THE VISUAL AND SEMANTIC COMPONENTS OF MEMORY FOR PICTURES

Ana Maria S. Bartolome, Melanie D.J. Caluma,
Melissa D.R. Felizardo, Ma. Concepcion I. Guevarra,
and Allan B.I. Bernardo
University of the Philippines

What do people remember about pictures? An experiment was conducted to determine whether people's memories of pictures include abstract representations of the pictures' meanings and visual representations of the pictorial information. Subjects were asked to study droodles which were either accompanied by a meaningful interpretation or not. Memory for the pictures was then tested using four different retrieval tasks. Each task was assumed to rely, in different specific ways, on the visual and semantic components of memory. The results are consistent with the notion that subjects have semantic and visual representations of pictures in memory, and are discussed in terms of the utility of having abstract propositional and concrete analogue representations in memory.

What do people remember when they remember paintings, pictures, and other complex forms of visual information? Why is it that everyone seems to remember what the painting of the Last Supper looks like, but, other than artists, art historians and art students, most people find it hard to distinguish one Jackson Pollock or one Mondrian painting from another? The difference between paintings like the Last Supper, on the one hand, and non-representational art, on the other, seems to be that people are able to capture the meaning of the former but not of the latter. This observation has prompted psychologists to propose that people do not remember pictures as exact visual objects. Instead, people remember some abstract representation that captures the meaning of the picture (Pylyshyn, 1973).

Several studies have found support for the position that people remember abstract representations of visual information. Wiseman and Neisser (1974), for example, showed their subjects a picture which appeared to be a random collection of dots and stains. The subjects did not see any meaning in the pictures and they showed very poor memories for these. However, when the subjects were told to detect the hidden figure in the pictures, they showed much better memory for the picture. Similarly, Bower, Karlin, and Dueck (1975) showed their subjects sets of droodles or line drawings that seem to have no meaning. Some subjects were given verbal labels which were meaningful interpretations of the droodles, while the rest were not. In a subsequent recall test, subjects who were given verbal labels showed much better recall than those who were not given the labels.

Evidence for the abstract representation argument include findings which show that subjects extract features of a picture that are important to its meaning, but not those details that are trivial and irrelevant. Mandler and Ritchey (1977), for example, had their subjects study pictures of scenes. One was a scene of a geography class, with a teacher, a student, a globe, a world map, and other classroom furniture. After viewing a set of these scenes, the subjects were tested for their memory for the pictures using a recognition task. The recognition task involved the studied scenes and two types of distractors. The token distractors were identical to the studied scene except for one minor detail. For example, the pattern of the teacher's skirt in the geography class scene was different. The type change distractors, on the other hand, were identical to the studied scene except for one important detail that changes the meaning of the scene; for example, the world map was changed to an art drawing which changes the subject of the class. Subjects were able to reject the token distractors only 60% of the time (which is just about chance levels) but were able to reject the type change distractors 94% of the time. The results are consistent with the argument that subjects remember an abstract representation of the pictures that extracts the physical elements of the picture related to the picture's meaning.

There is some evidence, however, that people seem to process images or representations of visual information in the same way perceptual information is processed. For example, in visual scanning experiments (Finke, 1980; Kosslyn, Ball & Reiser, 1978; Kosslyn & Pomerantz, 1977; Reed, Hock & Lockhead, 1983), the time it takes a subject to scan between two objects in an image was found to be a function of their distance from each other in the actual picture. If information about the location and distance between objects is represented only in propositional form, then scanning time between two objects should not be dependent on veridical distance. Furthermore, mental rotation experiments (e.g., Cooper & Shepard, 1973; Shepard & Metzler, 1971; Tarr & Pinker, 1989) indicate that when given pictures of patterns or objects in different orientations, subjects seem to manipulate images of these objects in order to recognize them. For example, the time it takes a subject to determine whether a pattern was a normal letter or its mirror image depends on how much the pattern was rotated from its upright position (Cooper & Shepard, 1973). These results could be better explained on the basis of visual representations in memory than on the basis of abstract representations.

Kosslyn & Pomerantz (1977) argued that people do have visual representations of visual information, but they are also interpreted and analyzed in a manner that characterizes abstract representations. This experiment attempts to show that memory for pictures has two components that correspond to the analogue visual representation and the abstract semantic representations of the visual information. The experimental procedures are based on Bower, Karlin, and Dueck's (1975) experiment with droodles.

Subjects were asked to study pairs of droodles like those shown in Figure 1. Half of the subjects were given verbal labels of the droodle pair, and the other half were not. The assumption was that subjects who were given verbal labels would represent both the visual and the semantic components of the picture. The other subjects would only have the visual representations.

The subjects' memories for these droodles were then tested using one of four different memory retrieval tasks: free recall, cued recall, correct cue recognition, and wrong cue recognition. In particular, the different tasks tested the subjects' memory for the picture on the right side of each pair presented in the study set. For the free recall task, the subjects were required to recall as many droodles on the right side of the pairs by drawing these in a blank sheet of paper. In the cued-recall task, the subjects were given the left droodle of the pair and they were asked to recall the corresponding droodle on the right

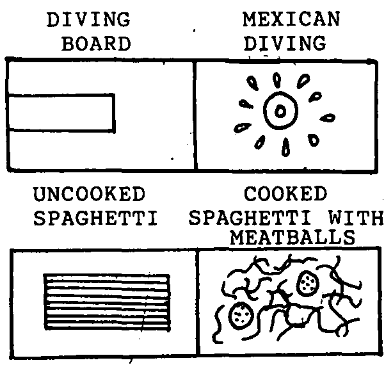

Figure 1. Examples of droodle pairs

side. Examples of these items are shown in Figure 2a. The recognition tasks consisted of 40 droodle pairs and the subjects were required to encircle those pairs in which the droodle on the right side appeared in the study set. In the right cue recognition, 20 of the 40 droodle pairs were exactly those which were presented in the study set, and 20 were completely new pairs. Examples of these items are shown in Figure 2b. For the wrong cue recognition the pairs, the target droodles on the right side were presented with left droodles other than that used in the study set, hence the term "wrong cue." Examples of these items are shown in Figure 2c.

The different retrieval tasks involve different modes of processing. Dyne, Humphreys, Bain, and Pike (1990), for example, argue that recall and recognition are fundamentally different tasks in the sense that recognition allows direct access to memory, whereas recall does not. In particular, according to Jacoby (1983), free recall tasks are conceptually driven retrieval tasks. In free recall tasks, subjects are given no cues to guide performance. Therefore, the subjects have to rely on stored concepts to facilitate remembering. Factors that have been identified as effective in aiding free recall usually involve generative or elaborative processing, and can therefore be viewed as facilitating conceptually driven processing.

Recognition tasks, on the other hand, are said to be more a combination of conceptually driven processing and data-driven processing (Jacoby & Dallas, 1981; Johnston, Dark & Jacoby, 1985; Mandler, 1980). In standard recognition tasks, aside from using stored concepts, subjects also have to rely on the perceptual record of the stimuli. To correctly recognize a previously experienced stimulus, subjects rely both on some semantic representation of the stimulus as well as on matching the surface features of the original stimulus with the items in the recognition set. Furthermore, factors that have shown to affect recognition are all related to qualities of the perceptual record of the infor-

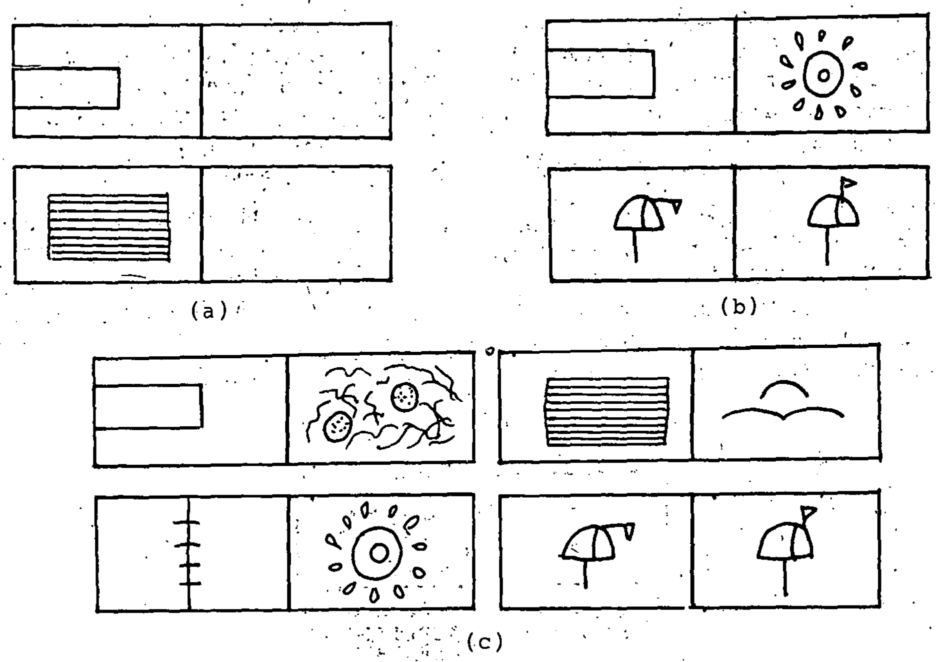

Figure 2. (a) Examples of test items for cued recall task. (b) Examples of test items for correct cue recognition task. First two items are correct targets and last two items are incorrect targets. (c) Examples of test items for wrong cue recognition task. First two items are correct targets and last two items are incorrect targets.

nation (e.g., persistence of the stimuli, surface similarity/dissimilarity, etc.).

The other two retrieval tasks differ from free recall and standard recognition in specified ways, as well. Cued recall tasks, like free recall, are also conceptually driven tasks. However, the cue could serve to facilitate or interfere with the conceptual processing depending on whether the cue is conceptually associated with the target information. The wrong cue recognition tasks, like standard recognition, are also both conceptually driven and data driven. However, the inappropriate cue should interfere with efficient retrieval of the semantic and visual information.

The rationale for varying the memory retrieval tasks is that for each retrieval task, subjects would have to rely on the visual and semantic components of their memory for pictures in different ways. We can, therefore, make very specific predictions about how subjects would perform in each of the retrieval tasks depending on the kind of information that was presented when the droodles were first studied. Generally, for example, recognition should be better than recall because subjects have both the visual and semantic components of their memory to use in recognition and only the semantic component in recall. Also, performance in all the retrieval tasks should be better when the droodles are presented with verbal labels than when they are not. The verbal label strengthens the semantic component of the memory for the droodles that should facilitate the conceptually driven processes in all the tasks.

If the droodles were presented without verbal labels, performance in free recall should be better than in cued recall, because the cue is not meaningfully associated to the target. Therefore, it will interfere with the retrieval process instead of facilitating it. However, if the droodles are presented with labels, the cues become meaningfully associated to the targets and should facilitate the conceptual processing. Therefore, performance in cued recall should be better than in free recall. A corollary of these two predictions is that the benefit of having verbal labels should be greater in the cued recall tasks than in the free recall tasks, because the cue becomes a useful conceptual cue to retrieval in cued recall.

Performance in the correct cue recognition task should also be better than in the wrong cue recognition task, because the inappropriate cue in the latter would interfere with the conceptually driven and data-driven processing. Having verbal labels associated with the droodles, however, could offset the effects of the wrong cue. This effect would imply that the benefit of having verbal labels should be greater for the wrong cue recognition than for the correct cue recognition.

To summarize, this experiment was designed to test whether people's memories for droodles contain both a visual representation of the visual information as well as semantic representation of the meaning of the visual information. The visual and semantic components of memory for pictures are tested by using four different memory retrieval tasks that tap, in different specified ways, these two components. The following hypotheses were tested:

(1) Memory for the droodles should be better if they are studied with verbal labels;
(2) Performance in recognition tasks should be better than in recall tasks;
(3) Performance in free recall should be better than in cued recall when the droodles are studied without verbal labels. The reverse should be true when the droodles are studied with verbal labels. The benefit of verbal labels should be greater for the cued recall task; and
(4) Performance in correct cue recognition should be better than in wrong cue recognition. The benefit of verbal labels should be greater for the wrong cue recognition task.

METHOD

Subjects

Subjects were 80 Introductory Psychology students from the University of the Philippines, Diliman, who participated in the experiment as part of a class requirement. They were randomly assigned to the label condition (N=40) or to the no label condition (N=40). The subjects in each group were further randomly assigned to one of the four retrieval tasks (for each labelling condition, N=10 for each task).

Materials

Twenty pairs of nonsensical pictures or "droodles" were presented to the subjects on 5 x 8 inch cards. Droodles are drawings that seem meaningless but turn out to have funny interpretations. The droodles in each pair are related in simple ways, as specified by their labels. Figure 1 shows some examples of droodle pairs with their corresponding labels.

The different retrieval tasks were presented in individual sheets of paper. For the free recall task, the test sheets contained 20 blank boxes. For the cued recall task, the test sheets contained 20 pairs of boxes, the left sides of which consisted of the 20 droodles shown on the left side in the study set. On the right side, the boxes were blank (see Figure 2a for examples). For the correct cue recognition task, the test sheet contained 40 droodle pairs, the 20 pairs shown in the study set and 20 new pairs (see Figure 2b for examples). The old and new pairs were arranged randomly in the test sheet. For the wrong cue recognition task, the test sheet also consisted of 40 droodle pairs. These pairs were divided as follows: (1) for 10 pairs, the right droodle appeared in the study, but the left side was a totally new droodle; (2) for 10 pairs, the right droodle appeared in the study set and the left droodle was also in the study set, but not the corresponding left droodle; (3) for 10 pairs, the right droodle was a new droodle and the left droodle was one that appeared in the study set paired with a different droodle; and (4) for 10 pairs, both the right and left side pictures were new (see Figure 2c for examples). The 20 pairs in sets (1) and (2) were the correct target items, and the rest in sets (3) and (4) were the incorrect targets. The correct and incorrect targets were arranged randomly in the test sheet.

A one page multiple choice test with ten trivia questions was also used as a distractor task.

Procedure

The subjects participated in the experiment in groups of five. They were informed that the experiment was about memory for pictures. They were told that they would be presented pairs of pictures for 10 seconds each and they were supposed to study these pairs. For the subjects in the label condition, the subjects were verbally given a meaningful interpretation for each droodle in the pair with the presentation of each pair. For the subject in the no label condition no interpretations accompanied the pictures. After the 20 pairs in the study set were presented, the subjects were given the distractor task. The subjects were given two minutes to complete the distractor task.

Immediately after the distractor task, they were tested for their memory for the study set. Subjects who were given the free recall task were instructed to draw the gist of as many of the pictures they remember to have appeared on the right side of the pictures they saw in the study set. They were told to draw their answers on the blank boxes in the sheet they were given. They were given seven minutes to complete the task. Subjects who were given the cued recall task were asked to do the same task as in the free recall task. However, the subjects were told that their test sheet included the drawings on the left side of the pairs. They should draw the correct picture that was earlier presented with the drawing on the left side. They were also given seven minutes to complete the task. Subjects in the correct cue recognition and wrong cue recognition tasks were told that their test sheet contained drawing pairs with old and new drawings. Their task was to choose and place a checkmark before each pair in which the right side picture appeared in the study set. They were given three minutes to complete the task. The answer sheets were collected after the times elapsed. The subjects were debriefed about the nature and the hypotheses of the experiment.

RESULTS

For the recall tasks, a drawing was coded as a correct recall if the drawing resembled and/or captured the gist of the original target droodles. The number of correctly recalled or recognized drawings was counted for each subject (perfect score = 20). For each condition, the mean number of correct answers was computed. The means are summarized in Table 1.

Table 1. Mean proportion of correct answers (and standard error) as a function of labelling and retrieval task

	No label	Label
Free recall	8.0 (1.9)	11.8 (1.8)
Cued recall	6.6 (2.3)	16.4 (2.1)
Correct cue recognition	19.1 (1.0)	19.5 (0.9)
Wrong cue recognition	13.3 (2.0)	17.1 (2.7)

These data were analyzed using a 2 x 4 (labelling x retrieval task) Analysis of Variance for completely randomized factorial designs. Consistent with the predictions, the analysis showed a main effect of labelling (F,1,72 = 106.08, MSe = 3.73, p<.0001). Subjects given verbal labels with the droodles showed better memory than those who were not. This result replicates Bower, Karlin, and Dueck's (1975) findings. Also as predicted, there was a main effect of retrieval task (F,3,72 = 93.91, MSe=3.73, p<.0001), and the pattern of results suggest that subjects showed better memory in the recognition tasks than in the recall tasks. There was also a reliable interaction between labelling and retrieval task (F,3,72 = 20.48, MSe = 3.73, p<.0001), which suggests that the effects of labelling varied across the retrieval tasks.

To test directly the specific experimental predictions, the means were further analyzed using the Cicchetti approximation (Cicchetti, 1972) of the Tukey multiple-range test for means based on the interaction of at least two variables (familywise α = .05). When the subjects were not given verbal labels with the droodles, both types of recognition were better than both types of recall (19.1 & 13.3 vs. 8.0 & 6.6). When the subjects were given labels, correct cue recognition was better than both types of recall (19.5 vs. 11.8 & 16.4), and wrong cue recognition was better than free recall (17.1 vs. 11.8). Overall, these results support the prediction that recognition will be better than recall.

It was predicted that when subjects were not given labels, free recall would be better than cued recall. The means show the predicted pattern (8.0 vs. 6.6), but the difference is not statistically reliable. However, as predicted, when subjects were given labels, cued recall was better than free recall (16.4 vs. 11.8). Furthermore, the benefit brought about by having verbal labels was considerably greater for the cued recall task (increase of 9.8 points which is a 198% improvement) than for the free recall task (increase of 3.8 points which is a 48% improvement).

Also as predicted, correct cue recognition was better than wrong cue recognition when the subjects were not given labels (19.1 vs. 13.3). When the subjects were given labels, there was still an advantage of correct cue recognition (19.5 vs. 17.1), but this difference was not statistically reliable. This non-reliable difference, however, might be due to a ceiling effect for the correct cue recognition. Also as predicted, the benefit of having labels is much greater for the wrong cue recognition (increase of 3.8 which is a 29% improvement) than for the correct cue recognition, where there was no observed improvement. Again, the absence of a labelling effect in the correct cue retrieval task could be due to a ceiling effect. Ceiling effects notwithstanding, the results are consistent with the view that retrieval tasks tap both the semantic and visual components of memory representations for pictures.

DISCUSSION

The experiment was designed to find out if people's memories for pictures include an abstract representation of the meaning of the pictures as well as a visual representation of the picture itself. The subjects' performance in different retrieval tasks support the notion that people represent both the visual information and the meaning of this information.

The results show that, among all conditions, performance was worst in the free recall and cued recall tasks when the subjects were not given verbal labels with the droodles. In these recall tasks the subjects had to rely mainly on the semantic component of their memory representations to facilitate memory retrieval. The droodles, however, were not given meaningful interpretations in these conditions, therefore, the semantic component was weak. They may have visual representations, but these were not particularly useful for recall tasks which are primarily conceptually driven processes.

However, performance in both recall tasks improved substantially when the droodles were presented with meaningful interpretations. These meaningful interpretations strengthened the semantic component of the memory representations, and therefore, facilitated retrieval in the conceptually driven recall tasks. Having a cue that was conceptually related to the target picture in the cued recall tasks further enhanced memory retrieval. With the cue, subjects no longer had to rely solely on stored concepts to facilitate memory. The cue allowed better access to the semantic component of the memory representations. Hence, the stronger labelling effect for the cued recall task than for the free recall task.

Performance was markedly better in the recognition tasks. This result is consistent with the view that in the recognition tasks, subjects are relying on both the semantic and visual components of their memory for pictures. When the droodles were not presented with meaningful interpretations, performance in the recognition tasks was reliably better than in the corresponding recall tasks. While subjects had to depend on a weak semantic component of their memories in the recall tasks, they relied on the weak semantic component and the visual component of their memories in the recognition tasks. The weaker performance in the wrong cue recognition task, compared to the correct cue recognition task, suggests that the visual component of subjects' memories might have consisted of a visual representation of the pair of droodles and not of individual representations of the droodles in the pair. This idea is consistent with the view that the visual representations are analogues of the original visual information instead of being abstracted or organized in a meaningful way.

Since the subjects were already effectively using the visual components of their memories in the recognition tasks, giving meaningful interpretations with the droodles and strengthening the semantic component did not lead to as dramatic labelling effects as it did with the recall tasks. In fact, for the correct cued recognition task, there was no such labelling effect. For the wrong cued recognition task, the additional semantic component helped the subjects overcome the interference brought about by the incorrect visual cues.

Knowing that people's memory for pictures has semantic and visual components has some clear practical implications. For example, this knowledge should guide study strategies used by students of geography, anatomy, biology, and the visual arts. More importantly, however, this knowledge reveals to us important insights into the workings of the human mind. Looming behind the issue of visual and semantic representations of visual information is the larger issue of whether representations in general are concrete analogue representations or abstract propositional representations.

There are those who have advocated for both verbal/propositional and visual representations, but, at the same time, proposed the primacy of the more concrete forms of representations (see e.g., Paivio, 1971). Others have argued that abstract propositional representations are more enduring and are therefore more important in terms of people's long term knowledge (see e.g., Anderson, 1983; Pylyshyn, 1973). The results of this experiment suggest that both forms of representations can be used effectively for different types of activities. The results show the superiority of recognition performance in conditions in which the subjects were not given labels. In these tasks, the subjects only had the visual representation to rely on their memory because the visual information was essentially meaningless to them. Nevertheless, subjects' memory for the pictures was very good. These results show that people can rely almost solely on visual representations when given tasks that involve matching surface properties of the information. Trivial as it may seem, the task of matching surface properties of a stimulus to stored representations is found in several of human being's important activities: when a surgeon decides that the particular organ her scalpel is about to excise is the correct body part her textbook told her to cut; when witnesses pick the face, features, and built of the suspect for a crime in a police line-up; when an infant begins to learn that the particular configuration of features is her mother's face; when the hunter-gatherers had to know which shapes, colors, and sizes of plants are edible or which shapes, colors, and sizes of animals are intent on devouring them, and so on.

On the other hand, the results also show the strong effect of having meaningful verbal labels to accompany the visual information. These labels strengthen the semantic representations, and its effects are most marked in the recall tasks which rely mainly on stored concepts. The results demonstrate the importance of semantic representations of visual information in tasks that involve generating or accessing information from stored representations. These tasks are also found in several of human being's important activities: when a student answer's an essay question in an exam; when a politician gives an extemporaneous speech on her proposals; when an author or a painter renders a particular event or landscape; and our pre-literate ancestors told their next generation about their lore, history, knowledge, and traditions; and so on.

Visual and semantic representations of visual information, or the concrete analogue and abstract propositional representation of information in general, have important functions in human performance. One form of representation may prove superior to another in some activities, but in other activities the reverse may be true. It seems useless, therefore, to claim that one form of representation has primacy over the other. What seems to be important is that both forms of representations, used together or independently, allow human beings to function effectively in an environment that calls on them to rely on their knowledge of a large range of information at any given time.

AUTHOR'S NOTES

Preparation of this paper was supported in part by a University of the Philippines Diamond Jubilee Faculty Grant to the last author. Correspondence regarding this article should be addressed to Dr. Allan B.I. Bernardo, Department of Psychology, Palma Hall Annex, University of the Philippine, Diliman, Quezon City 1101.

REFERENCES

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Bower, G.H.; Karlin, M.B.; & Dueck, A. (1975). Comprehension and memory for pictures. *Memory & Cognition, 3*, 216–220.

Cicchetti, D.V. (1972). Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximate solution. *Psychological Bulletin, 77*, 405–408.

Cooper, L.A. & Shepard, R.N. (1973). Chronometric studies of the rotation of mental images. In W.G. Chase (Ed.), *Visual Information Processing*. Orlando, FL: Academic Press.

Dyne, A.M.; Humphreys, M.S.; Bain, J.D.; & Pike, R. (1990). Associative interference effects recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 813–824.

Finke, R.A.(1980). Levels of equivalence in imagery and perception. *Psychological Review, 87*, 113–132.

Jacoby, L.L. (1983). Remembering the data: Analyzing the interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22*, 485–508.

Jacoby, L.L. & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *3*. 306–340.

Johnston, W.H.; Dark, V.J.; & Jacoby, L.L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 3–11.

Kosslyn, S.M.; Ball, T.M.; & Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 47–60.

Kosslyn, S.M. & Pomerantz, J.R. (1977). Imagery, propositions, and the form of internal representations. *Cognitive Psychology*, *9*, 52–76.

Mandler, G. (1980). Recognizing: The judgement of previous occurrences. *Psychological Review*, *87*, 252–271.

Mandler, G. & Ritchey, G.H. (1977). Long-term memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 386–396.

Paivio, A. (1971). *Imagery and Verbal Processes.* New York, NY: Holt, Rinehart and Winston.

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

Reed, S.K.; Hock, H.; & Lockhead, G.R. (1983). Tacit knowledge and the effect of pattern configuration in mental scanning. *Memory & Cognition*, *11*, 137–143.

Shepard, R.N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 702–703.

Tarr, M.J. & Pinker, S. (1989). Mental rotation and orientation dependence in shape recognition. *Cognitive Psychology*, *21*, 233–282.

Wiseman, S. & Neisser, U. (1974). Perceptual organization as a determinant of visual recognition. *American Journal of Psychology*, *87*, 675–681.