JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 19, 416–429 (1980)

The Effects of Orienting Tasks on Recognition, Recall, and Modality Confusion of Pictures and Words

FRANCIS T. DURSO AND MARCIA K. JOHNSON

State University of New York at Stony Brook

Subjects were presented with a list of pictures and words and performed tasks that oriented processing toward the concept as an image, the concept as a verbal item, or toward underlying referential information associated with the concept. Recognition (Experiment 1) for concepts presented as pictures was superior only when the task required subjects to orient to the concept as a verbal item per se; a word superiority effect was observed when orientation was to the concept as a picture per se. The mode of presentation did not influence concept recognition when subjects had focused on the referential meaning of the items. Memory for form was also influenced by the task, with subjects more likely to claim they saw a word as a picture than vice versa, but only in the referent tasks. When subjects were asked to recall, rather than recognize the concepts (Experiment 2), there was a similar, though not identical, pattern of results. The sensory-semantic model of D. L. Nelson, V. S. Reed, and C. L. McEvoy (*Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 485–497) had difficulty with some aspects of these data. The availability of information regarding cognitive operations performed on the input seems to be an important component of memory traces.

Interest in how learning of, and memory for, conceptual information is influenced by the mode in which the information is presented has led researchers to focus their interests on the similarities and differences in picture and word processing. One of the most frequently cited findings from these research efforts is that concepts are more likely to be remembered if presented in a pictorial form compared to a verbal form (e.g., Paivio, Rogers, & Smythe, 1968; Shepard, 1967).

However, picture superiority is not an invariant result when memory for pictures and words are compared in the laboratory. For example, Nelson, Reed, and Walling (1976) eliminated and in some conditions reversed, the picture superiority effect simply by increasing the schematic similarity of the pictures serving as stimuli in a paired associate list. Other studies have also failed to yield a picture superiority effect (Gadzella & Whitehead, 1975; Hasher, Riebman, & Wren, 1976; Shepard, 1973; Wicker, 1971) indicating that the conditions necessary for the effect remain unclear.

Attempts to explicate the picture superiority effect and its accompanying boundary conditions have led to a number of models of picture and word processing (e.g., Nelson, Reed, & McEvoy, 1977; Paivio, 1971; Snodgrass, Note 1). The sensory-semantic model (Nelson et al., 1977) has been particularly successful in accounting for the data in the current literature, and was partly responsible for stimulating the present studies. This model explains the usual superior memory for pictures by assuming that (a) pictures have more distinctive sensory codes than do words and (b) pictures are more likely to undergo semantic processing than are their verbal labels. Both of these, distinctive sensory codes and semantic processing, should confer an ad-

The authors would like to thank Warren Dodge and Gary Miller for their help in data collection. Thanks also to our artists Jo Firman and Joe Ruzcek for their assistance. This research has received support from Grant BNS-7813054 from the National Science Foundation. Requests for reprints should be addressed to Francis T. Durso, Department of Psychology, University of Oklahoma, Norman, OK 73019.

0022-5371/80/040416-14$02.00/0

vantage to pictures in memory tasks. Assumption b is based on a model of semantic access in which, for most tasks, pictures directly activate meaning whereas words typically first activate a phonemic code and subsequently (but not necessarily) activate meaning. Thus assumption b is an extension of the levels of processing (Craik & Lockhart, 1972) postulate that semantic processing results in more durable memories than phonemic processing.

The purpose of the first experiment was to determine the effect of different orienting tasks on the subsequent recognition of items presented as pictures or words. While assumption a should continue to confer an advantage to pictures through each of the tasks, orienting tasks allow manipulation of the information focused on during acquisition and thus allow us to maximize or minimize the effects of assumption b. We included a pair of tasks designed to encourage subjects to deal explicitly with a verbal label or the name code per se and a pair of tasks designed to encourage subjects to deal explicitly with a picture or image code per se. For the first pair of tasks, the sensory-semantic model predicts a large picture superiority effect by assuming that pictures have more distinctive sensory codes and that, in these tasks, the probability that pictures receive semantic processing is much greater than the probability that words do. These two factors should combine to produce a large superiority for pictures in recognition. For the second pair of tasks, the sensory-semantic model would assume that the probability of words receiving semantic processing is greater than the probability that pictures receive such processing (reversing assumption b), thus reducing, eliminating, or reversing (depending on the relative potency of assumption a) the picture superiority effect. This is because subjects would presumably not have to consult semantic information to make an image of a picture they had just seen. On the other hand, making an image from a verbal label should require first activating information about the meaning of the verbal label and then converting this meaning to pictorial information. In addition, Experiment 1 explored the effect of six tasks which required the subjects to consider different information about the referent of concepts. The specific orienting tasks will be described below, but in general, the sensory-semantic model would expect that these tasks will result in a picture superiority effect whose magnitude should lie somewhere between the large effect expected in the verbal tasks and the small effect expected in the imagery tasks. These tasks should tend to equate semantic processing of pictures and words thus minimizing the influence of assumption b.

There is already some research that considers how different acquisition tasks affect memory for pictures and words. Most of this work was directed at the effect of instructions to image (Hasher et al., 1976; Paivio & Csapo, 1973; Snodgrass, Burns, & Pirone, 1978; Wicker, 1971; Wicker, Edmonston, & McClure, 1973) or with the difference between intentional and incidental memory instructions (Paivio & Csapo, 1973). One study (D'Agostino, O'Neill, & Paivio, 1977) was concerned with how recall of pictures and words varied under structural, phonemic, and semantic orienting tasks. Semantic processing of words improved performance to a level comparable to pictures that had undergone either semantic or phonemic processing.

The current referent tasks asked subjects to make judgments of size, animacy, function, naturalness, relevant features, or physical complexity of the real world referent. In the D'Agostino et al. study it cannot be determined which nor in what proportion these types of semantic processing occurred. D'Agostino et al. had subjects decide if a stimulus could be inserted into a sentence frame to yield a meaningful sentence. Two meaningful sentence frames were constructed for each stimulus; each subject saw the test stimulus twice, once for each frame. It is possible that the failure of D'Agostino et al. to find differences between pictures and words in the semantic orienting task was due to averaging over a number of different *types* of semantic processing, some of which might produce a picture superiority effect, some of which might produce no difference, or even some of which might produce a word superiority effect. For example, if D'Agostino et al. had considered separately the sentence frames, if any, which focused on the physical characteristics of the referent, they may have found that such processing would result in superior memory for words whereas consideration of sentences which focused on the, say, animacy of the referent may have produced a picture superiority effect. Another facet of the D'Agostino et al. study that limits generalizability is the fact that the concepts were repeated. Repetition would tend to benefit the weaker trace relative to the stronger one and could have resulted in equivalent recall, not because picture and word traces are similar, but rather because one accrued the benefits of repetition more rapidly (cf. Durso & Johnson, 1979). Thus, we felt that with single presentations of stimuli, covering a range of semantic processing tasks, important differences generally ignored under the rubric of "semantic" might be revealed.

In addition to observing the number of times concepts were recognized as having previously occurred, another major dependent variable was the number of times the mode of presentation was incorrectly reported. Studies have typically shown that subjects are quite accurate in remembering the form in which the concept was presented (Madigan, 1974; Snodgrass, Wasser, Finkelstein, & Goldberg, 1974; Potter, Valian, & Faulconer, 1977), although in some situations they do make errors (Snodgrass & McClure, 1975).

The sensory-semantic model (Nelson, 1979) contends that when an item is processed, some codes will be activated even though these codes are not focal in performing the task. These nonfocal codes will subsequently influence memory and should vary with the nature of the orienting task, that is, with the nature of the focal code. If saying "*picture*" to a concept presented as a word or saying "*word*" to a concept presented as a picture varies with the task, this can provide us with information concerning the nature of the codes activated during acquisition—information which could not be gleaned from the recognition data alone. For example, if people in the referent tasks identify many pictorially presented items as having been presented as words, this would be an indication that the name code was a relevant aspect of pictures when making judgments about the referent. This would be consistent with the claim that pictures are more likely to be coded in both an imaginal and verbal code than are words (Paivio, 1971). The other possibility is that subjects will more likely confuse verbally presented items, claiming they were presented pictorially. This could come about if processing a word for information about the referent activates nonfocal codes similar to the codes usually activated when a picture is processed.

Experiment 1

The first experiment assessed recognition for concepts presented in a verbal or pictorial form. Three groups of incidental orienting tasks that required focusing on the verbal label per se, pictorial characteristics per se, or on the underlying referent were used. Recognition memory was assessed by having subjects respond *picture*, *word*, or *new* to auditorially presented labels of concepts that had been presented earlier as a picture or word or to labels of concepts that were not on the acquisition list.

Method

Stimuli. The stimuli consisted of 140 words chosen from the Kučera and Francis (1967) word norms. The words ranged widely in frequency count (1 to 591 parts per million, ppm) and averaged 36 ppm.

Only words that could serve as picture labels were chosen. A line drawing of the concept was created for each of the 140 words. Line drawings that did not unambiguously elicit the corresponding label were redrawn until a group of 20 judges agreed on the name of the drawing. The words were typed in all upper case letters on white, 5 × 8-in. cards. The drawings were made with a black felt tip pen on the 5 × 8-in. cards. Either one drawing or one label appeared on a card.

The stimuli were then divided into two lists, each consisting of 70 drawings and 70 labels. The concepts represented by drawings in one list were represented by words in the other and vice versa. In this way each list included each of the 140 concepts once and only once. The assignment of concept to list position was random with the restriction that within every four items two pictures and two words occurred. A concept had the same ordinal position in both lists, thus any difference between memory of that concept as a word or as a line drawing could not be a consequence of its serial position nor of the concepts in neighboring positions. In addition to the 140 items, each list was preceded by four fillers and succeeded by three fillers to reduce the effects of primacy and recency.

To assess recognition, a tape consisting of the labels of the 140 concepts and 70 new concepts was made. The 70 new items also had an average frequency of approximately 36 ppm and each was capable of being represented as a simple line drawing. The 210 items were ordered randomly on the tape with the restrictions that within every six items two old-picture labels, two old-word labels, and two new labels occurred and the average lag between presentation on the list and occurrence on the tape was comparable for pictures and words. The items were read at a rate of 5 seconds by a male, native English speaker.

Tasks. A subject could participate in one of 10 possible orienting tasks during acquisition (see Table 1). The tasks could a priori be grouped into three classes on the basis of

TABLE 1

DESCRIPTION OF THE 10 ORIENTING TASKS

Tasks	Description	Response[a]
Verbal		
Naming	What is the name of the item?	knife
Last Letter	What is the last letter of the name?	e
Imaginal		
Artist Time Judgment	How long would (did) it take to draw the object?	3
Explicit Imagery	Create an image. How "good" is the image?	9
Referential		
Size	How large is the real world object?	2
Physical Complexity	How physically complex?	1
Animacy	How much movement does the object go through daily?	2
Function	What is the object used for? You ________ it.	cut with
Relevant Features	What must the object have to be that object?	blade
Category Membership	Is the object artificial (i.e., man-made) or naturally occurring?	Art

[a] All number responses were on a scale from 1 to 10.

whether the task required information about the verbal label, an image or picture, or the referent.

The two verbal tasks were Naming and Last Letter. In these tasks the subjects wrote the name or the last letter of the name during the acquisition phase of the experiment. According to the Nelson et al. model these tasks should require more or deeper processing for pictures, since the model assumes that phonemic information can be accessed by pictures only after some semantic processing has taken place while words can access phonemics immediately after physical processing.

The two pictorial tasks were Artist Time judgment and Explicit Imagery. The Artist task required subjects to use a 10-point scale to decide how long it took for the artist to make the line drawing or, in the case of words, to decide how long it would take for the artist to render a line drawing of the concept. The Explicit Imagery task required subjects to make an image and then to give a number from 1 to 10 reflecting how "good" the image was. No specification of what *good* meant was made. Insofar as these tasks succeeded in directing subjects to consider the stimulus when presented with a line drawing or an image when presented with a word, the sensory-semantic model would contend that the words would undergo semantic processing in route to redintegrating an imaginal code while physical processing of the line drawing would not require semantics.

The referent tasks required subjects to make judgments of Size, Physical Complexity, Animacy, Function, Relevant Features, or Category Membership of the real world object. As indicated in Table 1, some tasks required judgments in the form of a number from a 10-point scale whereas other tasks required the subject to supply a word or phrase about the referent.

These particular tasks were chosen because they seemed to cover a broad range of semantic information. Additionments are of particular theoretical interest in that the semantic information could be more available to one mode or the other (Paivio, 1971). For example, it has been proposed that size information is more readily accessable from a pictorial input.

Procedure. The orienting task was a between subjects factor while the mode in which the concept was presented was within subjects. Half of the subjects in each task received one list and half received the other list, that is, half the subjects saw a particular concept as a picture and the others saw the same concept as a word.

Subjects were run in groups of one to three. The experimenter read aloud the appropriate instructions for the task. It was emphasized that we were interested in their judgments and no mention was made of any subsequent memory test. The experimenter then went over a practice judgment which illustrated that the response should be the same regardless of whether the item was presented as a picture or a word.

During the acquisition phase, the experimenter held the cards up one at a time to the subjects. The subjects were presented a new stimulus at a rate of 5 seconds per item. A tape recorder which emitted a tone every 5 seconds served to pace the experimenter and the subjects through the acquisition phase.

Following acquisition, the tape consisting of the 210 verbal labels, one every 5 seconds, was played to the subjects. Subjects were instructed to indicate whether the concept was seen earlier as a picture, as a word, or whether it was a new item. In this testing procedure, subjects could respond *picture*, *word*, or *new* to items that actually were presented as pictures, words, or neither.

Subjects. Twelve subjects served in each of the 10 orienting tasks for a total of 120 participants. The subjects were volunteers from an introductory psychology course at the State University of New York at Stony Brook who received extra credit for their

Results and Discussion

Recognition. A response of either *picture* or *word* was taken as an indication that the item was remembered as having been presented during acquisition, regardless of the actual mode of presentation. Thus, in these analyses it was of interest whether subjects responded *old* (i.e., picture or word) or *new* to items that were originally presented as a picture, word, or distractor.

Task × Mode analyses of variance were conducted on the unadjusted correct old responses (Hits) for each group of tasks: verbal, imaginal, and referential. Analyses were also conducted on Hits adjusted by subtracting the number of old responses to new items (False Positives) from the Hits for each subject. This correction could only produce disagreement with the unadjusted analyses in detecting a main effect of Task and will be noted when it occurs. Significant results are at an α level of .05 or better.

Figure 1 shows the results for each group of tasks while Table 2 presents a breakdown by task of the results shown in the figure. For the verbal tasks, Naming and Last Letter, pictures were recognized more frequently than were words, $F(1,22) = 85.19$, $MS_e = 76.97$. Mode did not interact with task. The tasks were comparable when unadjusted Hits were considered; however, Naming produced superior recognition when the rates of false positives were taken into account, $F(1,22) = 9.83$, $MS_e = 113.13$. The verbal tasks clearly demonstrated superior recognition for pictures.

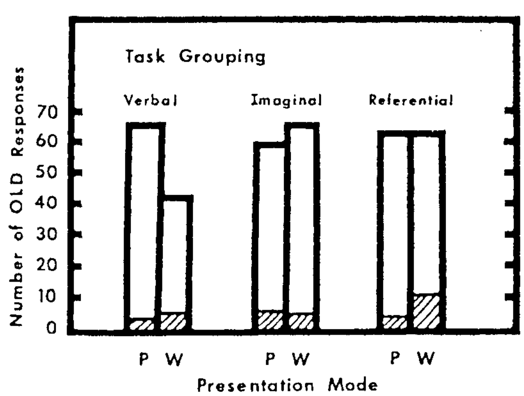

FIG. 1. Number of items recognized as old for each group of tasks and each input mode. Shaded regions are the number of correct old responses that were accompanied by an incorrect identification of the presentation mode (confusions).

For the imagery tasks, Explicit Imagery and Artist, the sensory-semantic model predicts the smallest picture superiority effect. In fact, as Figure 1 shows, the picture superiority effect was not only substantially reduced but was reversed. There was a main effect of modality, $F(1,22) = 48.28$, $MS_e = 9.97$. Unlike the verbal tasks, with the imagery tasks, subjects were more likely to recognize a concept that had been presented as a word than a concept that had been presented as a line drawing. By changing the orienting task from one that focused on the label to one that focused on the concept as a picture, we were able not only to eliminate the picture superiority effect but actually to reverse it. There was also a main effect of task, $F(1,22) = 9.51$, $MS_e = 14.36$, but only when no correction was made on the Hits, and the Task × Mode interaction was significant, $F(1,22) = 16.18$, $MS_e = 9.97$. The interaction reflects that the advantage of words was greater when subjects performed the Artist task than when subjects performed the Explicit Imagery task. However, consideration of each task separately indicated that the word superiority effect was present in the Explicit Imagery task, $F(1,11) = 6.67$, $MS_e = 6.39$, as well as the Artist Time judgment task, $F(1,11) = 44.29$, $MS_e = 13.55$.

For the referent tasks, Figure 1 shows that the magnitude of the picture superiority effect was clearly between the large effect observed in the verbal tasks and the reversal observed in the imagery tasks. There was no difference in recognizability of pictures and words. Over the six tasks, 90.24% of the line drawings were recognized as old and 90.22% of the words were recognized as old. The largest difference

TABLE 2

NUMBER OF CORRECT RECOGNITIONS, FALSE POSITIVES, AND CONFUSIONS FOR EACH TASK AS A FUNCTION OF THE INPUT MODE AND RESPONSE

	Number of Hits		Number of False Positives		Number of confusions	
Tasks	Old|P	Old|W	P|N	W|N	W|P	P|W
Verbal						
Naming	66.67	43.83	1.83	5.08	2.67	5.75
Last Letter	66.08	42.17	1.08	14.33	4.67	4.67
Imaginal						
Artist Time judgment	56.00	66.00	2.33	3.42	5.42	3.50
Explicit Imagery	63.58	66.25	3.00	7.17	6.25	6.33
Referential						
Size	59.25	61.00	3.83	7.33	7.58	12.00
Physical Complexity	63.33	64.08	2.50	4.17	3.25	7.67
Animacy	65.67	65.67	5.00	8.17	4.25	9.92
Function	65.25	64.92	4.17	6.25	3.42	13.25
Relevant Features	64.08	64.17	3.17	3.50	2.08	11.08
Category Membership	61.42	59.08	4.17	4.42	2.75	10.33

Note. P, Picture; W, word; old, P or W; N, new. Maximum cell entry is 70.

(3.3%)[1] was observed in the natural–artificial task while absolutely no difference was present in the animacy task.

Overall, the relative recognition performance for pictures and words in each of the three task groupings can be explained by a number of models of picture and word processing, though in most cases such explanations would have a distinct post hoc flavor. For example, the original dual-code model (Paivio, 1971) could claim that in the imagery tasks, the usually low probability of dually coding words was increased beyond the usually high probability of dually coding pictures and that in the referent tasks, the probabilities were more or less equated. Such an explanation would reduce the distinction between that model and the sensory-semantic model. The reformulation of the dual-code model (Paivio, 1978) has more success, and in fact, for these data, is not discriminable from the sensory-semantic model.

The predictions of the sensory-semantic model were confirmed. When subjects focused on the name of the concept, sensory distinctiveness of pictures and a higher probability of semantic processing for pictures combined to produce a large picture superiority effect. In the imagery tasks the fact that we observed a word superiority effect and the fact that the effect (9%), though reliable, was not as large as the picture superiority effect observed in the verbal tasks (33%) are consistent with the notion that words are more likely to receive semantic processing than are pictures in the imagery tasks but that the sensory distinctiveness of pictures is working against observing a large word superiority effect. Finally, when decisions were based on the referent, performance was comparable for

[1] Percentages reported in this article use the number of items actually presented during acquisition as a base. Experiment 1 used 70 pictures and 70 words while Experiment 2 used 20 of each type of stimulus. When a difference is reported as a percentage, including those characterizing the picture superiority effect, the percentage is simply the percentage of pictures minus the percentage of words remembered.

pictures and words. Verbal tasks produced the largest picture superiority effect (33%); imagery tasks produced the smallest picture superiority effect (−9%); referent tasks produced an effect between these two (0%).

One could argue that we did not observe a picture–word difference in the referent tasks because performance was so close to ceiling that any differences would not be detectable. This seems unlikely given that both picture and word performance was raised still higher in the verbal and imagery tasks, respectively. One might also argue that the level of recognition of words for each of the tasks was due, in part, to the fact that there was a greater change from acquisition to test in the case of pictures than in the case of words. However, this factor should have been present in each task grouping and could not account for the relative ordering of the tasks in terms of the magnitude of the picture superiority effect, although it might account for the point at which a reversal occurs or the point at which equal performance for pictures and words is observed.

Though there was no effect of mode in the referent tasks, the tasks differed in the overall recognition performance. In fact, the only significant variability was contributed by the task factor of the unadjusted Hits analysis, $F(5,66) = 3.96$, $MS_e = 35.04$. The variation in the tasks can be seen in Table 2, and though perhaps not surprising, does serve to point out the problems that may arise if a researcher decides to use only one semantic orienting task or decides to consider a number of types of semantic processing under the generic heading of semantic. For example, D'Agostino and associates' subjects decided whether or not the insertion of the stimulus into a sentence frame produced a meaningful sentence. D'Agostino et al. found that recall of pictures in their semantic task was equal to recall of pictures in a naming task. They reasoned that if naming is at a deeper level than meaning for pictorial stimuli, then the naming tasks should have produced superior recall. Since there was no difference, they argued that picture and word differences are better viewed as differences in the number of codes activated, and not as differences in the depth of processing.

If we follow the same logic in the current study, our conclusions would depend on which semantic or referent task we used to compare with the naming task. If we use Relevant Feature as our semantic task, we observe almost no difference between the semantic and the naming tasks. From this we would agree with D'Agostino et al. and take it as support for dual coding. If we use natural–artificial as our semantic task, then naming yields significantly better recognition of pictures than does our semantic task whether comparing adjusted, $F(1,22) = 8.60$, $MS_e = 33.36$, or unadjusted, $F(1,22) = 10.77$, $MS_e = 15.34$, recognition scores. This would appear to be in agreement with a depth of processing notion.

Overall, in the current study, we only observed a substantial picture superiority effect when the task could be completed by dealing with the label of a concept as a sign per se. Whenever the task required treating the label as a symbol for a referent, the picture superiority effect was eliminated. In fact, when the task allowed the line drawing to be treated as a sign per se, the picture superiority effect was reversed and a word superiority was found.

The literature has not always reported a picture superiority effect (e.g., Hasher et al., 1976). The findings of the current study suggest that whether pictures are easier to remember than words would be directly influenced by the tack taken by the subject to remember the materials. The results suggest that subjects usually employ a strategy more akin to naming pictures and reading words than to semantically processing both items.

Confusions. An aspect of the recognition performance not captured in the previous analyses are confusions. In the recognition analyses, subjects were scored correct if they said either *picture* or *word* to an old item. A confusion refers to the correct recognition of a concept as old, accompanied by the incorrect identification of the original mode of presentation. If a concept was presented as a picture and the subject claimed to have seen it as a word, this is a W|P (word given picture) confusion; if the concept was actually presented as a word and the subject claimed to have seen it as a picture this is a P|W (picture given word) confusion. These confusions are represented in Figure 1 as the shaded area at the base of each bar.

The first thing to note is that each task grouping showed a number of confusions. Over 6% of the items originally presented in the Verbal tasks, over 7% in the Imagery tasks, and over 9% in the Referent tasks were recognized but were misidentified with respect to the origin of the information.

Analyses of confusions proceeded in a manner similar to the analyses of recognition. The analyses of variance were again task by mode analyses, this time with the number of confusions serving as cell entries. For the verbal and the imagery analyses, no reliable differences in confusability were detected. In these tasks, subjects were as likely to claim that they saw a word as a picture as they were to claim they saw a picture as a word.

For the referent tasks, subjects were much more likely to say *picture* to something shown as a word than the converse, $F(1,66) = 68.95$, $MS_e = 24.28$. No effect of task or interaction with task was present. The greater frequency of P|W confusions relative to W|P confusions appears to be both quite robust and quite large. In fact, when each of the six tasks is considered separately (see Table 2) five of them show the effect at a significant level and the sixth (i.e., size) approaches significance. Over the six tasks, 15.3% of the 70 words were reported as having occurred as pictures, while 5.6% of the 70 pictures were reported as having occurred as words. The most confusions occurred after subjects had decided on the function of the concept; the subjects claimed that 19% of the words presented to them were actually pictures whereas they claimed that they had seen pictures as words 5% of the time. If the percentages are based on the number of items actually recognized, rather than the number presented, the percentages increase but, in these tasks, the difference in P|W versus W|P confusions remains about the same.

This confusion effect is not due to a bias for subjects in the referent conditions to respond *picture*. In fact, responses to the distractor items reveal that there is a strong bias to respond *word*, $F(1,66) = 8.37$, $MS_e = 14.46$, and this is not a function of the task.

It should be noted that, on the surface, the results of our verbal and imaginal tasks do not agree with the results of Snodgrass and McClure (1975) who instructed their subjects to rehearse words and labels or to form images from words and pictures. Snodgrass and McClure found more P|W than W|P confusions in both the verbal and the imagery conditions. Their study differed from ours in that their subjects were aware that a recognition test would follow. One possible resolution of this discrepancy is that subjects may be more likely to use some of the strategies employed by our subjects from the referent tasks when they expect a memory test.

One of the substantial contributions made by Nelson and his colleagues to the depth of processing literature, which is reflected in the processing assumptions of the sensory-semantic model, is that the memory of an event will be comprised of codes other than the focal code and these other codes could produce facilitation or interference when the memory is tested. In terms of the sensory-semantic model, it appears that either the referent tasks lead to imaginal/pictorial codes in addition to the focal semantic code or the focal code is itself imaginal/pictorial (cf. Paivio, 1978). When the subjects attempt to remember the origin of the information, the representation of both words and pictures now have a number of imaginal/pictorial codes activated. Because these imaginal/pictorial codes are usually by-products of having processed a picture, the subjects will claim that they were presented the information in a pictorial format. This explanation of performance on the referent tasks follows from the sensory-semantic model and merely requires the reasonable assumption that making decisions about referents leads to the creation of sensory codes more like those of the referent than those of the corresponding label. The confusion data suggest that the referent tasks used in the current study caused the creation of distinctive sensory codes for verbal inputs similar to those usually accompanying pictorial inputs. Judgments concerning real world referents produced nonfocal sensory codes for not only pictorial inputs but for a number of the verbal inputs as well.

These data suggest a limitation of the sensory-semantic model. If the referent tasks lead to picture-like codes, then the imagery tasks should be even more likely to produce such codes and yield even greater differential confusions. However, in fact, the imagery tasks do not show such a pattern. It should be noted that the original dual-code model also has difficulty with this pattern of confusions. If pictures are more likely to be dually coded then it should be the case that subjects will be more likely to report having seen pictures as words than words as pictures, especially in the verbal tasks where dual coding is apparently assured.

Experiment 2

The sensory-semantic model did quite well in explaining the recognition data of Experiment 1. One problem was the inconsistency of some of the confusion data with predictions from the model.

Experiment 2 was concerned with extending the results and conclusions of Experiment 1 from recognition to recall. If the findings generalize, then superior recall of pictures relative to words should be the greatest in the verbal tasks, smallest in the imagery tasks, and intermediate in the referent tasks. In addition to exploring the generality of the results of Experiment 1, a recall test should be less subject to potential ceiling effects compared to a recognition test, as well as less subject to any problems inherent in using an auditory verbal label to assess memory. In addition, although we expect fewer confusions in recall simply because overall recall performance should be less than recognition and therefore fewer confusions would be possible, it should be the case that differential confusability of pictures and words occurs if the confusion data reflect differences in the encoding of the stimuli rather than how the memory is tested.

Method

Stimuli. Forty concepts were randomly sampled from the stimuli used in Experiment 1. The order of the concepts and the creation of two lists differing in how the concept was represented followed the procedure of Experiment 1. Except for these changes and the elimination of the recognition tape, the materials for this experiment were identical to those of the first.

Procedure. Except for the decrease in the number of stimuli the acquisition phase was identical to that of Experiment 1. The test was changed by requiring subjects to free recall the items presented earlier. Subjects were allowed 7 minutes to write down the labels of the concepts. Following recall, subjects were instructed to go over their list and indicate the origin of the concept by writing *picture* or *word* next to the recalled items. This phase required approximately 2 minutes in the most extreme case. Finally subjects were presented with the line drawings they had seen before and were asked to name the concepts. This information was used to give recall credit when subjects used labels other than might be expected for a picture. However, rarely (less than 1%) did this information disagree with the labels supplied by the norming group.

Subjects. Only six subjects per task, for a total of 60, were utilized in this experiment since we were primarily concerned with the extent to which the overall patterns of Experiment 1 generalized to recall. Subjects were again from an introductory psychology course at the State University of New York at Stony Brook. None had participated in Experiment 1.

Results and Discussion

Recall. Task × Mode analyses of variance were performed on the number of recalled items for each of the groups of tasks. Figure 2 shows the recall performance for each group of tasks while Table 3 gives performance for individual tasks.

Overall, the pattern of results of Experiment 2 agreed with the recognition results. The verbal tasks produced the largest picture superiority effect (27%), imagery tasks produced the smallest picture superiority effect (3%), and the referent tasks produced an intermediate effect (5%). As in recognition, the effect in the verbal tasks was reliable, $F(1,10) = 79.49$, $MS_e = 2.28$, while the effect in the referent tasks was not different from zero. Unlike recognition, however, a reversal was never observed, and though statistical equivalence was observed

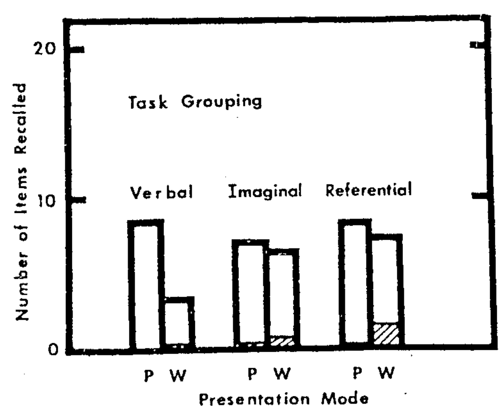

FIG. 2. Number of items recalled for each group of tasks and each input mode. Shaded regions are the number of recalled items that were accompanied by an incorrect identification of presentation mode (confusions).

TABLE 3
NUMBER OF ITEMS RECALLED AND NUMBER OF CONFUSIONS FOR EACH TASK AS A FUNCTION OF THE INPUT MODE

	Number of items recalled		Number of confusions	
Tasks	P	W	W|P	P|W
Verbal				
Naming	9.83	3.67	.00	.33
Last Letter	7.50	2.67	.00	.17
Imaginal				
Artist Time judgment	6.83	6.67	.17	.33
Explicit Imagery	7.17	6.00	.17	1.17
Referential				
Size	6.00	6.00	.00	1.17
Physical Complexity	8.50	9.00	.17	.50
Animacy	9.83	8.67	.00	.67
Function	8.33	6.67	.00	3.83
Relevant Features	9.17	6.00	.00	1.17
Category Membership	8.50	7.50	.00	1.33

Note. P, Picture; W, word. Maximum cell entry is 20.

between pictures and words in the imagery and the referent tasks, recall of pictures was superior in six of the eight tasks.

Confusions. As expected, the number of confusions here was less than in recognition, given the smaller number of opportunities for a confusion to be recorded. However, the number of P|W responses were of a nontrivial magnitude especially in the referent tasks where they occurred for over 7% of the 20 items originally presented as words. This is even more important given that average recall for words in the referent tasks was less than 37%; thus, one out of every five words recalled was thought to have been presented pictorially.

Analyses of variance were again conducted on the P|W and W|P responses. Neither the Imaginal tasks nor the Verbal tasks showed reliable differences. Analyses of the Referential tasks yielded a main effect of mode, $F(1,30) = 25.06$, $MS_e = 1.44$, a main effect of task, $F(5,30) = 3.02$, $MS_e = 1.41$, and an interaction, $F(5,30) = 3.21$,

$MS_e = 1.44$. Subsequent tests of each reference task separately revealed that only the Function task produced reliably more P|W confusions than W|P confusions, $F(1,5) = 11.35$, $MS_e = 3.88$; however, the effect was in the same direction for each task and, in fact, for five of the six referent tasks there was never a W|P confusion (see Table 3), while the P|W confusions ranged from 2.5 to 19.2% of the 20 presented words. It is interesting to note that, as in recognition, considering the function of a referent produced the greatest difference in P|W compared to W|P judgments. In fact, for function judgments, more than one out of every two words recalled was thought to have been presented pictorially while recalled pictures were never misidentified.

General Discussion

One important contribution of the present studies is that they point out that whether pictures or words show superior memorability is a function of the orienting task. The "usual" picture superiority effect was clearly present only when subjects oriented toward a verbal label per se and not toward the underlying referent. When, however, we induced the subject to consider the concept as a line drawing per se and not necessarily the meaning of the underlying referent, words were recognized better than pictures and recall was comparable. When subjects considered the real world referent in making their decisions, pictures and words were recognized and recalled equally often. These results imply that the magnitude of the picture superiority effect, or whether it is obtained at all, depends on the tack taken by subjects to encode the stimuli. Depending on how many subjects, and the number of times each subject decides to name the stimuli, imagine the stimuli, or consider the meaning of the stimuli, the picture superiority effect could be quite large or quite small.

The pattern of results observed in the present studies is generally consistent with the sensory-semantic model of picture and word processing. However, the model does not handle the pattern of confusions very well. With respect to confusions, the sensory-semantic model would claim that a number of nonfocal sensory codes, similar to the sensory codes activated during the processing of a picture, were activated from the verbal input items, accounting for the relatively high likelihood of subjects claiming that words had been presented as pictures in the referent tasks. The model would, however, generate the expectation that the imagery tasks would be even more likely to lead to such sensory codes. The greater frequency of these codes should make it even more likely that a word trace would be identified as a picture than the converse. Rather than the robust effect that would be expected, we did not find a preponderence of P|W confusions in the explicit imagery tasks.

Therefore, the confusion data invite the conclusion that the memories created in the explicit imagery tasks have available something in addition to sensory information that the memories created in the referent tasks do not. This additional component could give an advantage to subjects in the imagery conditions when they are deciding the origin of the pictorial information produced for word stimuli. When subjects were asked to identify whether an item was a picture or a word, and they previously created images for the words, they were being asked, in effect, to discriminate perceptually derived from imaginally derived pictorial information (Johnson, Raye, Wang, & Taylor, 1979). A number of investigators have considered the possibility that information about the cognitive operations performed at input is preserved in memory (e.g., Kolers, 1973, 1975; Raye, Johnson, & Taylor, 1980; Russo & Wisher, 1976; Johnson & Raye, Note 2). Furthermore, it seems likely that when pictorial information is explicitly or intentionally created in response to a verbal input, the memory will include more highly available information concerning cognitive operations than when such information is implicitly or automatically produced during stimulus processing (Johnson & Raye, in press). In the orienting tasks of our studies, subjects in the explicit imagery task or the artist time judgment task would have been purposefully creating images; subjects in the referent tasks, on the other hand, would have been incidentally doing so. The more salient cognitive operations information included in the memories created by the explicit imagery subjects should therefore have helped them in determining the origin of particular memories. So, deciding that a knife is used to cut with will create a memory trace containing sensory information relevant to the object even if the word, and not the object, served as the stimulus for the decision. Similarly, the trace resulting from imaging a knife will contain sensory information. However, in the latter case, the trace contains the additional information that the sensory information was generated and not perceived.

In summary, the sensory-semantic model of Nelson et al. has many advantages, most notably, the capability of dealing with a wide range of data while being reasonably specific. However, those findings of the present experiments suggesting cognitive operations are an important component of memories, point to a potentially interesting omission of the model.

REFERENCES

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.

D'AGOSTINO, P. R., O'NEILL, B. J., & PAIVIO, A. Memory for pictures and words as a function of level of processing: Depth or dual coding? *Memory and Cognition*, 1977, 5, 252–256.

DURSO, F. T., & JOHNSON, M. K. Facilitation in naming and categorizing repeated pictures and words. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, 5, 449–459.

GADZELLA, B. M., & WHITEHEAD, D. A. Effects of auditory and visual modalities in recall of words. *Perceptual and Motor Skills*, 1975, 40, 255–260.

HASHER, L., RIEBMAN, B., & WREN, F. Imagery and the retention of free-recall learning. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 172–181.

JOHNSON, M. K., & RAYE, C. L. Reality monitoring. *Psychological Review*, in press.

JOHNSON, M. K., RAYE, C. L., WANG, A. Y., & TAYLOR, T. H. Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, 5, 229–240.

KOLERS, P. A. Remembering operations. *Memory and Cognition*, 1973, 1, 347–355.

KOLERS, P. A. Specificity of operations in sentence recognition. *Cognitive Psychology*, 1975, 7, 289–306.

KUČERA, H., & FRANCIS, W. Computational analysis of present-day American English. Providence, R.I.: Brown Univ. Press, 1967.

MADIGAN, S. Representational storage in picture memory. *Bulletin of the Psychonomic Society*, 1974, 4, 567–568.

NELSON, D. L. Remembering pictures and words: Appearance, significance, and name. In L. Cermak & F. I. M. Craik (Eds.), Levels of processing. Hillsdale, N.J.: Erlbaum, 1979.

NELSON, D. L., REED, V. S., & MCEVOY, C. L. Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 485–497.

NELSON, D. L., REED, V. S., & WALLING, J. R. Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, 2, 523–528.

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

PAIVIO, A. Mental comparisons involving abstract attributes. *Memory and Cognition*, 1978, 6, 199–208.

PAIVIO, A., & CSAPO, K. Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, 1973, 5, 176–206.

PAIVIO, A., ROGERS, T. B., & SMYTHE, P. C. Why are pictures easier to recall than words? *Psychonomic Science*, 1968, 11, 137–138.

POTTER, M. C., VALIAN, V., & FAULCONER, B. A. Representation of a sentence and its pragmatic implications: Verbal, imagistic, or abstract? *Journal of Verbal Learning and Verbal Behavior*, 1977, 16, 1–11.

RAYE, C. L., JOHNSON, M. K., & TAYLOR, T. H. Is there something special about internally generated information? *Memory and Cognition*, 1980, 8, 141–148.

RUSSO, J. E., & WISHER, R. A. Reprocessing as a recognition cue. *Memory and Cognition*, 1976, 4, 683–689.

SHEPARD, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 156–163.

SHEPARD, W. O. Pictures vs. words: Some discrepant results. *Psychological Reports*, 1973, 32, 619–624.

SNODGRASS, J. G., BURNS, P. M., & PIRONE, G. V. Pictures and words and space and time: In search of the elusive interaction. *Journal of Experimental Psychology: General*, 1978, 107, 206–230.

SNODGRASS, J. G., & MCCLURE, P. Storage and retrieval properties of dual codes for pictures and words in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1975, 1, 521–529.

SNODGRASS, J. G., WASSER, B., FINKELSTEIN, M., & GOLDBERG, L. B. On the fate of visual and verbal memory codes for pictures and words: Evidence for a dual coding mechanism in recognition memory. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 27–37.

WICKER, F. W. Pictures, words, and imagery mediation in paired-associate learning. *Perceptual and Motor Skills*, 1971, 33, 135–144.

WICKER, F. W., EDMONSTON, L. P., & MCCLURE, H. D. Stimulus familiarization with pictures and words in paired-associate learning. *American Journal of Psychology*, 1973, 86, 617–626.

REFERENCE NOTES

1. SNODGRASS, J. G. Toward a model for picture and word processing. Paper presented at the meeting on Processing of Visible Language, Niagara-on-the-Lake, Canada, September 1979.
2. JOHNSON, M. K., & RAYE, C. L. A working model of reality monitoring. Annual meeting of the Psychonomic Society, San Antonio, Tex., November 1978.

(Received December 26, 1979)