Journal of Experimental Psychology
1970, Vol. 84, No. 2, 268–272

IMAGERY AND ABSTRACTNESS IN SHORT-TERM MEMORY[1]

DELOS D. WICKENS[2] AND RANDALL W. ENGLE

Ohio State University

Two-hundred and fifty-six *S*s were run in a short-term memory study designed to investigate whether or not words are encoded according to their imagery-abstract characteristic. No clear evidence in support of encoding by means of this dimension was found. Performance on high-imagery words was superior to performance on abstract words on the second and third trial of the experiment, but not on the first, suggesting that the superior performance on high-imagery words is due to the fact that they produce less interitem interference than do abstract words.

Much evidence has been amassed during the past decade demonstrating proactive inhibition (PI) and the release from PI in short-term memory (STM). Keppel and Underwood (1962) have shown that with the retention interval held constant, performance declines from the first to the third or fourth trial. Wickens, Born, and Allen (1963), using a Peterson and Peterson (1959) paradigm, have shown that this PI buildup does not generalize to material of a different class. They presented *S*s with either CCCs or NNNs (digits) for three trials and then switched to the contrasting material on the fourth trial. They tested this against control groups who were presented either CCCs or NNNs for all four trials. When the item in the fourth trial was of the same class as the preceding trials, noticeable interference was obtained. However, when the fourth trial item was of a different class, there was no evidence of PI. These findings have since been extended to triads of words from different poles on the semantic differential, triads of words on different dimensions of semantic differential dimensions, sense impressions (Wickens, 1970), and taxonomic categories (Loess, 1967).

It appears, therefore, that the process of perceiving a verbal item involves encoding that item into one or more superordinate classifications or categories. If a series of items comes from the same category, the items will interfere with each other and depress retention performance. However, for words or items that are encoded differently (i.e., into a new category), interference is absent and retention performance increases (Wickens, 1970).

In recent years the concept of imagery has been revitalized and it has been abundantly demonstrated that words which are high in capacity to produce images are more effective stimuli and responses as learning materials than are words which are low in their capacity to produce images (Paivio, 1969). It has been suggested by Paivio, Yuille, and Rogers (1969) that the superior performance of *S*s on high-imagery words arises from the use of the image as an alternative coding system for memory representation, this alternate encoding assisting recall in one or another fashion. The STM technique described previously appears to be highly sensitive to differences in the way in which words are encoded, and the purpose of the present experiment was to use the STM technique to determine whether or not high-imagery words are encoded differently than are words low in imagery.

The method is a fairly simple one. Triads of words all homogeniously high in imagery or low in imagery are presented in the Peterson and Peterson (1959) paradigm with a constant recall interval for three trials and on the fourth trial the triad is shifted to words drawn from the opposite end of the imagery-abstract dimension. Control groups are given homogenious triads from one or

[1] This research was supported in part by Grants OEG-5-9-450276-0053 and 2 TOl MH 08526-05 from the United States Department of Health, Education, and Welfare.

[2] Requests for reprints should be sent to Delos D. Wickens, Department of Psychology, Ohio State University, Columbus, Ohio 43210.

the other end of the dimension throughout the four trials. If the performance of the experimental groups is higher than that of their appropriate controls on the fourth trial, then it can be assumed that *S*s are encoding differently on that trial than they had been during the first three trials.

Method

Subjects.—The *S*s were 260 male and female introductory psychology students at Ohio State University who chose to serve in the experiment to meet part of their course requirements. The data for four *S*s were discarded: for two *S*s because they misunderstood the instructions and for two *S*s because of equipment failure. The *S*s were assigned to one of the four groups in the order they came into the laboratory, the assignment rotating successively through the four conditions. There were 64 *S*s in each of the groups.

Materials.—The lists were chosen from the 925 words in the Paivio, Yuille, and Madigan study (1968). Because imagery (I) and concreteness (C) correlated so highly (.83) and the group of words that did differ on I and C were very unusual, I and C were allowed to covary. The words were paired, one high I-C and one low I-C, with meaningfulness (M), number of syllables, and Thorndike-Lorge frequency made as equal as possible for the two words. A pool of 48 high and 48 low words was chosen: the mean and standard deviation on the imagery scale for the high words were 6.36 and .23; for the low-imagery (high-abstract) words, the respective measures were 3.29 and .23. The mean and standard deviation on the covarying concrete scale were 6.69 and .41 for the high words; for the low words, the respective measures were 2.56 and .75. The high I-C words had a mean of 2.10 syllables and mean M scale of 5.92. The low I-C words had a mean of 2.27 syllables and a mean M rating of 5.85.

The words were placed into 16 high (Hi) and 16 low (Lo) I-C triads. Four lists were composed out of these 16 triads, each list using four triads. An equal number of *S*s was run on each list. The sole purpose of using four lists was to obtain as large a selection of words from each class as was deemed possible.

To form a triad the three words were required to meet the following criteria: (*a*) One of the words could not be an obvious association of one of its counterparts; (*b*) all three of the words could not be of the same frequency; (*c*) all three of the words could not have the same number of syllables; (*d*) one of the words could not have the same initial letter as one of its counterparts; and (*e*) one of the words could not be a form of, or a root of, one of its counterparts.

Apparatus.—The equipment used in the experiment consisted of a Kodak Carousel slide projector, a Gerbrands tape timer, an electric metronome, and 2×2 in. slides on which the material was printed from top to bottom, each word being one space to the right of the one above it.

Procedure.—After *S* was seated in the experimental room, he was read the instructions which are typical for the Peterson paradigm. The *S* was then asked if he had any questions, and, if not, the experiment began.

An asterisk appeared for 2 sec. as a signal to *S* that a triad was about to be presented. The triad was then presented for 3 sec. and *S* recited it once. The next slide appeared for 20 sec. and consisted of a two-digit number. The *S* was to count backwards by threes from the number. Next, a question mark appeared for 6 sec. and *S* was to recall the three words in the triad, in order if possible. Immediately after the question mark, the asterisk appeared signaling the next trial. All *S*s received four trials in this manner, the intertrial interval being 30 sec.

Each *S* was given one of four programs. One group, hereafter called H/L, received three triads of Hi I-C words and shifted to a Lo I-C triad on the fourth trial. The control group for H/L, hereafter called L/L, received four triads of Lo I-C words. One group, L/H, received three triads of Lo I-C words and shifted to Hi I-C words for the fourth trial. The control group for L/H, hereafter called H/H, received four triads of Hi I-C words. Each of the 16 Hi I-C triads and 16 Lo I-C triads appeared equally often, and the order of appearance was counterbalanced to insure that each would appear equally often on each of the four trials.

In analyzing the data, an *S* was given a score of four if the three words were given in the correct order. If not in order, *S* received a score of three, one for each correct word. No bonus for order was given if fewer than three words were recalled.

Results

The percentage correct recall of the four groups on all four trials is presented in Fig. 1.

Production of PI.—A two-way analysis of variance for the first three trials showed a significant interaction effect, $F(6, 756) = 3.14$, $p < .01$, indicating that the change in performance across trials was different for the four groups, so separate one-way analyses were done for all four groups to test for PI development. H/H, H/L, L/L, and L/H all showed significant PI, with $F(2, 756) = 43.12$, 27.85, 58.77, and 76.14, respectively, p's $< .001$.

Because, for the first three trials, Groups H/H and H/L received the same treatment and Groups L/L and L/H received the same treatment, the groups were combined to give an overall picture of PI buildup in the high- and low-imagery conditions. Figure 2 gives the results of this combination. The analysis of variance again showed significant interaction effects, F (2, 756) = 5.00, $p < .01$, indicating that the change in performance of the Hi I-C and Lo I-C groups was different across trials. Further analysis of individual trials indicated that the two groups did not differ from each other on Trial 1, but did so at the $p < .001$ level on Trials 2 and 3.

Release from PI.—The results for Trial 4 can be seen in Fig. 1. Scheffé tests for post-hoc comparison (Hayes, 1963) were conducted to test for release from PI. When Group H/L is compared with the appropriate control, Group L/L, on Trial 4, there is no significant difference. Likewise, there is no significant difference between Group L/H and its appropriate control, Group H/H. If the two groups that received Hi I-C triads on Trial 4 are combined and the two groups that received Lo I-C are combined, the two points are significantly different ($p < .01$). However, the combined experimentals do not differ from the combined controls.

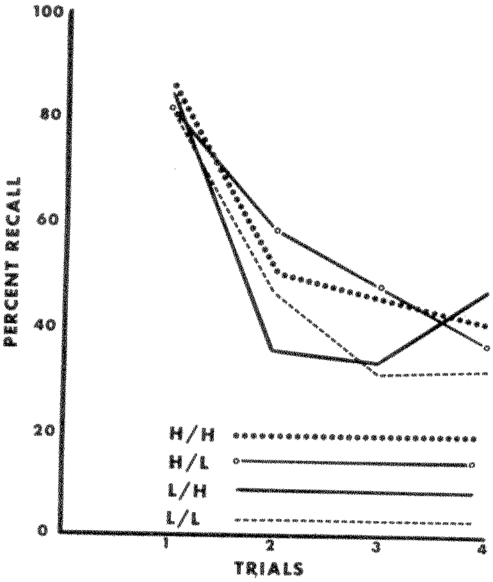

Fig. 1. Recall performance of all groups across all trials.

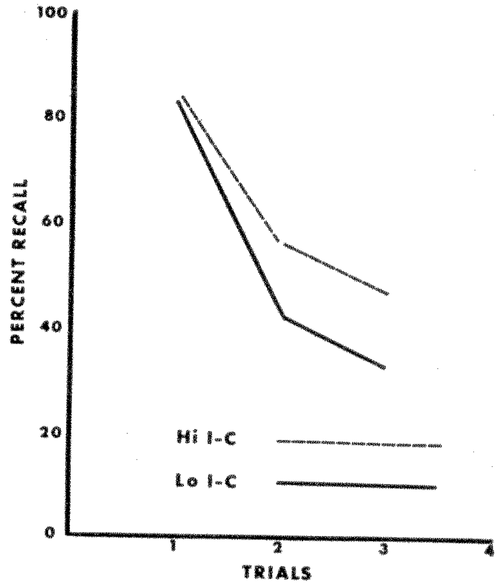

Fig. 2. Recall performance of all high- and low-imagery groups during the first three trials.

Discussion

Imagery and encoding.—If one may assume that a release from PI is indicative of a shift in the manner of encoding and a failure to obtain release indicates that the two logically different classes of materials are encoded in the same manner in STM, then the experiment offers little evidence for the view that words of high imagery value are, as a class, encoded differently from words of low imagery value. An inspection of Fig. 1 shows that the performance of the experimental groups on Trial 4 was determined primarily by the materials experienced on that trial rather than the differential histories across the first three trials. It is true that the L/H group improves from Trial 3 to Trial 4, but its performance on Trial 4 is not significantly above that of the control, while the H/L group actually declines in performance from Trial 3 to Trial 4.

It should be noted, however, that in each instance the performance of the experimental group slightly excels the appropriate control, suggesting a slight amount of release from PI—though, as noted in the Results section, this difference does not attain significance even when the groups are combined to achieve a total of 128 Ss in the experimental and control groups. Assuming the gain is a real one, it is still quite small as compared with the effects produced by such other dimensions as taxonomic classes, Osgood Semantic Differential dimension, sense impression, and word frequency (Wickens, 1970). It would seem justifiable to conclude that imagery-abstractness is a relatively unimportant encoding dimension in a word perception situation of the sort represented by the present STM experiment.

The preceding conclusion is in no way meant to deny the occurrence of imagery, or to imply that it cannot be a useful tool in learning; but it is intended to raise doubts about the extent of its use in our ordinary daily dealing with words. According to a study by Moore (1915), it takes about 2.5 sec. for a word to generate an image; and in the present experiments, the three words were presented for only a total of 3 sec., or a little more than enough time to conjure up an image of but a single word. Ordinary speech also flows too rapidly to permit one to dawdle over an item while awaiting its image. Thus, the demands of the linguistic environment may be such as to discourage the use of imagery as an encoding dimension.

Imagery and word recall.—Although there is little evidence in support of encoding by imagery, Fig. 2 and the statistical analysis of those data clearly indicate that retention performance on high-imagery words is superior to that on abstract words—thus supporting, in general, the typical finding. Of particular interest in these data are the significant Trials × Conditions interaction and the subsequent tests of the individual trials which show a significant difference between the two conditions on Trials 2 and 3, but not on Trial 1, where the two groups are much alike—with a recall score of 84 for the high-imagery condition and 83 for the abstract condition. It is apparent that these relationships could be attributed to a ceiling effect on Trial 1, and if this conclusion is accepted, no more need be said on the topic.

It should be pointed out, however, that both groups are a fair distance from the ceiling and that there *is* room for the imagery condition to show its superiority over the abstract condition on the first trial, just as it does on the succeeding trials. If the empirical data are correct reflections of *S*s' capacity to perform on each trial, they offer an understanding of why high-imagery items usually produce better performance than do abstract items.

The decline in recall which is so characteristic of this STM paradigm (Wickens, 1970) can best be attributed to interference produced by the earlier items on later trials. Applied to the present data, this interpretation would state that on Trial 1, interference was minimal, the only source being whatever small amount of intralist interference would arise from three words, and thus the two groups are essentially alike. On Trials 2 and 3, however, the amount of interference has been increased for both groups, so both of them decline. The decline is less for the high-imagery groups, thus implying that *high-imagery items produce less interference than do abstract items,* and this is the psychological mechanism responsible for the usual superiority of performance on high-imagery material over abstract material. It is especially important to note that the occurrence of only one previous item—a mere three words—is sufficient to result in a marked decline in performance. Studies on imagery, whether in serial, paired-associate, or free recall learning, tend to use at least 8 or 10 items, and thus by the time of completion of the first trial, very close to the maximum amount of interference will have been introduced, the interference being intralist in nature. Since the present experiment suggests that interference is greater for abstract than for concrete items, the groups should differ from each other from the first trial onward.

That abstract items interfere with each other more so than do concrete items is not surprising. The concept of abstractness implies that the word in question is broad and inclusive in its meaning. Thus, the probability of obtaining semantic overlaps between several abstract words is far greater for abstract than for concrete words. Even a randomly chosen list of abstract words might be expected to have more interitem overlap, and hence interference, than a randomly chosen list of concrete words. It will be noted that the word "concrete" has been substituted for "high-imagery" in the preceding argument, but it should also be noted that these two concepts correlate .83 in the Paivio et al. (1968) norms.

An additional support for the view that the basis for superior performance on imagery items over abstract items is due to differential interference is found in the results of the shift aspect of the experiments. This portion of the experiment offers no real evidence in favor of the view that there is an overall differential in the way abstract and imagery words are encoded. This negative finding raises doubt about a differential encoding interpretation, whereas the data on the differential buildup of PI implicate an interference interpretation.

REFERENCES

HAYES, W. L. *Statistics*. New York: Holt, Rinehart & Winston, 1963.

KEPPEL, G., & UNDERWOOD, B. J. Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, 1962, **1**, 153–161.

LOESS, H. Short-term memory, word class, and sequence of items. *Journal of Experimental Psychology*, 1967, **74**, 556–561.

MOORE, T. V. The temporal relations of meaning and imagery. *Psychological Review*, 1915, **22**, 177–224.

PAIVIO, A. Mental imagery in associative learning and memory. *Psychological Review*, 1969, **76**, 241–263.

PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 1968, **76**(1, Pt. 2).

PAIVIO, A., YUILLE, J. C., & ROGERS, T. B. Noun imagery and meaningfulness in free and serial recall. *Journal of Experimental Psychology*, 1969, **79**, 509–514.

PETERSON, L. R., & PETERSON, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, **58**, 193–198.

WICKENS, D. D. Encoding categories in short-term memory: An empirical approach to meaning. *Psychological Review*, 1970, **77**, 1–15.

WICKENS, D. D., BORN, D. G., & ALLEN, C. K. Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 440–445.

(Received October 20, 1969)