predictability by position; 6 of the 38 trials were of this kind. The remaining 32 target words appeared equally often at each of the 4 critical positions and were defined equally often by either 1 or 2 attributes in each between-*Ss* condition. For presentation purposes, the total number of word sets was divided into 4 blocks which were counterbalanced with regard to presentation order and the level of attribute number. In effect, all *Ss* received the same target words, and no target word was differentially favored with respect to order of presentation or experimental condition.

Approximately 2 min. after the final trial in the recognition sessions, *Ss* were unexpectedly asked to write down, in any order, all of the target words they could remember.

*Materials and procedure.* Of the 32 critical target words, 27 were selected from the Battig and Montague (1969) category norms and the remaining 5 were chosen by *E*. Those drawn from the published norms had a mean frequency of occurrence of 16%. This relatively low value was chosen in order to minimize *Ss*'s chances of guessing the target word when given category names as the target-defining attributes. The frequency count for nontarget words ranged between 5-20/million (Thorndike & Lorge, 1944). All words, both targets and nontargets, were 2 syllables in length.

Two sets of presentation lists were constructed for each of the 4 levels of attribute lag. The difference between the 2 sets lay in the nontarget words in each; specifically, where target words were defined by 2 attributes, the preceding nontarget words always included 1-3 words that contained one or the other of the 2 defining phonemes and belonged to the single but not to the nested category. This procedure was designed to ensure that *Ss* had to utilize both given attributes when targets were so defined.

The word lists were typed in for presentation via an IBM *c* and a closed-circuit television system. Laboratory Timer and Signal *Ss* both trigger the presentation or also to start, simultaneously, a V stopclock. The Birkbeck Laboratory Signal Source was operated contingent upon *Ss*'s having responded recently presented. In practice, *V* appeared at about a 2-sec. rate.

The *Ss*'s vocal response, amplified microphone, stopped the Vennel was instructed to respond yes or as rapidly but as accurately as addition, *Ss* were instructed that after they were to report the target decision latency measure, however from the onset of the target word.

The *Ss* were initially familiar paratus and procedure. They instructed as to the nature of materials. Prior to the experiment received 6 practice trials in which appeared once at each possible. Each trial began with *E* reading defining attributes for the word; example, the case where the POLAND. If the target was defined *E* said either "the target contains" or "the target is a country," for semantic conditions, respectively target was defined by 2 attributes "the target contains the sounds" "the target is a European country" of other target words and their 2-attribute definitions were: LEC animal; /p/ and /d/, or a wild /b/, or a city; /b/ and /n/, or a cap /k/, or a disease; /k/ and /t/, TEMPLE; /m/, or a building; /m/ gious building. Following target word in the set was presented, or Examples of 2 presentation sets METEOR, ATTIC, POLAND, and ATTIC, POLAND. The latter exact distractor which contains 1 of the phonemes and belongs to the nested category; sets including distractors were presented when defined by 2 attributes. Each after *S* had correctly identified target; the next trial began immediately. The *Ss* were told that the purpose was to investigate the effect of target words by 1 or 2 "values" on the speed with which they could target. No *S* reported subsequently anticipated having to recall the target. Unknown to *Ss*, only decision latencies to target words was recorded. In the target-recognition session "corrected" orally by *E* at the occurrence.

periment was conducted using level of processing (phonemic or between-Ss variable, called here attribute type." Sixteen Ss were  $n$  to each between-Ss condition. sign comprised a factorial combination of 2 levels of "attribute number" (attribute lag). Attribute number was defined by either 1 or 2 on one hand, or by a single or two, on the other hand. Attribute position of the target within each

presented with 38 word sets in the session. Prior to the presentation in each set, the target for  $S$  either as a member of a single set for Ss in the semantic condition containing either 1 or 2 given in the phonemic condition. The ways the final word in the set, it was preceded by 1-6 non-critical positions of the target were 3, 4, 5, and 6. Positions led to reduce target predictability the 38 trials were of this kind. target words appeared equally in 4 critical positions and were defined by either 1 or 2 attributes in condition. For presentation number of word sets was divided which were counterbalanced with presentation order and the level of effectiveness, all Ss received the same target word was differentially to order of presentation or ex-

min. after the final trial in the Ss were unexpectedly asked in order, all of the target words

edure. Of the 32 critical target words from the Battig and Monty norms and the remaining 5 Those drawn from the pub-mean frequency of occurrence which low value was chosen in Ss' chances of guessing the given category names as the attributes. The frequency count ranged between 5-20/million (e, 1944). All words, both sets, were 2 syllables in length. presentation lists were constructed of levels of attribute lag. The dif-2 sets lay in the nontarget conditionally, where target words attributes, the preceding non-included 1-3 words that concern of the 2 defining phonemes single but not to the nested edure was designed to ensure e both given attributes when d.

The word lists were typed in uppercase letters for presentation via an IBM electric typewriter and a closed-circuit television system. A Birkbeck Laboratory Timer and Signal Source was used to both trigger the presentation of each word and also to start, simultaneously, a Venner millisecond stopclock. The Birkbeck Laboratory Timer and Signal Source was operated manually by  $E$ , contingent upon  $S$ 's having responded to the word currently presented. In practice, words within a set appeared at about a 2-sec. rate.

The  $S$ 's vocal response, amplified from a throat microphone, stopped the Venner timer. Each  $S$  was instructed to respond *yes* or *no* to each word as rapidly but as accurately as he could. In addition, Ss were instructed that after each *yes* response they were to report the target word aloud. The decision latency measure, however, was of course from the onset of the target word to  $S$ 's *yes* response.

The Ss were initially familiarized with the apparatus and procedure. They were also fully instructed as to the nature of the stimulus materials. Prior to the experimental trials, each  $S$  received 6 practice trials in which the target word appeared once at each possible target position. Each trial began with  $E$  reading aloud the target-defining attributes for the word set. Consider, for example, the case where the target word was POLAND. If the target was defined by one attribute,  $E$  said either "the target contains the sound /l/," or "the target is a country," for Ss in the phonemic and semantic conditions, respectively. Where the target was defined by 2 attributes,  $E$  said either "the target contains the sounds /l/ and /n/," or "the target is a European country." Examples of other target words and their corresponding 1- or 2-attribute definitions were: LEOPARD: /p/, or an animal; /p/ and /d/, or a wild animal; DUBLIN: /b/, or a city; /b/ and /n/, or a capital city; RICKETS: /k/, or a disease; /k/ and /t/, or a bone disease; TEMPLE: /m/, or a building; /m/ and /p/, or a religious building. Following target definition, each word in the set was presented, one word at a time. Examples of 2 presentation sets were: FATHOM, METEOR, ATTIC, POLAND, and FATHOM, JAPAN, ATTIC, POLAND. The latter example includes one distractor which contains 1 of the 2 target-defining phonemes and belongs to the main but not to the nested category: sets including one or more such distractors were presented when the target was defined by 2 attributes. Each trial terminated after  $S$  had correctly identified and named the target; the next trial began immediately.

The Ss were told that the point of the experiment was to investigate the effects of defining target words by 1 or 2 "values along a dimension" on the speed with which they could identify the target. No  $S$  reported subsequently that he had anticipated having to recall the target words.

Unknown to Ss, only decision latency in response to target words was recorded. Errors made by  $S$  in the target-recognition session were noted and "corrected" orally by  $E$  at the time of their occurrence.

## RESULTS

For both the decision latency data and the recall data, Target Positions 3, 4, 5, and 6 (attribute lag) were collapsed into 2 levels of 3 plus 4 (short lag) and 5 plus 6 (long lag) in order to obtain more reliable estimates. The  $E$  errors and apparatus failures accounted for less than 2% of all responses in the target-recognition session;  $S$  errors accounted for 5% of responses in the semantic condition and 6% in the phonemic condition. Error data were excluded from all analyses. The results are shown in Figure 1.

Consideration is given first to the latency data. The basic datum was median reaction time for correct responses from the 8 trials each  $S$  received in the  $2 \times 2 \times 2$  combination representing attribute type (phonemic or semantic), attribute number (1 or 2), and attribute lag (short or long). The mean median scores are shown in the left-hand panel of the figure; each point is based on approximately 120 observations. These data were subjected to an analysis of variance. The main effects due to attribute type,  $F(1, 30) = 14.9$ , attribute number,  $F(1, 30) = 43.5$ , and attribute lag,  $F(1, 30) = 8.8$ , were all significant at or beyond  $p < .01$ . The only significant interactions were the following: Attribute Type  $\times$  Number,  $F(1, 30) = 12.3$ ,  $p < .01$ , and Attribute Type  $\times$  Lag,  $F(1, 30) = 4.4$ ,  $p < .05$ .

No detailed interpretation of these results will be attempted. From the point of view of the subsequent argument, the most relevant feature in these data is that Ss took considerably longer to process target words at the phonemic level than they did when targets were defined semantically. This result might seem surprising in view of Shulman's (1970) finding that, in a probe-recognition task, homonyms were recognized more rapidly than synonyms. However, it should be noted that whereas homonyms can be identified on a "holistic" basis, target words in the present paradigm were identified on the basis of individual, constituent phonemes.

The right-hand panel in Figure 1 shows the percentage of target words which Ss

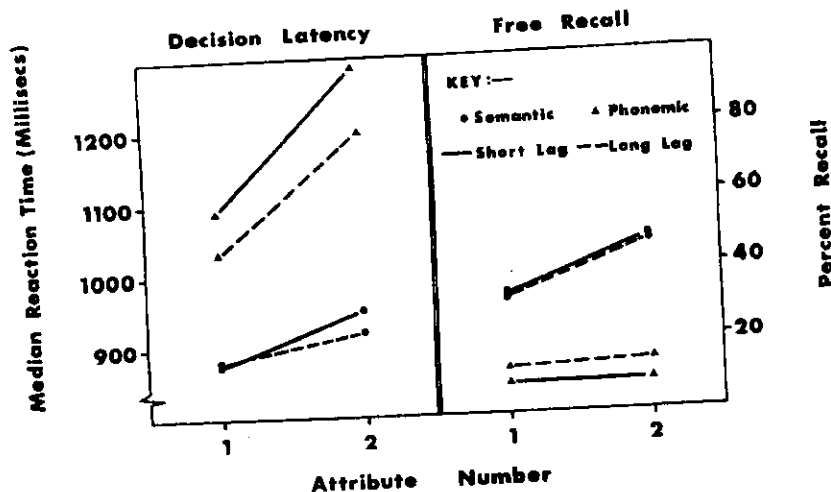


FIGURE 1. Decision latency and subsequent recall as a function to attribute type, number, and lag.

recalled in the final test, with the lag variable again collapsed into short and long lag. In order to obtain "workable" numbers for the purposes of analysis, the recall data supplied by each consecutively tested pair of *Ss* were combined to form 8 macro-*Ss* in each between-*Ss* condition. An analysis of variance carried out on the macro-*S* recall scores confirmed that the main effect due to attribute type was highly significant,  $F(1, 14) = 73.2, p < .001$ . The main effect due to attribute number was also significant,  $F(1, 14) = 7.2, p < .05$ , but the main effect due to attribute lag was not ( $F < 1$ ). The only significant interaction was Attribute Type  $\times$  Number,  $F(1, 14) = 5.7, p < .05$ .

Not surprisingly, these results confirm that target words were recalled much better following semantic processing, as compared with phonemic processing, in the initial session. Recall of targets following phonemic processing is virtually at floor level. The results also indicate that the number of phonemes involved in processing the target had no effect upon recall performance. This finding contrasts with the comparable result in the semantic condition, where recall is better following target definition by a nested category than by a single category name.

Finally, it is apparent that the length of time the target-defining attributes were held in working memory (short or long attribute lag) did not affect subsequent recall of the target word. Neither the main effect due to attribute lag nor any interaction term involving lag approached significance.

#### DISCUSSION

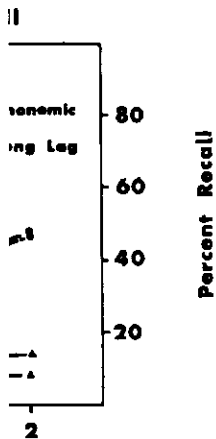
These results have several implications for the levels-of-processing approach discussed earlier. First, it is clear that processing target words at a semantic level led to good subsequent recall of the targets, whereas targets processed at a phonemic level were very poorly recalled. Similar differences in memory performance following semantic as opposed to structural processing have been reported by Craik (1973), Hyde and Jenkins (1969), and Schulman (1970). Such findings are in good agreement with the levels-of-processing view.

Although further analysis at a semantic level resulted in a corresponding improvement in recall (nested vs. single category names), additional analysis at a phonemic level did not benefit recall. Thus, as Craik (1973) pointed out, the persistence of the memory trace should not be conceived to be simply a function of the number of analyses performed on the stimulus, but rather, a function of the degree of meaningfulness extracted from it.

The markedly poor recall that had been processed at was associated with much latencies for these targets task. Clearly, the pattern of in the recall test cannot be a responding differences in the following target-word recog required to report the tar This procedure ensured that were processed to at least a The poor recall of phonemical must reflect the manner in were processed, rather than that the target word as a perceived.

It was also found that the the target-defining attribute working memory did not a recall of the target words relates well to Craik and L suggestion that maintaining a "fixed" level of processing rarely lead to any improvement trace. More generally, the monstrate that duration at nearly so crucial for registra as are the levels of analysis evidence was found to sup that processing time, whether or for target-defining attribut factor for subsequent reca The results imply that proc only be important insofar as i or deeper levels of analysis to

In summary, the present a situation in which *S* was required to recall the target identified in a perceptual task. The main findings were the target words was better for as opposed to phonemic analy Second, whereas further analy



ion to attribute type,

apparent that the length of time for holding target-defining attributes were held in working memory (short or long lag) did not affect subsequent recall of the target word. Neither the length of time to attribute lag nor any lag involving lag approached

## DISCUSSION

have several implications for word processing approach discussed here. It is clear that processing target words at a phonemic level led to good subsequent recall of the targets, whereas targets processed at a semantic level were very poorly recalled. Similar differences in memory performance between semantic as opposed to phonemic processing have been reported by Craik and Jenkins (1969), and by Craik and Tulving (1975). Such findings are in good agreement with the levels-of-processing view. Craik's (1973) further analysis at a semantic level showed a corresponding improvement in recall (e.g., single category names) as opposed to single category names at a phonemic level did not. Thus, as Craik (1973) suggested, the persistence of the memory trace may be conceived to be simply a function of the number of analyses performed on the stimulus, but rather, a function of meaningfulness extracted

The markedly poor recall of target words that had been processed at a phonemic level was associated with much longer decision latencies for these targets in the orienting task. Clearly, the pattern of results obtained in the recall test cannot be attributed to corresponding differences in the processing time for target words. It should be noted that following target-word recognition, Ss were required to report the target word aloud. This procedure ensured that all target words were processed to at least a "nominal" level. The poor recall of phonemically defined targets must reflect the manner in which the words were processed, rather than the possibility that the target word as a whole was never perceived.

It was also found that the length of time the target-defining attributes were held in working memory did not affect subsequent recall of the target words. This finding relates well to Craik and Lockhart's (1972) suggestion that maintaining information at a "fixed" level of processing does not necessarily lead to any improvement in the memory trace. More generally, the present data demonstrate that duration at encoding is not nearly so crucial for registration in memory as are the levels of analysis involved. No evidence was found to support the notion that processing time, whether for target words or for target-defining attributes, was a critical factor for subsequent recall performance. The results imply that processing time may only be important insofar as it permits further or deeper levels of analysis to be carried out.

In summary, the present study explored a situation in which S was unexpectedly required to recall the target words he had identified in a perceptual decision-making task. The main findings were, first, recall of the target words was better following semantic as opposed to phonemic analysis of the target. Second, whereas further analyses at a semantic

level led to a further improvement in recall, additional analyses at a phonemic level did not benefit recall. Third, neither processing time for target words, nor the length of time for which the target-defining attributes were held in working memory, appeared to be critical factors in determining subsequent recall performance. These results provide further support for the levels-of-processing view (Craik & Lockhart, 1972). It was concluded that the memorial consequences of the operations carried out in the perceptual decision-making task were attributable solely to the level of analysis these operations involved.

## REFERENCES

- BATTIG, W. F., & MONTAGUE, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 1969, 80(3, Pt. 2).
- 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.
- 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.
- SCHULMAN, A. I. Recognition memory for targets from a scanned word list. *British Journal of Psychology*, 1970, 62, 335-346.
- SHULMAN, H. G. Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 499-508.
- THORNDIKE, E. L., & LORGE, I. *The teacher's wordbook of 30,000 words*. New York: Columbia University, Teachers College, Bureau of Publications, 1944.

(Received April 9, 1973)