INSTITUTE FOR THE

STUDY OF

INTELLECTUAL

BEHAVIOR

Effects of Different Kinds of Semantic Processing on Memory for Words

Janet L. Packman and William F. Battig

University of Colorado

Program* on Cognitive Factors in Human Learning and Memory Report No. 77

Institute for the Study of Intellectual Behavior

University of Colorado

Boulder, Colorado 80309

December, 1977

ABSTRACT

Recall and recognition memory were shown to be significantly better for words rated on pleasantness than on any of the other six semantic dimensions (concreteness, imagery, categorizability, meaningfulness, familiarity, and number of attributes) recently used for scaling of 2854 words by Toglia et al. (1978). Pleasantness ratings are also relatively uncorrelated with ratings on these other six dimensions, and the pattern of memory differences between these seven dimensions corresponds closely to differences in dimensional distinctiveness, as indexed by the average correlation of each dimension with the other six dimensions as reported by Toglia et al. (1978). Word subsets with high and low mean ratings on all seven dimensions showed comparable dimensional differences in memory, but high words were both recalled and recognized better than were low words.

*Supported in part by NSF Grant BNS 72-02084

Several recent experiments have demonstrated quite convincingly that free recall and/or recognition performance for individual words is much improved following allegedly deeper semantic processing, as compared with nonsemantic processing conditions (e.g., Craik & Tulving, 1975; Jenkins, 1974). Considerably less attention, however, has been directed toward the examination of possible differences between various types of semantic processing with respect to subsequent memory performance.

While there exists some recent evidence for memory differences between different semantic processing tasks or dimensions, such comparisons have shown little consistency across experiments and have typically represented only secondary concerns within experiments done primarily for other reasons. Nonetheless, there have been several demonstrations of superior memory of words processed for pleasantness or other affective properties, as compared with a variety of other semantic processing dimensions. More specifically, pleasantness has been shown to lead to better word memory than processing for frequency (Postman & Kruesi, 1977), imagery (Paivio, 1975), concreteness (Thieman, Note 2) and for categorizability and number of attributes as well as for imagery and concreteness (Battig & Weiss, Note 1). Klein and Saltz (1976), however, reported no consistent superiority for pleasant-unpleasant ratings as compared either with the closely related happy-sad dimension or the less related fast-slow and big-little dimensions. Moreover, Shaughnessy (1976, Exp. 3) found no significant differences between pleasantness, imagery, and frequency processing conditions. Thus even the memory superiority for pleasantness over other types of semantic processing has not been consistently demonstrated.

The present experiment was designed to provide a more direct and comprehensive comparison of free recall and recognition memory following processing on the seven semantic dimensions of concreteness (CON), imagery (IMG), categorizability (CAT), meaningfulness (MNG), familiarity (FAM), number of attributes (NOA), and pleasantness (PLS). These are the seven dimensions on which 2854 different words were scaled recently in the Colorado laboratories (Toglia, Battig, Barrow, Cartwright, Posnansky, Pellegrino, Moore, & Camilli, 1978). To avoid possible interactive effects produced through the concurrent use of multiple semantic dimensions, each subject rated all words for only one of these seven dimensions.

The Toglia et al. (1978) norms were also used to select equal numbers of words representing two different high and low levels of scale values across all seven of these dimensions, thereby permitting a direct evaluation of possible differences between different word classes in relative memory performance across the seven semantic dimensions. An even more important use of the Toglia et al. (1978) normative information involved the evaluation of memory performance as a function of the degree to which semantic ratings on a given dimension were correlated with ratings on the other six dimensions. As the results will show, there is a strong indication that memory varies directly with the extent to which the semantic processing dimension is unique or unrelated to other semantic dimensions, thereby extending the recently emphasized importance of distinctiveness of encodings of individual items for effective memory (see Cermak & Craik, in press) to apply also to distinctiveness of the processing dimensions used for these encodings.

METHOD

Subjects and Design

All 105 subjects were introductory psychology students fulfilling a course requirement option through their participation, each serving for a single session of about 40 minutes. Subjects were run in small groups of 2-18. Equal-sized groups of 15 subjects used each of the seven Toglia et al. (1978) dimensions for their word ratings. All of these dimensions, however, were used concurrently for different subjects within the same group session, through the usage of different instruction sheets.

Procedure

Each subject was first given the appropriate instruction sheet for the assigned one of the seven semantic dimensions, along with a response sheet for the actual word ratings. All instructions were identical to those used and reproduced by Toglia et al. (1978). The 50 words to be rated by all subjects were then presented by slide projector for 5 seconds each, each word being preceded by the 5-second presentation of a slide containing the number from 1-50 representing the position of the word in the presentation sequence as designated on the subject's rating sheet. These rating sheets contained 50 rows of numbers from 1 (very low) to 7 (very high), and subjects were instructed to circle the one of these seven numbers which represented their rating for the presented word.

After all words had been rated and the rating sheets turned in, all subjects performed for 5 minutes on a nonverbal filler task, which consisted of 16 sets of four similar rectangular arrays of six x and o symbols. For each of these 16 items, subjects were instructed to circle the one array which they judged to be most different from the other three.

Immediately after completion of the filler task, all subjects were given a surprise written free-recall test for the 50 words that they previously had rated, with a time limit of 5 minutes. Subjects were instructed to write any words remembered together on the same line of the recall sheet, as suggested by Buschke (1977). Recall was followed immediately by a recognition test in which 100 words were presented in alphabetical order including all 50 rated words intermixed with

50 distractors, each of which was semantically and/or structurally similar to one or more of the rated target words. For each word, subjects were asked to circle one of the following four symbols: Y (quite sure the word was rated), YG (unsure, but best guess is that the word was rated), NG (unsure, but best guess is that the word was not rated), or N (quite sure the word was not rated).

Materials

All subjects rated exactly the same 50 words in the same presentation order. Equal halves of these words represented relatively high and low rating values across all of the seven dimensions. This was accomplished by selecting 25 words from two of the eight homogeneous clusters in which the word ratings are presented by Toglia et al. (1978). More specifically, Cluster 8 consists of words with uniformly high ratings on all seven dimensions, whereas Cluster 4 contains words with average or below average ratings on each dimension. Thus the mean rating differences between the 25 high and 25 low words ranged from 2.42 for concreteness down to 0.51 for familiarity, with high-low differences for the other five dimensions (ordered from largest to smallest) being categorizability (2.05), pleasantness (1.85), imagery (1.71), number of attributes (0.74), and meaningfulness (0.56). These 25 high and 25 low words were unsystematically intermixed in the rating presentation order, such that no more than four words of either type were ever presented in succession.

RESULTS

Presented in Table 1 separately for each of the seven dimensions are mean proportions of rated words that were correctly recalled (first row), and correctly recognized with Y responses (second row), along with mean d' measures from a signal-detection analysis performed on the recognition data (third row). Also presented in the last row of Table 1 are the means of the six correlation coefficients for each dimension with the other six dimensions as reported by Toglia et al. (1978, Table 1), which will become important in interpreting the present results.

Insert Table 1 About Here

Memory Performance

Since the recall measure showed the most significant differences between the seven dimensions, $\underline{F}$ (6, 98) = 3.58, $\underline{MS_e}$ = 30.82, $\underline{p}$<.005, dimensions are ordered from low to high recall performance in Table 1. Duncan tests on the recall data showed that pleasantness was significantly superior to all other dimensions, but there were no significant differences between any of these six dimensions. Moreover, the superiority of pleasantness over the other six dimensions accounted for 75.5% of the total between-groups variance.

The recognition d' measure showed a closely comparable pattern of differences between the seven dimensions, with an overall $\underline{F}$ (6, 98) = 3.46, $\underline{MS_e}$ = 0.53, $\underline{p}$<.005, with pleasantness alone differing significantly from each of the other six dimensions by Duncan test. This difference accounted for an even larger percentage of the total between-groups variance (81.2%) than was found for the recall measure.

As for percentage of target words correctly recognized, Table 1 indicates the magnitude of differences between the seven dimensions to be closely comparable to the recall differences. However, this percent recognition measure also produced an error variance over twice as large as that for recall, yielding an overall $\underline{F}$ (6, 98) = 2.74, $\underline{MS_e}$ = 65.26, $\underline{p}$<.02. Moreover, Duncan tests indicated pleasantness not to be significantly superior to number of attributes, although pleasantness did differ significantly from each of the other five dimensions, and the difference of pleasantness from the other six dimensions still accounted for 70.3% of the total between-groups variance.

Thus all three memory measures in Table 1 showed a significant superiority for pleasantness over the other dimensions, although any differences between the other six dimensions were much smaller and nonsignificant.

To determine whether the foregoing memory differences might reflect greater recall organization of words rated for pleasantness than the other semantic dimensions, further analyses were performed on numbers of cases where two words were recalled together on the same line of the recall sheet, and on adjusted-ratio-of-clustering scores (e.g., Roenker, Thompson, & Brown, 1971) based on recall of pairs of high or of low words on adjacent lines of the recall sheet. Neither organization measure showed any evidence of differences between groups, or any significant overall organization. Groups also did not differ in recall intrusions.

Memory as Related to Dimensional Relatedness

The superiority of all memory measures for pleasantness over the other six dimensions, combined with the fact that pleasantness ratings show the smallest correlations with other semantic dimensions (Toglia et al., 1978), suggests that the extent to which a semantic dimension is interrelated with the other types of semantic ratings may be an important determinant of memory performance. In an attempt to evaluate this possibility, the correlations between mean word ratings on all possible pairs of these seven dimensions obtained by Toglia et al. (1978, Table 1) were used to estimate the level of interrelatedness between each one of the present dimensions and the remaining six dimensions.

These dimensional relatedness estimates were obtained by finding the mean value of the six correlation coefficients (r) obtained by Toglia et al. (1978)

between the specified dimension and each of the six other dimensions. The resulting mean correlation values are given in the last row of Table 1. Because of the questionable scale characteristics of the correlation coefficient itself, parallel comparisons were performed using both the z transformations of these individual _r_ values, and also the corresponding squared correlation coefficients which represent directly the proportion of variance accounted for by the correlation between the two dimensions. Since all three of these analyses yielded closely comparable results, only the results for _r_ will be reported here.

Correlations were computed between the dimensional relatedness values in the last row of Table 1, and each of the recall and recognition memory measures given also in Table 1, in an attempt to obtain a quantitative estimate of the extent to which memory performance differences between the seven dimensions are associated with differences in dimensional relatedness. These correlations were all very large and negative in value, being significant at $p<.01$ for both the recall (-.941) and recognition d' measures (-.879) and falling just below the .01 level for recognition percent correct (-.846), despite the fact that the seven dimensions yield only five degrees of freedom for these comparisons. Nonetheless, 71-89% of the total variance between the present memory measures across the seven dimensions can be accounted for by differences in the extent to which their ratings are correlated with ratings on the other semantic dimensions. Since much of this relationship reflects differences of pleasantness from the other six dimensions (which also accounts for 80.1% of the variance between the seven mean _r_ values in Table 1), it is noteworthy that substantial negative relationships hold even when pleasantness is excluded, all _r_ values being at least -.31 and as large as -.738 for recall.

Differences in Memory for High and Low Word Subsets

As described in the method section, equal halves of the present 50 words were selected to represent high and low scale values on all seven of the present dimensions, so that high and low words were evaluated as an additional factor in the percent recall and recognition analyses. Such high-low word comparisons were not possible with the recognition d' measure, because distractors could not be clearly specified with respect to the high and low word subsets.

Table 2 contains the mean percent recall and recognition measures separately for the high and low subsets for each of the seven dimensions. Overall recall

Insert Table 2 About Here

performance proved to be nearly twice as great for the high as for the low word subsets, F (1, 98) = 105.40, MS_e = 16.8, p<.0001. A closely comparable overall superiority was obtained in recognition for high over low words, F (1, 98) = 78.20, MS_e = 25.28, p<.0001. Thus overall recall and recognition memory were markedly and comparably superior for words rated higher on the present semantic dimensions.

Of primary concern for present purposes, however, is the extent to which the superiority of pleasantness over the other six dimensions was consistent across the high and low subsets. Pleasantness showed the highest recall and recognition performance both for high and low word subsets, and this pleasantness superiority accounted for at least 56% of the total variance between the seven dimensions for all except recognition of the high-valued words. In the latter case, the 96% pleasantness recognition performance was so close to maximum that a ceiling effect may well be responsible for the reduced superiority for pleasantness over the other dimensions. In any event, this pleasantness superiority accounted for only 29.3% of the total variance between dimensions in recognition of high words.

In line with the foregoing description, there was no significant interaction of high-low words with the seven dimensions for recall, F (6, 98) = 1.33, $p<.25$. This interaction was significant, however, for percent correct recognition, F (6, 98) = 2.75, MS_e = 25.28, $p<.02$. Further indicative of some discrepancy between the pattern of dimensional differences for recognition of high words as compared to other high and low recall and recognition measures is the fact that only for high-word recognition was there a nonsignificant correlation with the dimensional relatedness estimates given in the last row of Table 4.

Consequently, it appears that the present recall differences between the seven dimensions are quite consistent across the high and low word subsets, and that recognition of low words likewise yields comparable dimensional differences. That the somewhat different pattern of dimensional differences for recognition of high words may represent a ceiling effect, and may not reflect any fundamental differences between recall-recognition measures and/or high-low word types, is indicated by the results for a three-way analysis of variance including recall and recognition measures as a factor along with dimensions and high-low words. This showed no evidence for interactions of the seven dimensions with high-low word sets, or of these two variables with recall-recognition, all Fs $\leq$ 1.21, p >.25.

The interaction of the seven dimensions with recall-recognition, however, was significant, F (6, 98) = 2.87, MS_e = 33.79, p <.02. This latter interaction proved to be entirely attributable to the significantly smaller superiority shown in Table 1 for recognition over recall within the familiarity group (59.9) than for any of the other six groups, F (1, 98) = 9.19, p <.01, which accounted for over half (53.3%) of the total interaction variance. These recall-recognition differences, however, were not significantly different across the other six dimensions, F (5, 98) = 1.61, p >.10.

DISCUSSION

The major finding of this experiment is the significant superiority in both recall and recognition memory produced by pleasantness processing as compared with any of the other six semantic dimensions. Since pleasantness ratings also show substantially lower correlations with ratings on the six other dimensions (Toglia, et al., 1978), these results suggest the greater distinctiveness or uniqueness of pleasantness as contrasted with the other six dimensions to be at least partially responsible for the present memory differences. The importance of distinctiveness of the encoding(s) for effective memory is currently receiving heavy emphasis by levels-of-processing researchers (e.g., Cermak & Craik, in press; Fisher & Craik, 1977; Klein & Saltz, 1976).

The present results and proposed distinctiveness interpretation, however, represent an important extension beyond the distinctiveness of the particular encoding of an individual word or event, to encompass also the higher "level" of distinctiveness of the semantic processing task or dimension which is used for that encoding. More specifically, pleasantness appears to produce better memory because it is less related to or more distinctive from at least the other six semantic dimensions used in this experiment, which results in more unique and consequently more memorable encodings of the individual words. At a more general level, it would appear that selection of a dimension or attribute which is maximally different from other dimensions or attributes may provide an effective way of enhancing memorability.

Of the other demonstrations of superior memory following pleasantness encoding cited in the introduction, only Postman and Kruesi (1977) offer a real explanation for the superiority of pleasantness over their alternative semantic dimension of judged frequency. They characterize pleasantness as a subjective dimension, for which ratings of a particular word are based upon comparisons with previously rated words, with consequent added "displaced rehearsals" of these previously rated words. Frequency ratings, on the other hand, are presumed to represent more objective comparisons with a few specific prototypical words that are likely to be external to the list, so that frequency comparisons produce less rehearsal of previously rated words.

Most if not all of the present semantic dimensions appear more subjective than objective in the Postman-Kruesi sense, and it is difficult to understand how pleasantness ratings could involve more subjective comparisons and displaced rehearsal of previous words than ratings for all of the other six dimensions (concreteness, imagery, categorizability, meaningfulness, familiarity, and number of attributes). Of these six dimensions, familiarity (which is closest to Postman and Kruesi's frequency) and number of attributes would seem to be the closest to objective dimensions with ratings based on comparisons with external standards. Table 1, however, shows that familiarity ranks second to pleasantness in recall performance, while number of attributes is second best on both recognition measures. So although the Postman-Kruesi interpretation may well reflect an important factor in effective memory, it clearly cannot account for the present pattern of memory differences as well as does our proposed distinctiveness interpretation.

Also deserving of further discussion is the evidence in Table 2 showing markedly superior recognition as well as recall for the high over the low word subset. Mean Kucera-Francis (1967) frequency counts are over four times as large for the present high (85.40) as for the low word subset (20.24), F (1, 48) = 7.46, MS_e = 7113, p <.01. Consequently, the present recognition differences represent another exception to the general finding (e.g., Gregg, 1976) that recognition is better for low- than for high-frequency words, even though the present low words are probably higher in frequency than those typically employed in such comparisons. That the present high words are also high on all seven of the rated semantic dimensions may well have been sufficient to counteract any tendency toward superior recognition of low-frequency words, although such differences on other semantic dimensions also characterize other experiments showing a low-frequency recognition superiority. It should also be noted that the present results represent an exception to Gregg's (1976) conclusion that the recall superiority for high-frequency words is limited to homogeneous lists, and is not found in mixed high- and low-frequency lists like those of the present study.

Although there is no evidence here for overall differences between high and low words in recognition as compared with recall memory, the present processing dimension (familiarity) that is most closely related to frequency does show a significantly smaller increment from recall to recognition than any other dimension, as well as a smaller difference favoring high over low words on both recall and recognition than the other six dimensions. This suggestion that frequency-related processing may attenuate the magnitude of recall-recognition and high-low word differences deserves further investigation, but it may bear some relevance to the inconsistent pattern of high-low frequency differences in recall and recognition memory that has received considerable attention in recent years (e.g., Gregg, 1976).

REFERENCES

Buschke, H. Two-dimensional recall: Immediate identification of clusters in episodic and semantic memory. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 313-324.

Cermak, L. S., & Craik, F. I. M. (Eds.) Levels of processing and human memory Hillsdale, N. J.: Lawrence Erlbaum Associates, in press.

Craik, F. I. M., & Tulving, E. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 1975, 104, 268-294.

Fisher, R. P., & Craik, F. I. M. Interaction between encoding and retrieval operations in cued recall. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 701-711.

Gregg, V. Word frequency, recognition, and recall. In J. Brown (Ed.), Recall and recognition. New York: Wiley, 1976.

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

Klein, K., & Saltz, E. Specifying the mechanisms in a levels-of-processing approach to memory. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 671-679.

Kucera, H., & Francis, W. N. Computational analysis of present-day American English. Providence, Rhode Island: Brown University Press, 1967.

Paivio, A. Coding distinctions and repetition effects in memory. In G. H. Bower (Ed.), The psychology of learning and motivation. New York: Academic Press, 1975, Vol. 9.

Postman, L. & Kruesi, E. The influence of orienting tasks on the encoding and recall of words. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 353-369.

Roenker, D. L., Thompson, C. P., & Brown, S. C. A comparison of measures for the estimation of clustering in free recall. Psychological Bulletin, 1971, 76, 45-48.

Shaughnessy, J. J. Persistence of the spacing effect in free recall under varying incidental learning conditions. Memory and Cognition, 1976, 4, 369-377.

Toglia, M. P., Battig, W. F., Barrow, K., Cartwright, D. S., Posnansky, C. J., Pellegrino, J. W., Moore, T. J., & Camilli, G. A. Handbook of semantic word norms. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1978.

REFERENCE NOTES

1. Battig, W. F., & Weiss, K. R. Distinctiveness and elaboration of processing in recognition and recall memory. Paper presented at the Psychonomic Society meetings, 1977.

2. Thieman, T. J. What's so deep about deep processing: A critical analysis of the relation between incidental orienting tasks and retention. Unpublished manuscript, 1976.

Table 1

Mean Percent Words Recalled and Recognized, d' measures, and Correlations with other Dimensions for each of Seven Dimensions

Measure	MNG	CAT	IMG	NOA	CON	FAM	PLS
	Dimension						
Recall %	14.8	15.3	17.7	18.7	18.9	20.0	26.4
Recognition %	83.5	78.5	81.6	86.5	83.2	79.9	92.8
Recognition d'	2.42	2.01	2.38	2.48	2.38	2.34	3.17
Mean r	.595	.612	.638	.524	.519	.501	.288

Table 2

Mean Percent Words Recalled and Recognized for High and Low Word Subsets for Each of Seven Dimensions

	MNG	CAT	IMG	CON	NOA	FAM	PLS	Total
	Dimension							
Recall								
High	20.3	23.2	24.3	24.3	26.1	22.7	31.7	24.7
Low	9.3	7.5	11.2	13.6	11.2	17.3	21.1	13.0
Recognition								
High	90.9	85.1	92.8	88.0	94.0	82.1	96.0	89.8
Low	76.0	72.0	70.4	78.4	81.1	77.6	89.6	77.9