Pictures and words in visual search*

ALLAN PAIVIO

University of Western Ontario, London, Ontario, Canada N6A 3K7

and

IAN BEGG

McMaster University, Hamilton, Ontario, Canada

Ss in three experiments searched through an array of pictures or words for a target item that had been presented as a picture or a word. In Experiments I and II, the pictures were line drawings of familiar objects and the words were their printed labels; in Experiment III, the stimuli were photographs of the faces of famous people and their corresponding printed names. Search times in Experiments I and II were consistently faster when the array items were pictures than when they were words, regardless of the mode of the target items. Search was also faster with pictures than with words as targets when the search array also consisted of pictures, but target mode had no consistent effect with words as array items. Experiment III yielded a completely different pattern of results: Search time with names as targets and faces as search array items was significantly slower than in the other three conditions, which did not differ from each other. Considered in relation to several theories, the results are most consistent with a dual-coding interpretation. That is, items that are cognitively represented both verbally and as nonverbal images can be searched and compared in either mode, depending on the demands of the task. The mode actually used depends on whether the search must be conducted through an array of pictures or words.

The study investigated the speed of visual search when pictures and their verbal labels served both as target stimuli and as arrays through which the search was conducted. A major focus of the investigation concerned the asymmetrical picture-word comparisons because a successful search in these instances requires a cognitive transformation of the to-be-compared items into a common mode of representation. Formally, this common mode could be verbal, requiring implicit labeling of pictures; pictorial, requiring image generation to words; or some abstract form of representation underlying both pictures and words. Each alternative has received emphasis in different theories of the mediation processes involved in memory and perception, so that different predictions arise from the alternatives.

Verbal coding has been emphasized as a central aspect of perceptual processing by various theorists (e.g., Bruner, 1957; Glanzer & Clark, 1963; Haber, 1966). A strong verbal position would assert that pictures are identified essentially by naming them. Since words can be read faster than pictures can be labeled (e.g., Fraisse, 1968), word arrays should be more quickly searched than picture arrays, and words should be more effective targets than pictures. The predicted order of conditions is, thus, that the word-target word-array condition will be fastest, followed in order by picture-word, word-picture, and picture-picture comparisons.

In contrast, a strong imagery hypothesis would assert

that the target item, whether pictorial or verbal, is stored most effectively as a nonverbal image. The search task would involve a comparison of that image with the pictures in a pictorial array or with the images of the words in a verbal array. Since it takes longer to generate images to words (a second or more in different studies, see Paivio & Begg, 1971) than to recognize pictures perceptually (a mean of 0.73 sec was reported by Ernest, 1972, Experiment 6), the predictions are the reverse of the predictions arising from the strong verbal hypothesis. From fastest to slowest, the four target-array conditions should be picture-picture, word-picture, picture-word, and word-word.

A combination of verbal and imaginal coding is a third possibility following from a general dual-coding approach to cognition (Paivio, 1971) and more specific dual-coding analyses of performance in comparison tasks (e.g., Posner, Boies, Eichelman, & Taylor, 1969). The essence of the argument is that Ss will use either imaginal or verbal coding, depending on expectations aroused by contextual information in the experimental setting. Thus, the mode of the comparison stimuli encountered on previous trials (Posner et al, 1969; Tversky, 1969), instructions concerning the mode of the comparison stimuli, etc., may predispose the S to generate a visual image to a target item if the comparison stimulus is basically a visual entity, such as a picture, or vice versa if it is a verbal entity, such as a word. In the experiments described in this paper, the contextual information was provided by the mode of the search array as a whole (i.e., a matrix of pictures or words) and predictions concerning search time for the experimental conditions hinge on assumptions about the relative efficiency of a visual search through pictures as

*This research was supported by grants from the National Research Council of Canada (A0087 and A8122) and the University of Western Ontario Research Fund. We are grateful to Ann Anas for her assistance with the experiments. Reprint requests should be sent to A. Paivio, Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 3K7.

compared to a visual search through words. This must also depend on the relative complexity and discriminability of the stimuli involved.

Where the comparison involves simple line drawings and their printed names, a search through pictures might be favored because differentiation among items and identification of a particular item are possible on the basis of relatively abstract visual features which to some extent could be detected peripherally as well as foveally. Each word, however, is a unique configuration of elements shared by other words and identification would require a more detailed inspection involving foveal fixation of items. Thus, pictures could be discriminated more readily and perhaps searched with a broader visual "sweep" (more items searched in parallel), with more rapid exclusion of items with inappropriate features than would be possible with printed words. The latter would more likely involve an item by item serial search. To this point in the argument, then, we have generated the prediction that picture arrays will be more quickly searched than word arrays, regardless of the mode in which the target is presented.

It is more difficult to generate predictions from the dual-coding hypothesis concerning effects of target mode, especially in the tasks used here, where Ss were given ample time to encode the stimuli in any way they chose prior to presentation of the search array. However, the picture-picture condition should have a small but consistent advantage over the word-picture condition, simply because Ss are more likely to generate variable images with words than with pictures as stimuli. Any lack of correspondence between the S-generated image of a word and the E-provided picture of which the word is a label must have the effect of increasing search time. The same reasoning, however, does not hold with word arrays, because the pictorial stimuli were chosen to maximize the commonality of the verbal labels. The prediction arising from the dual-coding hypothesis is, thus, that the search times will increase from picture-picture to word-picture and then to the other two conditions, which are not expected to differ. This predicted order is very different from the verbal encoding hypothesis but similar to the strong imagery hypothesis.

The fourth hypothesis states that both pictures and words must be transformed into some form of abstract representation before a match is made. Such an interpretation would be consistent with Osgood's (1968) analysis of meaning as a representational process consisting of a simultaneous distinctive bundle of fractional responses (of course, Osgood would not necessarily argue that stimulus comparison in the search task would take place at this abstract level of meaning). It is represented more specifically by Chase and Clark's (1972) hypothesis that pictures and their descriptive sentences are processed by being transformed into underlying semantic entities called propositions. Such a hypothesis obviously would predict no differences between any of the search conditions involved in the

present study, unless an additional assumption is made to the effect that the third entity is differentially accessible to words and pictures—in which case the hypothesis becomes formally similar to the verbal or the image coding theory, depending on one's assumption.

The problem has not been previously investigated as a search task, but Rosenfeld (1967) conducted a study that is highly relevant to the general problem. He measured the same-different comparison time between successive stimuli, which were either figures or their labels (e.g., three red triangles). In a given session, the second stimulus was always either verbal or pictorial, while the first stimulus could be either. All four combinations were used, with or without a delay between the successive stimuli. If we consider first the delay condition, which is most similar to the present experiments, predictions can be made. An abstract entity hypothesis assumed that both the pictures and the words are translated into a common third entity. Assume that the delay interval is long enough to translate the first stimulus. Then, if both types of stimuli are encoded into the third entity with equal rapidity, the prediction is that the null hypothesis is true. If, however, the third entity is more directly accessible for the pictures than for the words, the prediction is that figures will be faster comparison stimuli than words; the reverse prediction is made if the third entity is more accessible for words than for pictures.

Rosenfeld (1967) found that, for picture and word target stimuli, respectively, it was faster to make the comparison with pictures (656 and 675 msec) than with words (828 and 811 msec). With each type of comparison stimulus, the decision was faster if the successive stimuli were of the same type, but the differences were small. The order of conditions was picture-picture, word-picture, word-word, and picture-word. These results are most consistent with dual coding or with an abstract coding hypothesis in which it is assumed that the third entity is more accessible for pictures than words. Similar results have been reported more recently by Seymour (1973).

Reported here are three experiments that bear directly on the problem. The first experiment involved pictures and printed names of familiar objects as stimuli. The second was identical in design to the first, but the number of items in the target stimulus was varied. Likewise, the third experiment was identical in design to the others, but the pictured faces and names of famous people were used as stimuli. In each experiment, Ss were presented with either a pictorial or verbal target set and were allowed time to encode the stimulus. Then either a picture or word matrix was presented, and Ss searched for the appropriate matching item.

EXPERIMENT I

Method

Materials. One hundred pictures and their corresponding labels were selected from a pool of 260 items. The pictures consist of

black and white line drawings of common objects. Normative data, obtained from a group of 122 Ss, are available on the commonality or consistency of the picture labels. All of the selected pictures were given the same name by a high percentage (between 90% and 100%) of the Ss. The words used were the most common labels for the pictures, printed in uppercase executive type. Four 5-item by 5-item square matrices of the labels were constructed so that none of the 25 items in one matrix appeared in another. Word length was controlled so that short and long words were distributed about equally throughout the four matrices. Four additional arrays were constructed by rearranging the items in each of the original matrices. Each of the word arrays had a corresponding picture matrix, giving a total of 16 matrices. Two-inch-square duplicates of the selected items (both pictures and words) were then arranged to form eight of the above matrices and photographed to produce a separate 10 x 10 in. print of each array. Since each S was to be presented each array four times (see Procedure), with a different target each time for a total of 64 trials, four xeroxed copies were made of each matrix. Thus, one matrix was available for each trial. Each of these arrays was then pasted onto a 12 x 12 in. sheet of black Mayfair cover.

The target stimuli were then selected so that they occupied 16 different positions in the search matrices throughout the experiment, with the two top and bottom rows and the two left and right columns being sampled about equally often. The upper left corner, bottom right corner, and the center square were never used as target locations. The positions were assigned randomly to the four copies of each of the four word arrays, resulting in 16 arrays, each with the word target appearing in a different position. The same positions were used for target items in the 16 rearranged word arrays and in the corresponding picture arrays, with the proviso that each target item was used only once throughout the experiment as either a picture or a word target. Thus, 64 different items served as targets throughout the experiment for any given S. Half of those occurred as pictures and half as words in both the target set and the search array, in the picture-picture, picture-word, word-picture, word-word combinations called for by the factorial design. In all, four random sequences of these 64 target-array combinations were constructed, and four additional sequences resulted from the reversal of these random orders. A Latin-square design was applied to the first four target-array combinations, each of which represented one of the four conditions. For presentation each target stimulus was pasted onto a piece of 5 x 4 in. black Mayfair cover.

Procedure. The S was seated opposite E at a small table. Detailed instructions read to the S included a description of the four combinations of target mode and search array mode, and S was urged to be prepared for any of these combinations on any trial. This presumably encouraged Ss to store targets as both words and images or primed him to transform the memory target (if necessary) when the array was exposed. The 64 targets were all placed face down on S's left and the 64 arrays face down on his right. The S was instructed to look carefully at each target item as the E presented it and to say "ready" as soon as he recognized what it was. He was also asked to close his eyes immediately afterwards. While his eyes were closed, the E placed the appropriate matrix face up directly in front of the S and said "ready." This was the signal for the S to begin searching for the target item he had just seen. The search ended with the S pointing to the correct target item in the matrix. He then closed his eyes again, the E said "next item," and then presented the next target stimulus. This procedure was repeated for each of the 64 target-array combinations.

The search time was the interval between the "ready" signal and the moment when the S pointed to the target item. The timing was done by a stopwatch.

Subjects. The Ss were 20 introductory psychology students at the University of Western Ontario who participated as part of a course requirement.

Table 1
Mean RTs (Seconds) for Visual Search Through Picture or Word Arrays for Picture or Word Targets

Array		Target	
		Picture	Word
Picture	Mean	2.89	3.15
	SD	0.57	0.47
Word	Mean	4.18	4.46
	SD	0.47	0.77

Results and Discussion

Mean response latencies were determined for the four combinations of target and search array modes for each S by averaging over the 16 observations in each condition. Three observations (one from the picture-word and two from the word-picture conditions) were omitted from the calculation of means because S had selected the wrong item from the matrix. The overall mean latencies for each condition are presented in Table 1. The results show an orderly pattern for the four target-array combinations, with search time being fastest for picture-picture and progressively slower for word-picture, picture-word, and word-word. Thus, pictures facilitated search as targets and even more so as search array items.

The reliability of the pattern was confirmed by a 2 by 2 analysis of variance, which revealed significant main effects of target [$F(1,19) = 4.53, p < .05$] and search array [$F(1,19) = 100, p < .001$]. The interaction was not significant ($F = .007$).

These results are precisely in agreement with predictions from the strong imagery hypothesis. According to this theory, the target item is more effectively stored as a nonverbal image than as a word, and the search process is most efficient when the memory image can be compared directly with pictorial items in the array. The results clearly disconfirm the strong verbal hypothesis, which predicts exactly the reverse of the pattern that was actually observed. For example, if the search had involved an item by item comparison between the verbal labels of target items held in memory and each item in the array, search time should clearly be longer for the word-picture than for the picture-word combination, contrary to what was obtained. The results are also quite inconsistent with the abstract entity hypothesis, which does not make directional predictions. The imagery theory is also favored over dual coding, since the latter predicts search time should not differ for word-word and picture-word conditions.

The findings are generally consistent with Rosenfeld's (1967) data for the picture-word comparison task in the delay condition, with the main discrepancy being that Rosenfeld found the word-word condition to be faster than the picture-word condition. Rosenfeld interpreted this to mean that, in the word-word condition, words were simply compared as visual word forms rather than being decoded into their figural representations. Such a

Table 2
Mean Median RTs (Seconds) and SDs for Visual Search Through
Picture or Word Arrays for Picture or Word Targets of Sizes 1-3

Array	Picture Target		Word Target	
	Mean	SD	Mean	SD
Single Target				
Picture	2.33	0.41	2.52	0.38
Word	3.43	0.84	3.79	0.85
Double Target				
Picture	3.30	1.01	3.54	0.83
Word	4.56	0.98	4.18	0.79
Triple Target				
Picture	4.15	1.15	4.69	1.07
Word	5.50	1.20	5.65	1.05

possibility was anticipated in the present experiment as well, but it apparently did not occur. The discrepancy could be due to differences in item discriminability or memory load, or both, in the two experiments. Rosenfeld's experiment involved a same-different comparison task between a single target item and a single comparison stimulus that followed either immediately or after a delay of 2.5 sec. Item discrimination was not a problem and a visual trace of the printed word presumably could be held in memory for that interval, permitting a direct visual match. In the present study, however, Ss were required to discriminate among 25 printed words and hold a target item in memory for as long as 15 sec in order to perform a successful match. Perhaps visual memory for the orthographic features of a word had faded by that time and the S was forced to rely either on a name code or a visual image of the referent of the target item in order to retain it in memory during the search process. In either case, the essential features of the target item were stored less effectively for purposes of the matching task than when the target item was presented as a picture. Whether the coding involved in the word-word condition was imaginal or verbal in the present study remains in doubt, but the differential discrimination and memory load analysis resolves the apparent inconsistency between the two sets of results, at least in principle.

The analysis also suggests, however, that a firm choice between the imagery and dual-coding hypothesis is not possible on the basis of these results. That is, imaginal coding might have been involved whenever pictures were involved as targets or search items, whereas verbal coding was involved in the word-word condition, and search was hindered simply because the target was less effectively stored or because verbal coding of the target and of the search items created response competition that took up extra time (cf. Brooks, 1967). Here, too, a definitive answer requires further research, but it remains clear that the results cannot be readily explained without relying heavily, if not completely, on imagery as an explanatory concept.

EXPERIMENT II

The second experiment was identical in design to the first, except that the targets consisted of one, two, or three pictures or words. In the case of multiple targets, however, only one of the target items appeared in the search array. Since S did not know which target item was to appear, all targets on a given trial presumably had to be held in memory during the search. The main purpose of the experiment was simply to examine the replicability and generality of the results obtained in the first experiment.

Method

Materials and Procedure. Precisely the same search matrices as in the first experiment were used again, with 64 trials for each S. In the case of the double and triple targets, additional pictures and their labels were chosen from the previously mentioned pool, although the target stimulus which had a match in the search array was always the same in the different target conditions.

Likewise, the procedure was identical to the first experiment.

Subjects. Forty-eight student volunteers from McMaster University were paid \$2/h for their participation, with 16 Ss in each level of the target-size variable and 2 Ss in each of eight random orders of presentation.

Results and Discussion

Because a few of the search times were very long, some as high as 40 sec, median rather than mean response latencies were determined for the four combinations of target and search array modes for each S. There were 16 observations for each S in each condition. A total of 24 responses of the 3,072 made were discarded because of failure to make a correct identification (16 of the errors were from the triple-target condition, with the highest error rate coming in the picture-target word-array condition with a triple target, with seven errors in 256 trials). The overall means of the medians for each condition are presented in Table 2.

The single-target condition was analyzed first by a 2 by 2 analysis of variance with target and array mode as repeated factors, since it was a direct replication of the first experiment, involving a different E and a different population of Ss. Both the main effect of target [$F(1,15) = 4.90, p < .05$] and array [$F(1,15) = 50.9, p < .001$] were as in the first experiment, with latencies for the different conditions increasing in the order picture-target/picture-array, word-picture, picture-word, and word-word. The interaction was not significant ($F = 0.35$). Thus, the pattern of results obtained in the first experiment was clearly replicated.

The overall analysis of the data was a 3 by 2 by 2 analysis of variance with number of targets as an independent factor and target and array mode as repeated factors. The main effect of number of targets [$F(2,45) = 35.3, p < .001$] simply reflected increasing times as target number increased, with respective times of 3.02, 3.89, and 5.00 sec. Picture arrays were searched

faster than word arrays [$F(1,45) = 65.4, p < .001$]. Finally, target number and target mode interacted [$F(2,45) = 4.32, p < .05$], since picture targets led to shorter latencies than word targets in the single-target condition (2.88 vs 3.16 sec) and the triple-target condition (4.54 vs 5.17 sec) but not in the double-target condition (3.93 vs 3.86 sec). In each condition picture-picture searches were fastest, with word-picture next. In the single- and triple-target conditions, word-word searches were slowest, while in the double-target condition picture-word searches were slowest.

Before drawing conclusions from the two experiments, it should be noted that we conducted another experiment with target set size of one and two, the same four combinations of pictures and word targets and arrays, and with six consecutive searches through each array. Analysis of the experiment revealed a main effect of target set size ($p < .05$), a main effect of both target mode ($p < .05$) and array mode ($p < .001$) as found in the other experiments, and a Target by Array Mode interaction ($p < .001$), reflecting the order of conditions, from fastest to slowest, as picture-picture, word-picture, word-word, and picture-word. Note in particular that the order of the two slowest conditions is reversed relative to what was generally observed in Experiments I and II. The experiment is not reported in detail here because it does not provide enough new information, but the results will be taken into account in drawing conclusions.

First, picture arrays are consistently associated with faster times than word arrays, regardless of the mode in which the target stimulus was presented. The result can be interpreted to mean that pictures can be scanned in parallel, while word arrays require a sequential word by word search. Second, pictures are consistently better targets than words when the array to be searched is a matrix of pictures. This could simply reflect the greater variability of images generated by words as opposed to pictures. However, neither pictures nor words have a consistent advantage as targets when searches are conducted through matrices of words.

The asymmetrical comparisons in the case of the word-picture and picture-word conditions were consistently in favor of the former condition, as expected by both the strong imagery and dual-coding hypotheses. In general, then, the results are quite consistent with the dual-coding hypothesis, although the strong imagery hypothesis cannot be ruled out.

In order to make a stronger test of the two hypotheses, a third experiment was conducted.

EXPERIMENT III

The basic problem was investigated in Experiment III using names and faces of famous people as items. The design was generally identical to that used in Experiment I, with the difference that a

prefamiliarization condition was added because the faces and names were not as uniformly familiar to the sample of Ss involved in the experiment as were the object drawings used previously. Moreover, different items were unfamiliar to different Ss, so that a common pool of easily named faces was difficult to obtain. The problem was solved by basing the analysis only on search trials for which Ss indicated that they knew the target name or face.

The strong imagery hypothesis would make identical predictions to those made for the previous experiments. That is, face-face searches would be fastest, followed by name-face, face-name, and name-name in order. The dual-coding hypothesis relied on the assumption of parallel scan to predict a main effect of picture arrays. However, faces, like names, may require foveal fixation for specific identification, in which case the array effect obtained in the previous experiments should not occur in this experiment. However, the face-face comparison should still be faster than the name-face comparison, because of encoding variability in the images generated to the names. Such variability would not be expected in the case of verbal coding of face targets in the face-name condition. Thus, the general dual-coding prediction is that search time will be slowest for the name-face condition and essentially equal for the other three conditions.

Method

Materials. The basic material again consisted of a series of target items and search arrays. Each target item was a photograph of the face or the full name of a well known public personality. Each search array was a 5 by 5 matrix of faces or surnames (first names were excluded from the search array) which included the target item for a given trial.

The target items were chosen in the following manner. Fifty undergraduate psychology students were required to write the names of famous people whose faces they could easily picture. They were given response sheets containing 13 categories of well known public figures and personalities (e.g., comedians, sports figures, monarchs, etc.) to use as guidelines for the task. For each category they were given 1 min to write as many names as they could of people whose faces easily came to mind. The number of instances of each name was calculated and the 32 most frequently named male personalities were chosen as target stimuli for the study. They were Pierre Berton, Humphrey Bogart, Richard Burton, Fidel Castro, Winston Churchill, Gary Cooper, William Davis, John Diefenbaker, Albert Einstein, D. D. Eisenhower, Clark Gable, Dustin Hoffman, Bob Hope, Gordie Howe, Bobby Hull, L. B. Johnson, J. F. Kennedy, Nikita Krushchev, Abraham Lincoln, Dean Martin, Mao Tse-Tung, Paul McCartney, Richard Nixon, Lester Pearson, Prince Philip, John Robarts, Frank Sinatra, Robert Stanfield, Pierre Trudeau, Dick Van Dyke, John Wayne, and Flip Wilson. Full-face photographs of these people and of 68 lesser known individuals (to serve as filler items in the search arrays) were obtained from various photograph and portrait libraries in the city to bring the total number of items to be used in the study to 100. These were rephotographed to produce 2 x 2 in. prints, which were then arranged into four separate 5-item by 5-item matrices so that none of the 25 items in one appeared in another.

The 32 target items were distributed in such a manner that each of the four basic matrices contained eight different target stimuli, with none of the targets in one matrix appearing in

Table 3
Means and SDs of Search Time (Seconds) for Faces and Names in Face and Name Arrays for Items With and Without Pretraining

Array	Face Target		Name Target	
	Mean	SD	Mean	SD
No Pretraining				
Face	3.94	0.82	4.99	1.30
Name	3.71	1.25	3.75	0.98
Pretraining				
Face	4.40	1.14	4.87	1.02
Name	4.04	0.97	4.11	0.80

another. (The selection of their positions is described below.) Four more arrays were constructed by randomizing the items in each of the original matrices. Each of these eight matrices was photographed and two 10 x 10 in. copies were made of each. These 16 prints were pasted on 12 x 12 in. sheets of black Mayfair cover. An additional 16 matrices were then made using the last names of the personalities that appeared in the face arrays. The names were arranged in the same orders as the eight face arrays on 10 x 10 in. plain white paper, which had been divided into 2-in. squares. Two xerox copies were made of each of the typewritten arrays and these copies were also pasted onto 12 x 12 in. black Mayfair cover. This brought the total number of search matrices to 32, with a face or name available as a target for each matrix.

The positions of the target stimuli in the arrays had been selected before the construction of the arrays in order to place the target stimuli in the appropriate positions when the arrays were being constructed. Eight different positions were chosen, two different ones for each of the four original face arrays. In this way, each area of the matrix was sampled throughout the experiment as in Experiment I: Positions 1 and 2 appeared in Copies 1 and 2, respectively, of Face Matrix A, Positions 3 and 4 in Copies 1 and 2 of Matrix B, Positions 5 and 6 in Copies 1 and 2 of Matrix C, and Positions 7 and 8 in Copies 1 and 2 of Matrix D. The same positions were also distributed in the same way throughout the four rearranged face matrices. With the name arrays, however, the allocation of the eight positions to the arrays was reversed, i.e., Positions 1 and 2 appeared in Copies 1 and 2 of Name Matrix D, Positions 3 and 4 in Name Matrix C, etc. This insured that the targets differed for the name and face arrays. Thus, each S would receive 32 different target-array combinations, 8 of which would be face-target/face-array, 8 face-target/name-array, 8 name-target/face-array, and 8 name-target/name-array.

Each of the 32 target stimuli was available both as a name typed on a 2 x 2 in. square of white paper and as a 2 x 2 in. photograph of the personality, each pasted onto a 4 x 5 in. sheet of black Mayfair cover. Since the search arrays also could appear as names or as faces, any given target-array combination could be an observation in any one of the four conditions.

Two random orders were made of the 32 target-array combinations each S was to receive. Two additional orders were provided by reversing the original two. For each of these four orders, four new presentation conditions were provided by insuring that with four Ss each of the 32 target-array combinations had appeared once in each of the four face-face, face-name, name-face, name-name conditions.

Experimental Procedure and Subjects. Each S received all 32 different target-array combinations. For the first seven Ss, the procedure was exactly the same as in Experiment I. The S sat opposite the E at a small table. The 32 targets and matrices were face down on the S's left and right, respectively. The four target-array combinations were described in detail and S was urged to be prepared for any one of them. He was instructed to look carefully at each target item and to say "ready" as soon as

he knew the person's name, if presented his photograph, or could easily picture the person's face, if presented his name. He was instructed to tell the E if he was not sure who the person was. As in Experiment I, he was asked to close his eyes as soon as he was sure of the target stimulus and wait for the E to say "ready," which was the signal for S to begin searching the arrays for the target and E to begin timing with the stopwatch. As soon as the S pointed to the correct item in the matrix, the watch was stopped and he was told to close his eyes again. The matrix was then turned face down and the time of search was recorded. The results obtained from these seven Ss were examined to see in how many instances they were unsure of the identity of the target person. Each S missed an average of three to four items. It was accordingly decided that another group of Ss would be given pretraining with the target items. Prior to the experimental procedure, they were shown the faces of the target personalities, one at a time, and asked to give the appropriate name. The names were then similarly presented and the S was allowed to study each face and name together for about 5 sec. Following this, the regular procedure described above was carried out. Seventeen Ss were tested under the pretraining condition. An additional 10 Ss were then tested without pretraining to equate the Ns in the two groups.

The Ss were 34 introductory psychology students from the University of Western Ontario who participated as part of a course requirement.

Results and Discussion

Mean search latencies were computed for each S for each of the four target-array combinations. Each cell of the factorial design was represented by approximately the same total number of observations, ranging from 116 to 130 for the different conditions. Departures from equality resulted only from the fact that trials on which the S could not identify the target item were not included in the analysis. The means for each cell of the design are presented in Table 3. These data were analyzed by a 2 by 2 by 2 repeated-measure analysis of variance, with target mode, search mode, and training condition (pretraining vs no pretraining) as factors.

The results revealed that the training variable had no effect approaching significance either independently or in interaction with the other factors ($F_s \leq 1.43$). Significant effects were obtained for both target and search mode: Search time was faster with pictures than with names as targets [$F(1,32) = 9.52, p < .01$] and faster with names than with pictures as search array items [$F(1,32) = 11.3, p < .01$]. The finding for target mode is consistent with Experiments I and II. The inferiority of pictures as search array items, however, is contrary to the other experiments.

The specific prediction from the dual-coding hypothesis was that search would be slowest for the name-picture condition and about equal for the remaining conditions. This would be reflected in the above analysis as an interaction of Target by Search Mode, which approached significance [$F(1,32) = 3.71, p < .10$]. The prediction justified further individual comparisons among the four cell means (Winer, 1962, p. 208). This was done by a Newman-Keuls test following a one-way analysis of variance which showed that the four means differed significantly [$F(3,102) =$

8.16, $p < .001$]. The pairwise comparisons revealed that search time for the name-picture condition significantly exceeded the times for each of the other three conditions ($p < .01$), which did not differ significantly among themselves. Thus, the results are precisely as expected from the dual-coding hypothesis, given the initial assumptions that faces, like names, would require foveal fixation to be specifically identified and that variability in the images generated to name targets would retard the comparison process in the name-face condition.

GENERAL DISCUSSION

The overall results of the three experiments are most consistent with a dual-coding theory. Items that are represented cognitively both verbally and as nonverbal images can be searched and compared in either mode, depending on the demands of the task. This is consistent also with the theorizing and results presented by Posner et al (1969) and Tversky (1969). The results of Experiments I and II are also consistent with Rosenfeld's (1967) strong decoding (or imagery) hypothesis, but that hypothesis would incorrectly predict the same results for Experiment III. The findings appear to be quite inconsistent with any hypothesis that would emphasize either verbal coding or abstract representations as the sole basis of the search and comparison processes.

Although the abstract entity hypothesis is insufficient to account for the results, some kind of abstraction presumably occurred in the search task. It was central to the theoretical argument that line drawings and the imaginal representations corresponding to them are schematic in nature, with certain general features being essential to the core representation (or representations) corresponding to a particular generic object such as a cross. These features are by definition relatively simple, being abstracted from the detailed features of numerous specific exemplars of the generic objects. Faces, however, derive their identity from a more complex combination of features that is unique to each individual (cf. Scapinello & Yarmey, 1970). In this respect they resemble words, which also involve a unique combination of orthographic (or phonemic) features. The nature of the structures also differs, with faces involving a hierarchical spatial pattern or nested set (e.g., nostrils within a nose within a face), whereas words involve a sequential or serial structure of letters or phonemic units, but this is not the most relevant point here. What is relevant is the assertion that visual structures and their corresponding imaginal representations, whatever their nature, can vary in abstractness and simplicity (cf. Attneave, 1954; Paivio, 1971, Chapter 2; Posner, 1969).

The above analysis has several implications in regard to the present experiments. First, S should find it easier to generate and retain the more schematic visual images aroused by object drawings or their labels than the more detailed images of specific faces or of words as visual patterns. Second, it might be easier for the S to compare a schematic memory image of an object directly with the object pictures in the search array than to do so in the case of the more complex faces or words. Third, the drawings might be advantageous during the search process because they can be discriminated from each other and identified more quickly than faces or words, even if the S engages in an item by item serial search. Alternatively, if a parallel search process is assumed (cf. Neisser, 1967), schematic pictures might be identified more peripherally, permitting a broader visual scan of the search array than is possible in the case of faces and words. These and other possibilities remain to be explored.

REFERENCES

- Attneave, F. Some informational aspects of visual perception. *Psychological Review*, 1954, 61, 183-193.
- Brooks, L. R. The suppression of visualization in reading. *Quarterly Journal of Experimental Psychology*, 1967, 19, 289-299.
- Bruner, J. S. Neural mechanisms in perception. *Psychological Review*, 1957, 64, 340-358.
- Chase, W. G., & Clark, H. H. Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
- Ernest, C. Spatial imagery ability and the recognition of verbal and nonverbal stimuli. Unpublished PhD thesis, University of Western Ontario, 1972.
- Fraisse, P. Motor and verbal reaction times to words and drawings. *Psychonomic Science*, 1968, 12, 235-236.
- Glanzer, M., & Clark, W. H. Accuracy of perceptual recall: An analysis of organization. *Journal of Verbal Learning & Verbal Behavior*, 1963, 1, 289-299.
- Haber, R. N. Nature of the effect of set on perception. *Psychological Review*, 1966, 73, 335-351.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- Osgood, C. E. Toward a wedding of insufficiencies. In T. R. Dixon and D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
- Paivio, A. *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston, 1971.
- Paivio, A., & Begg, I. Imagery and comprehension latencies as a function of sentence concreteness and structure. *Perception & Psychophysics*, 1971, 10, 408-412.
- Posner, M. I. Abstraction and the process of recognition. In G. H. Bower and J. T. Spence (Eds.), *Advances in learning and motivation*. Vol. 3. New York: Academic Press, 1969.
- Posner, M. I., Boies, B. J., Eichelman, W. H., & Taylor, R. L. Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monograph*, 1969, 79(1, Part 2).
- Rosenfeld, J. B. Information processing: Encoding and decoding. Unpublished PhD thesis, Indiana University, 1967.
- Scapinello, K. R., & Yarmey, A. D. The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science*, 1970, 21, 329-331.
- Seymour, P. H. K. Pictorial coding of verbal descriptions. *Quarterly Journal of Experimental Psychology*, 1973, in press.
- Tversky, B. Pictorial and verbal encoding in a short-term memory task. *Perception & Psychophysics*, 1969, 6, 225-233.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received for publication October 24, 1973;
revision received January 7, 1974.)