Journal of Experimental Psychology: Human Learning and Memory

Vol. 2, No. 3 May 1976

A Processing Approach to the Dual Coding Hypothesis

Stephen M. Kosslyn
The Johns Hopkins University

Keith J. Holyoak and Catherine S. Huffman
Stanford University

This paper explores the notion that imagery encoding differs from other forms of encoding with regard to the particular processes invoked. Words encoded as images were recalled after words that were rehearsed, and words clustered in output according to processing mode. Clustering by modality occurred even when grouping according to semantic categories conflicted with grouping by processing mode. In addition, more words were recalled if both verbal and visual encoding was utilized, compared with when only one mode was used. Two selective interference tasks revealed that silent verbal rehearsal seems to involve some of the same processing mechanisms activated during encoding of phonemic information. However, these tasks revealed no effects of visual perception selectively interfering with visual imagery encodings.

People often have been credited with the ability to encode information either verbally and/or in terms of mental images (see Paivio, 1971). Recently, this notion has been attacked by theorists who postulate a single, abstract mode of internal representation (e.g., Pylyshyn, 1973). Rather than attempting to determine the ultimate form of representation in long-term memory, it may be more profitable to see whether there are functional differences between different modes of representing information. Instead of focusing on the end products of encoding, we will consider whether the encoding process itself differs in important ways when people encode visual images of the referents of words as opposed to simply encoding the words themselves (i.e., rehearsing verbally). The present approach to the dual coding hypothesis centers on the idea that encoding visual images as opposed to verbal representations requires *different processing channels.* We explore the implications of this idea with regard to the nature of encodings shortly after they are formed. The present study can be divided into two parts. In the first two experiments we investigate whether imagery and verbal encoding use different processing mechanisms. The final two experiments then attempt to discover whether the processes underlying the use of imagery to retain words are also involved in like-modality perception.

If imagery and verbal encoding require different processing channels, we might expect material being encoded to become organized in terms of modality. For example,

We wish to thank E. Smith, A. Glass, and H. Egeth for helpful remarks. We especially wish to thank Gordon Bower who extended both perceptive comments and financial support (NIMH Grant No. MH13950-08). In addition the present research was supported by Biomedical Sciences Grant 3 S05 RR07041-09 awarded to The Johns Hopkins University by The Division of Research Resources, DHEW. Finally, this research was completed while K. Holyoak was supported by NIMH Grant MH 20021 to Herbert Clark.

Requests for reprints should be sent to Stephen M. Kosslyn, Department of Psychology, Johns Hopkins University, Baltimore, Maryland 21218.

items which are imaged visually will share certain properties by virtue of their common mode of representation. Items that are processed in like manner may thus come to be associated together. Common modality should therefore be a source of similarity between remembered items as a result of mutually utilized types of cognitive processing.

This consideration led us to expect that if subjects learned lists of words in which some items were visually imaged and others rehearsed, later free recall would exhibit clustering according to modality. If modality is a salient property of a representation, its organizational effects might be expected to persist even when they conflict with the powerful effects of semantic categories, which typically largely determine recall clustering (Bousfield, 1953). In addition, the properties of different modalities should be important in predicting order of recall. Visual images, once constructed, seem relatively stable, whereas auditory images of words must be rehearsed continuously. We therefore expected that subjects would "unload" the more fragile auditory representations first, and "read off" the visual images only subsequently.

Experiment 1

Method

Subjects heard 20 lists of eight words each. Each word was preceded by the ringing of a bell or a snap of the fingers, which cued subjects either to image or to rehearse the word (for two instruction groups), or simply to include the item in one of two arbitrary groups (for the third group). After the eight words had been presented, subjects wrote down all of the words they could remember in the order that they recalled them.

The lists were composed of concrete nouns belonging to various simple semantic categories (e.g., clothing, animals). Each list included four words each of two different categories. Four of the words in each list were paired with a bell, and four were paired with a snap. For half of the lists (the congruent condition), the division according to snap or bell coincided with the division by semantic category. For the other half of the lists (the noncongruent condition), two words from each category were paired with a bell, and two words from each category were paired with a snap. The words within each list were then ordered randomly.

Six sets of test booklets were prepared. These differed only in the particular learning instructions printed on the first page. Three basic instructions were used, with two versions of each one differing only in the assignment of the bell and snap cues to conditions. Twenty subjects were assigned to each of the three instruction groups (half of each group receiving one version of the instructions and half receiving the other). One group was instructed to form separate images of the referents of words preceded by one cue and to verbally rehearse words preceded by the other cue. They were told that the verbal method was "to repeat the words silently to yourself as rapidly as possible until you are asked to recall the list," while the imagery method was "to make separate visual images, or 'mental pictures', for each of the words." It was emphasized that "no image should touch any other image." Subjects were told that the best procedure was to picture each of the four imaged objects in a separate compartment of a horizontal bookshelf. A second group received identical instructions, except that they were told to form images that did not remain separate but interacted and touched in some way. The third group served as a control. These subjects were not given any specific instructions about learning strategies; rather, they simply were told to organize the items into two groups in memory according to whether they were associated with a bell or a snap. The remainder of each test booklet contained spaces for subjects to write down the words they could recall from each list.

The experiment was conducted in two group sessions. In each session subjects were randomly assigned to the various instructional conditions. Subjects were asked to read their instructions carefully and raise their hands if any questions arose. Two practice lists were then administered. Practice trials used words that did not appear in the test lists but were like the test trials in all other respects. The procedure for the 20 test lists was as follows: The experimenter announced the list number and then read the words in the list (in the previously randomized order) at the rate of one word every 4 sec. Each word was immediately preceded by the appropriate cue (either bell or snap). Five sec after the last word of the list, the experimenter said "Write," and subjects were given 30 sec to recall the list. Subjects had been explicitly instructed that they could recall the words in whatever order seemed easiest, but that it was important that they wrote them down in that order. The experimenter then announced the next list number, and the procedure was repeated. At the end of the experiment subjects were debriefed and queried as to their compliance with instructions.

Sixty Stanford University introductory psychology students participated in the experiment as part of a course requirement.

Results

Both subjects and lists were treated as random effects in all analyses of variance reported in the present paper, necessitating use of the quasi-F statistic, F' (Clark, 1973; Winer, 1971). Because the quasi-F statistic is inherently conservative, the .05 level of significance was used. Following each F' value is the mean square error for subjects (SMS_e) and items (IMS_e) used in the error term for that comparison.

The various learning strategies (e.g., separate imagery, rehearsal) did not produce significantly different recall scores since there was a clear ceiling effect (overall, a mean of 7.15 words per list—out of the 8 possible—were correctly recalled). This was, of course, desirable because our interest was in clustering and output order. For each list we calculated a clustering score for the words associated with each of the two cues. The clustering measure we used actually was more a measure of disruption of clustering. Perfect clustering occurred when all words encoded in a given modality were recalled contiguously. When like-mode words were separated during output by words learned in the other modality, less clustering occurred. Because items recalled from a given modality could vary from zero to four, we calculated the mean *proportion* of different-cue items that were inserted between same-cue pairs during recall. This proportion was calculated in two ways: as the ratio of intervening different-cue items to the number of different-cue items recalled, and as the ratio of intervening different-cue items to the total number of items recalled from the entire list. These two measures produced essentially identical results, so only the former will be reported here. It is impossible to have a meaningful clustering score for a cue class in which only one or none of the items is recalled (in those cases there are no pairs between which items could intervene). For such cases (4.2% of the scores) the subject's mean score for that cue class in the remaining lists was used as a replacement.

The clustering results are presented in Table 1. Since clustering did not differ between rehearsed and imaged items (nor, of course, between the two arbitrary groups according to which control subjects learned each list), the data in Table 1 were obtained by collapsing across this variable. Note that the larger the number in the table, the greater the mean number of intervening items (i.e., the less clustering by modality). Clustering was considerably above chance level, with less than 25% of the potential intervening items on the average actually being recalled between like-modality pairs. Clustering did not differ between the separate and interactive imagery conditions. The lack of greater clustering with interactive imagery instructions, which presumably should produce greater explicit organizational effects, may have been due to an indirect ceiling in the amount of clustering that could be expected in such short lists. Subjects often recalled the last two or three words of the list first, regardless of modality or category, so that the effective range of clustering did not reach the theoretical maximum. The reduction in overall clustering resulting from this strategy should not be as great if longer lists were used, and a difference between interactive and separate imagery might then emerge.

TABLE 1

Recall Clustering by Learning Modality and Category–Cue Congruence

Learning modality	Congruent	Noncongruent
Interactive imagery/verbal rehearsal	.081	.166
Separate imagery/verbal rehearsal	.073	.169
Control (arbitrary grouping)	.054	.300

Note. Larger numbers indicate less clustering. (If recall was perfect, absolutely no clustering would be indicated by a score of about .417.)

The important comparisons involve the conditions in which the learning cues were congruent with the semantic categories (e.g., all animals were preceded by a cue to rehearse and all vehicles by a cue to image), as opposed to the conditions in which cues and categores were crossed (e.g., two animals and two vehicles were rehearsed, and two of each category were imaged). As would be expected, all instruction groups showed a high degree of clustering by cues in the congruent conditions when these cues corresponded to different semantic categories. In these conditions the three groups did not differ significantly in clustering. However, the situation was quite different in the noncongruent conditions. For the control group, for which the cues now represented two arbitrary groupings of the list items, clustering by cues was sharply reduced by the now conflicting effect of semantic categories. However, in the remaining two groups the cues signaled different modes of representation—imaging versus verbal rehearsal. Although clustering also was reduced for these subjects in the noncongruent condition, it was significantly greater than that shown by the control subjects, $F'(1, 70) = 19.8$, $SMS_e = .23$, $IMS_e = .03$. The overall interaction between category-cue congruence and learning instructions was highly significant, $F'(2, 70) = 13.4$, $SMS_e = .10$, $IMS_e = .02$. These results indicate that organization by modality persists in memory even when it conflicts with the well-known powerful effects of semantic categorization.

The second measure of interest was the effect of mode of representation on order of recall. In calculating this measure, cases where no words of a class were recalled were replaced by the subject's mean score for that condition; this occurred in only .4% of the lists. Table 2 presents the mean output order of items for the various conditions. As expected, rehearsed items tended to be recalled earlier than imaged items (whether separate or interactive), $F'(1, 67) = 64.7$, $SMS_e = 26.60$, $IMS_e = 5.35$. The control group showed no preferential ordering of the two arbitrary item groups signaled by the cues. This interaction was significant, $F'(2, 67) = 6.66$, $SMS_e = 26.60$, $IMS_e = 3.84$.

TABLE 2

MEAN OUTPUT BY LEARNING MODALITY

Learning modality	Imaged	Rehearsed
Interactive imagery/verbal rehearsal	5.08	3.19
Separate imagery/verbal rehearsal	4.60	3.42
Control (arbitrary grouping)	4.07	4.15

Note. These data are based on 8-word lists.

Discussion

Two results of interest emerged from this experiment. First, words encoded in a given modality tended to be recalled together. Second, rehearsed words were recalled prior to imaged words. These results are consistent with the view that separate processing capacities are involved in the two forms of encoding. However, it might also be argued that the results are simply due to quantitative differences in "depth-of-processing" (Craik, 1973; Craik & Lockhart, 1972). That is, perhaps rehearsal simply requires less of the same *sort* of processing required by imagery. Since these rehearsed words are less well encoded, and will be forgotten first, they are recalled as soon as possible (and thus tend to be recalled as a group). In order to counter this explanation, we need to demonstrate that using visual imagery in this situation does not always result in better memory, as a simple depth-of-processing account would lead us to expect. The present conceptualization suggests that if one of the encoding systems is overloaded with information to be processed, the other system should still be able to accommodate some additional information. In other words, encoding capacity should be maximal when a person is using both imagery and verbal rehearsal rather than just one mode of representation. Experiment 2 was designed to test this hypothesis.

EXPERIMENT 2

We sought to overload subjects' immediate memory by asking them to actively rehearse and then recall 12-word lists. On some trials they verbally rehearsed all 12 words, and on some trials they imaged the referents of all 12 words, whereas on other trials they imaged the referents of six words and verbally rehearsed the other six. Our prediction was that recall would be optimal in the mixed-strategies condition. Note that this hypothesis does not imply that the mixed condition will necessarily produce higher recall than the all-imagery condition, since we know from other work (e.g., Bower, 1972) that forming separate images is simply a more effective learning strategy than verbal rehearsal. Rather, the proper prediction is that recall of any half-list (six words learned by one method) will be better when the other half-list is learned by a different method. For example, the first six words of a list, if learned by verbal rehearsal, should be better recalled if the last six words are learned by imagery rather than verbal rehearsal. Conversely, recall of a second half-list learned by imagery should be better if the first half-list was learned by verbal rehearsal. Stated in another way, the hypothesis is that learning a half-list by a particular strategy will interfere less with recall of the other half-list if the latter is learned by a different strategy.

Two critical assumptions underlie the above experimental predictions. The first is that imaginal and verbal rehearsal use (to some degree) different processing resources. The second assumption is that neither mode of representation is involved in initiating the other. For instance, we must assume that it is not necessary to image the referents of words in order to generate their names for verbal rehearsal. This assumption seems straightforward, but its converse is not—that is, do people generate images for words without initially verbalizing their names? At least for our particular experimental situation, it seems likely that the answer is no. Subjects hear the words on tape and must maintain each word in some form until the information required to form an image is retrieved from long-term storage. Under these circumstances it appears probable that subjects will initially rehearse the names of words until they are able to produce an appropriate image.

This consideration suggested that order of strategies in the mixed-strategy condition might be critical in determining recall. Consider the case in which subjects image the first half-list and verbally rehearse the second. When the first half-list is being processed the verbal rehearsal system is free. The subject can therefore verbalize, if necessary, while forming images and then discard the verbal representation of the first half-list in time to rehearse the second. Accordingly, the predictions outlined above should hold for the imagery-verbal order. But now consider the situation in which the imagery block follows the rehearsed items. When the subject attempts to image the second half-list, his capacity for verbal rehearsal will already be taxed in maintaining the first half-list. So if he now verbalizes the names of words while forming images, he will draw valuable processing resources away from rehearsal of the first half-list; if he does not, his ability to image the second half-list may be impaired. This suggests that the mixed-strategy condition may only improve recall in the imagery-verbal order. Accordingly, the order of strategies in the mixed-strategy condition was systematically varied in Experiment 2.

Method

Subjects heard 24 lists of words recorded on tape, and after each list immediately attempted to recall as many list words as possible. Each list consisted of 12 concrete nouns. Approximately two thirds of these nouns were selected from those with rated imagery value greater than 6.0 (on a 7-point scale) in the norms of Paivio, Yuille, and Madigan (1968). The remainder were selected for their intuitive concreteness by the experimenters. Subjects heard the label "A" prior to the first 6 words of any list and then heard the label "B" prior to the second 6 words. An accompanying answer sheet indicated the learning strategy (verbal or imagery) to be used for Parts A and B of each list. For eight lists, both half-lists were to be learned by imagery; for another eight, both half-lists were to be learned by verbal rehearsal; and for the remaining eight lists, one part of the list was to be learned by one strategy and the other part by the other strategy. For half of the latter lists (the mixed-strategy conditions), the "A" half-list was learned by imagery and the "B" half-list by verbal rehearsal, whereas for the other half of the lists the order of strategies was reversed. Each subject thus served in all experimental conditions.

Six versions of the answer sheet were used, so that across subjects each list served equally often in each condition. Order of conditions within each list was again counterbalanced. Words were recorded at 5-sec intervals (except that 10 sec intervened between the two halves of the list). Five sec after the final word, the experimenter said "Write," at which point subjects were given 30 sec to recall the list. Two practice lists, illustrating the verbal-imagery and verbal-verbal conditions, preceded the 24 test lists.

The learning instructions given to subjects at the beginning of the experiment were essentially the same as the separate-image and verbal-encoding instructions used in Experiment 1. Subjects were told to picture mentally the six objects in a half-list to be learned by imagery as if they were in separate compartments of a horizontal bookcase; if both half-lists were to be imaged, subjects were told to imagine a second bookcase below the first.

Twenty-four Stanford University undergraduates, different from those in Experiment 1, participated in the experiment either for pay or to satisfy a course requirement.

Results

The mean number of words recalled for each half-list is presented in Table 3, broken down according to the four experimental treatments. Generally, the conditions were ordered from highest to lowest mean recall as follows: imagery-verbal, imagery-imagery, verbal-verbal, and verbal-imagery. As suspected, the mixed-strategy condition only facilitated learning in the imagery-verbal order. This conclusion was bolstered by the results of analysis of variance. Analyses were performed separately on means obtained by collapsing across subjects, on one hand, and across lists, on the other hand, and minimum quasi-F ratios were calculated (Clark, 1973). Order of strategies in the mixed condition significantly influenced the difference between the mixed- and single-strategy conditions, $F'(1, 45) = 5.49$, $SMS_e = 2.08$, $IMS_e = 2.84$. More words were recalled from the first half of a list learned by imagery if the second half was learned by verbal rehearsal, and more words were recalled from the second half of a list learned by verbal rehearsal if the first half was imaged (3.82 vs. 3.42 words per half-list). Yet, although mixed strategies increased recall if applied in the order imagery-verbal, recall in the verbal-imagery condition did not differ significantly from performance in the corresponding single-strategy conditions (3.13 vs. 3.25 words).

TABLE 3

Mean Words Recalled per Learning Strategy

Learning strategy	1st half	2nd half
Imagery-imagery	2.78	4.27
Verbal-verbal	2.23	4.06
Imagery-verbal	3.00	4.64
Verbal-imagery	1.96	4.30

Several other recall differences also were evident. First, generation of separate images produced overall higher recall than did verbal rehearsal (3.59 vs. 3.22 words per half-list), $F'(1, 34) = 4.17$, $SMS_e = 1.26$, $IMS_e = .32$. Second (and not surprisingly), there was a strong recency effect, such that recall generally was higher from the second half than from the first half of lists (4.32 vs. 2.49 words), $F'(1, 39) = 61.6$, $SMS_e = 1.83$, $IMS_e = .76$. Of interest is the fact that the effectiveness of imagery and verbal strategies in learning differed between the first and second halves of the lists. For the first half of the lists, imagery was the better strategy, whereas for the second half, verbal rehearsal was superior. This interaction was highly significant, $F'(1, 46) = 13.5$, $SMS_e = .36$, $IMS_e = .31$. This finding is consistent with those of Smith, Barresi, and Gross (1971), who found that imagery encoding was superior only when relatively long-term "secondary memory" rather than short-term "primary memory" was later assessed. Further, Mazuryk (1974) has shown that although semantic encoding enhances recall from long-term memory, it actually is detrimental to recall of information presumably in primary memory. If we consider the second half of a list to be in primary memory and the first half in secondary memory, then the present results also suggest that imagery produces superior recall only when secondary memory is tapped. It should be noted that these effects are independent of the comparisons between mixed and blocked strategies reported above. Appropriate counterbalancing and comparing half-lists occurring in the same position eliminate any confounding between list position and strategies.

Discussion

Immediate recall is greatest when different portions of a list are maintained by different modes of representation. This result confirms the major hypothesis of Experiment 2. The superior recall produced by the imagery-verbal condition supports our claim that imaginal and verbal rehearsal are distinct modes of representation in active memory and that the two modes to some extent tap different processing resources. The finding that mixing strategies is only effective in the imagery-verbal order and not when the order is verbal-imagery is consistent with the view that image formation in our experimental situation follows an initial period in which the name of the word is maintained verbally.

The above experiments demonstrate functional differences in encoding processes but tell us nothing about these processes themselves. Numerous researchers (Brooks, 1967, 1968; Segal & Fusella, 1970) have provided support for the idea that imaging and perceiving in the same modality utilize more common processing than imaging and perceiving in different modalities. This inference is based primarily on findings that visual imagery selectively interferes with visual perception more than with auditory perception. No evidence exists, however, that the imaginal processes involved in encoding *verbal* material utilize mechanisms usually active only during like-modality perception. Byrne (1974) showed that classification of imaged items was interfered with when spatial processing occurred, but not when mere pictorial processing was introduced. Atwood (1971) appeared to show that like-modality perception selectively disturbed memory for words encoded via imagery; unfortunately, however, this experimental result repeatedly has failed to replicate (see Anderson & Bower, 1973; Bower, Munoz, & Arnold, Note 1; Brooks, Note 2). Experiment 3 incorporated a selective interference paradigm in an attempt to discover whether imagery mnemonics in verbal learning utilize processing resources usually restricted to like-modality perception.

Experiment 3

Method

Subjects were asked to learn 8-word lists either by using imagery or by verbal rehearsal. After each list subjects performed an interpolated letter-scanning task, prior to free recall of the list. Two versions of the scanning task were used. One task was designed to require primarily verbal processing (identifying letters that rhymed with a target), whereas the other task was designed to require primarily visual processing (identifying letters with a particular visual feature). The prediction was that the rhyme-matching task would selectively interfere with recall of verbally rehearsed lists, whereas the visual-matching task would selectively interfere with recall of imaged lists.

Materials. The material to be memorized consisted of 24 8-word lists. These words were a subset of those used in Experiment 2. Words were assigned to lists randomly, and each word was used only once. The 24 lists were tape recorded, with a 5-sec interval between each word.

In addition, 24 lists of letters were constructed for use in the scanning tasks. Twelve of these lists were used in the rhyme-match task, and 12 were used in the visual-match task. Each list consisted of 48 letters. For each list, a target letter was placed randomly in one of the four positions within each block of four letters, so that exactly 12 target letters occurred in each list. The lists for the rhyme-match and visual-match tasks were yoked, so that for one list in each condition target letters were located in the same sequence, with the same number of letters intervening between each.

For the rhyme-match task, target letters consisted of those that rhyme with the word *key* (e.g., *b, B, c, C, e, E*). For the visual-match task, the targets were letters containing a curve concave to the right like a *c* (e.g., *a, c, C, d, e*). Both uppercase and lowercase letters were used in constructing the lists for both tasks, so that the visual-match task could not be reduced to simply a name match (e.g., lowercase *a* and *e* contain the target visual property, but uppercase *A* and *E* do not). For both tasks each possible target letter was used approximately equally often in constructing the lists. Each of the 24 lists of letters was typed in a single line across a separate page in an answer booklet.

Procedure. Each subject learned 12 of the word lists using imagery and the other 12 using verbal rehearsal. Within each of these learning conditions, six lists were followed by interpolated rhyme-match task and the other six by a visual-match task. Every subject thus served in all four experimental conditons.

At the beginning of the experiment subjects were given instructions concerning each learning method (verbal and separate imagery) as in the previous experiments. Subjects were instructed to "keep the two learning methods separate and distinct" (i.e., to avoid rehearsing when using imagery and to avoid using imagery when rehearsing). The two types of scannng tasks also were described. Subjects were told to go through each list of letters as quickly as possible, crossing out the appropriate target letters.

Subjects were given test booklets containing two pages for each list. For each list the first page named the appropriate learning method at the top and the scanning task which was to follow at the bottom (e.g., "verbal" and "rhymes with key"). The second page again contained the name of the scanning task, with the letter sequence typed in a line underneath. The bottom of the second page contained spaces to recall the eight words of the list.

The sequence of events on a given trial was as follows: First, subjects were given 5 sec to read the page identifying the learning method and type of scanning task. The recording of the 8-word list was then played and subjects attempted to memorize it by the indicated method. Five sec after the final word the exerimenter said "Go," at which point subjects turned the page and proceeded to the scanning task. They were allowed 12 sec to work at the scanning task. The experimenter then said "Stop and write," and subjects were given 30 sec to write down, in any order, as many list words as they could recall. They then turned to the next page of the booklet and prepared for the next list.

The initial instructions emphasized that for any list subjects were only to use the learning method indicated on the answer sheet. They were told that "it is extremely important that you use exactly the methods indicated on the answer sheet, even if they don't help you that much, since we are specifically interested in the effects of these particular strategies on learning." Two practice trials were given prior to the test trials, using non-test items. The practice trials illustrated each type of scanning and each type of encoding.

Four different sets of booklets were used, so that across subjects each word list occurred equally often with all learning and scanning conditions, and each scanning list occurred equally often after both imagery and verbal-rehearsal conditions. Within any booklet the four experimental conditions each occurred once in every block of four lists, and across subjects each condition was represented equally often at all positions within the booklets.

Thirty-six Stanford University undergraduates served as subjects either for pay or for course credit; none of these people had participated in either previous experiment. Testing was done in two group sessions, each lasting about 45 min.

TABLE 4

Mean Words Recalled and Mean Target Letters Identified per Learning Method and Scanning Task

Learning method and scanning task	Words recalled	Targets identified
Verbal rehearsal		
Rhyme-match	4.98	6.30
Visual-match	5.23	6.75
Imagery		
Rhyme-match	5.87	6.71
Visual-match	5.77	6.92

Note. Maximum of 8 words per list and 12 target letters per list.

Results and Discussion

The result of major interest concerns differential impairment of recall engendered by the two scanning tasks, depending on the type of learning instructions. Table 4 presents the mean number of words per list recalled for each of the experimental conditions, as well as the mean number of target letters correctly identified. The interaction between learning method and scanning task was significant, $F'(2, 15) = 3.75$, $SMS_e = 1.01$, $IMS_e = 1.19$. Under verbal-rehearsal instructions, recall was slightly lower after the rhyme-match task, $F'(1, 21) = 5.79$, $SMS_e = 1.39$, $IMS_e = 1.42$, suggesting a like-modality interference effect; but under imagery instructions, recall did not differ significantly as a function of the scanning task, $F' < 1$. This result suggests that items coded imaginally are less susceptible to acoustic interference; however, they apparently do not become significantly more susceptible to visual interference.

Pooling over interference tasks, subjects recalled more words after forming images than after verbally rehearsing the items (5.82 vs. 5.10 words per list), $F'(1, 27) = 11.5$, $SMS_e = 6.48$, $IMS_e = 3.37$. This superior recall under imagery instructions may be explained in terms of deeper levels of processing (Craik, 1973; Craik & Lockhart, 1972). Presumably, the formation of separate images of the referents of words requires a greater degree of semantic processing than does verbal rehearsal. That is, to generate an image of a word its reference must be comprehended. To rehearse the word itself, however, one need not necessarily comprehend the meaning at all.

The number of target letters correctly identified in the scanning tasks was also examined. The frequency with which subjects incorrectly checked a nontarget letter was extremely low ($< 1\%$), so the scanning data were analyzed without a correction for guessing. Subjects were able to identify significantly more target letters after learning a word list by imagery than after learning by verbal rehearsal (6.81 vs. 6.53 target letters per list), $F'(1, 26) = 6.00$, $SMS_e = 1.06$, $IMS_e = .17$. This difference supports the conclusion that imagery is a more effective mechanism for encoding words into memory than is verbal rehearsal. Presumably the subject can afford to transfer a larger portion of his processing capacity to either scanning task, and still be able to later retrieve the list words, after forming images than after verbal rehearsal. This result also suggests that letter scanning per se does not produce a great deal of visual interference. Finally, imagery did not disrupt visual matching more than rhyme matching; the interaction between learning method and target type was not significant, $F'(2, 27) = 1.43$, $p > .25$, $SMS_e = 1.91$, $IMS_e = .17$.

Experiment 4

Experiment 3 provided little evidence that engaging in a visual-search task selectively interferes with memory for verbal items that were learned by forming visual images. However, there is always the possibility that the interference tasks used simply were too easy to create any noticeable disruption of recall. In Experiment 4, therefore, we repeated the interference paradigm introduced in Experiment 3, with modifications intended to make the interference task more difficult.

Method

The materials and procedure employed in Experiment 4 were identical to those of Experiment 3, except for some changes in the scanning task. On each trial, subjects were instructed to cross out letters that rhymed with either of two targets, such as *buy* (*i, Y*) or *gay* (*A, K*), or letters that contained either of two visual features, such as a curve concave to the right (*C, d*) or a horizontal bar (*A, t*). A total of four different rhyme and four different visual targets were used, with all possible pairs occurring equally often across trials. Subjects did not know what targets they would be searching for on a particular trial until they turned the page in their test booklet after encoding the eight-word list.

In addition, the duration of the scanning task was increased from 12 to 25 sec. A maximum of 80 letters could be scanned on each trial, typed in two rows of 40. The interference tasks in Experiment 4 therefore involved greater uncertainty than those in Experiment 3 (eight rather than two possible targets), a heavier memory load during scanning (two targets on each trial instead of one), and lasted twice as long on each trial. These changes were expected to increase the amount of processing capacity that would be occupied by the scanning tasks, and hence increase the likelihood of detecting selective interference between visual imagery and visual scanning.

Subjects were 28 Stanford University undergraduates (different from those who served in previous experiments), who participated either for pay or for course credit.

Results

The mean number of words recalled for each of the experimental conditions, and the mean number of target letters correctly identified, are presented in Table 5. The recall results essentially mirror those obtained in Experiment 3, except that overall recall was somewhat lower, suggesting that the interference tasks in fact were more effective in Experiment 4. Under verbal-rehearsal instructions recall again was lower after the rhyme-match task, $F'(1, 14) = 7.60$, $SMS_e = 1.29$, $IMS_e = 2.32$, but recall of words encoded via visual imagery was the same following both types of scanning tasks. Overall, recall was higher after forming images than after verbally rehearsing $F'(1, 14) = 6.61$, $SMS_e = 5.75$, $IMS_e = 3.01$. An analysis of the number of target letters correctly identified revealed no effects close to significance. Notably, using visual imagery as a learning strategy did not impair visual matching more than rhyme matching.

General Discussion

The experiments reported here investigated two interrelated issues concerning the role of imagery in determining memory for verbal items. In the first two experiments

TABLE 5

Mean Words Recalled and Mean Target Letters Identified per Learning Method and Scanning Task (Experiment 4)

Learning method and scanning task	Words recalled	Targets identified
Verbal rehearsal		
Rhyme-match	4.50	6.95
Visual-match	4.92	7.16
Imagery		
Rhyme-match	5.29	6.90
Visual-match	5.29	7.45

Note. Maximum of 8 words per list and 20 target letters per list.

we asked whether imagery and verbal rehearsal tap different processing resources. The results of these experiments supported this hypothesis. Items that were encoded in the same mode tended to be recalled together, and rehearsed items were recalled prior to those that were imaged. Immediate recall was maximized when different portions of a list were maintained by different modes of representation, providing imagery encoding occurred first.

The two later experiments then investigated the nature of the processes underlying the effects of imagery on verbal recall. We asked whether the differences in processing observed previously were indicative of modality-specific perceptual processing mechanisms being called into play. To answer this question we attempted to find selective interference effects that would reveal a link between imagery and like-modality perception. We consistently found that a task requiring acoustic processing selectively interferes with recall of verbally rehearsed words; however, we did not find any evidence that a visual-search task selectively reduces recall of imaged words.[1] The conclusion best supported by these results is that the processes leading to greater recall under imagery instructions are not specifically linked to either acoustic or visual processing; rather, image formation may improve recall by increasing semantic elaboration (Anderson & Bower, 1973; Bransford & McCarrell, in press).

This conclusion can, of course, only be tentative because it rests on the acceptance of the null hypothesis with regard to selective interference effects. Nevertheless, increasing the difficulty of the interference tasks (Experiments 4 vs. 3) did not alter the pattern of recall in any way. And the present study adds to a number of others that have also been unable to find mutual interference between visual perception and the recall of words encoded imaginally (Elliott, 1973; Bower, Munoz, & Arnold, Note 1; Brooks, Note 2). In addition, the finding in Experiment 2 that imagery produces superior recall to verbal rehearsal only when long-term memory is tapped is also consistent with a semantic interpretation of imaginal encoding (see Mazuryk, 1974).

One must be very cautious, however, in generalizing the present conclusion. Though we have found no evidence that the perceptual qualities of imagery are central to the recall of unordered word lists, this certainly does not mean that such perceptual qualities do not exist; they may very well exist even if they do not play a direct role in verbal recall. Furthermore, such perceptual qualities of imagery may directly influence memory in other experimental paradigms, such as memory for details of pictures (Kosslyn, 1973; Marks, 1973) and the serial recall of spatially organized material (Brooks, 1968; Byrne, 1974). A full understanding of the processes involved in imaginal encoding depends on the investigation of both amodal and modality-specific properties of the internal representation.

---

[1] In a follow-up experiment, we also found no evidence of differential interference effects when we compared lists of items learned by forming images with other lists learned by forming a story to connect the words.

REFERENCE NOTES

1. Bower, G. H., Munoz, R., & Arnold, P. G. *On distinguishing semantic and imaginal mnemonics.* Unpublished manuscript, Stanford University, 1972.
2. Brooks, L. R. *Visual and verbal processes in internal representation.* Paper presented at Salk Institute, La Jolla, California, July, 1970.

REFERENCES

Anderson, J. R., & Bower, G. H. *Human associative memory.* Washington, D. C.: V. H. Winston & Sons, 1973.

Atwood, G. E. An experimental study of visual information and memory. *Cognitive Psychology,* 1971, *2,* 290–299.

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology,* 1953, *49,* 229–240.

Bower, G. H. Mental imagery and associative learning. In L. Gregg (Ed.), *Cognition in learning and memory.* New York: Wiley, 1972.

Bransford, J. D., & McCarrell, W. S. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In D. Palermo & W. Weimer (Eds.), *Cognition and the symbolic processes.* Washington, D. C.: Winston-Wiley, in press.

Brooks, L. R. The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology,* 1967, *19,* 289–299.

Brooks, L. R. Spatial and verbal components of the act of recall. *Canadian Journal of Psychology,* 1968, *22,* 349–368.

Byrne, B. Item concreteness vs. spatial organization as predictors of visual imagery. *Memory and Cognition,* 1974, *2,* 53–59.

Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* 1973, *12,* 335–359.

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.

Elliott, L. Imagery versus repetition encoding in short- and long-term memory. *Journal of Experimental Psychology,* 1973, *100,* 270–276.

Kosslyn, S. M. Scanning visual images: Some structural implications. *Perception & Psychophysics,* 1973, *14,* 90–94.

Marks, D. F. Visual imagery differences and eye movements in the recall of pictures. *Perception & Psychophysics,* 1973, *14,* 407–412.

Mazuryk, G. F. Positive recency in final free recall. *Journal of Experimental Psychology,* 1974, *103,* 812–814.

Paivio, A. *Imagery and verbal processes.* New York: Holt, Rinehart & Winston, 1971.

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

Pylyshyn, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin,* 1973, *80,* 1–24.

Segal, S. J., & Fusella, V. Influence of imaged pictures and sounds on detection of visual and auditory signals. *Journal of Experimental Psychology,* 1970, *83,* 458–474.

Smith, E. E., Barresi, J., & Gross, A. E. Imaginal versus verbal coding and the primary–secondary memory distinction. *Journal of Verbal Learning and Verbal Behavior,* 1971, *10,* 597–603.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1971.

(Received May 27, 1975; revision received August 13, 1975)