Memory & Cognition
1974, Vol. 2, No. 2, 224-232

Generation of a pictorial code*

PHILIP H. K. SEYMOUR
The University, Dundee, Scotland

In Experiments I and II direct measures were obtained of the time required to convert a picture or a sentence to a report of visualization or to a simple drawing. Latencies were faster for pictures than for sentences and were affected by the surface form of the sentence. In Experiment III, Ss matched pictures or sentences against a test picture under conditions of simultaneous or successive presentation. "Yes" and "No" response times were affected by sentence form under both conditions, although the difference between sentence-picture and picture-picture comparisons was virtually eliminated under the successive condition. In Experiment IV, Ss held a sentence or picture in memory over a 1,500-msec interval at the end of which a test picture or an instruction to draw the design was presented. Construction latencies were shorter in response to pictures than sentences, and there were effects of sentence form on verification time. Some implications of these results are discussed.

Comprehension of a simple locative assertion, such as "The square is inside the circle," might be exemplified in a number of ways. A person who understands such a sentence would be expected to perform successfully on such tasks as (a) drawing a picture of the scene described, or (b) determining whether or not a test picture was in semantic correspondence with the sentence. He might also be able to visualize an arrangement of objects corresponding to the sentence and to report on this experience.

It can be noted that a successful performance on a test of comprehension involves two distinct types of competence. There is firstly a question of interpretation of the intention of the utterance, or diagnosis of its function as a speech act (Searle, 1969). The sentence "The square is inside the circle" might function as an instruction to draw or visualize in one experimental situation but as a yes/no question about truth or falsity in another. Secondly, there is a question of the spatial or referential meanings of the nouns "square" and "circle" and the relational expression "inside."

In the present paper, the term *pictorial code* will be used to refer to a representation of spatial meaning that is utilized in tasks of pictorial construction, imagination, or recognition. The assumption here is that when S constructs or imagines a picture, or scans a picture to determine whether or not it satisfies some criterion, he consults a memory representation in which pictorial properties and relations are specified. The pictorial code is, in this respect, analogous to the image coding systems of Paivio (1971) and Bower (1972).

Experiments I and II of this report describe studies of the latencies of construction and visualization as responses to simple locative sentences and pictures of geometric shapes. It was expected that latencies in both tasks would be shorter when the stimulus was a picture than when it was a sentence, because access to the pictorial code is direct for pictures but involves a process of conversion or generation in the case of sentences. A second point concerned the effects of sentence structure on the latency of the conversion. Paivio and Begg (1969) have demonstrated that sentence structure may affect the latency of an imagery response to a sentence, and Huttenlocher and Strauss (1968) have shown effects of sentence form on the latency to construct the situation described in an active or passive sentence. In the present experiments, the materials used were locative sentences that asserted the position of one geometric shape as being "inside" or "outside" another. Three different surface realizations of the assertions "X inside Y" or "Y outside X" were tested: Type I. The square is inside the circle. Type II. Inside the circle is the square. Type III. The circle has the square inside. A principal aim was to determine, empirically, whether or not the three sentence types differed in their effects on response latency in the visualization and construction tasks. The general expectation was that the conversion would be faster for sentences of Type I than for sentences of Types II or III. This prediction is consistent with linguistic arguments that suggest that sentences of Types II and III are transformationally more complex than sentences of Type I (Fillmore, 1968).

EXPERIMENTS I AND II: VISUALIZATION AND CONSTRUCTION

Method

Subjects

The Ss were 16 volunteers from classes at the University of Dundee. Ten Ss served under instruction to visualize the spatial relationships represented by the sentence and picture stimuli, and six Ss under instruction to draw pictures exemplifying those relationships.

Apparatus

The sentence and picture stimuli were rear-projected by means

*This research was supported by the Social Science Research Council of the U.K. under Grant HR/1787/1. Thanks are due to Jenny Greenhalgh for her assistance in obtaining and analyzing the data.

Table 1
Summary of Mean Latencies (Milliseconds) to Close a Switch to Indicate Completion of Reading and Visualization Tasks in Experiment I

	Sentence Type			
	I	II	III	Picture
Read	727	760	786	478
Visualize	1407	1460	1590	930

of a Gaf automatic slide projector that had been modified to incorporate an electronically operated shutter. A series of timing circuits was used to present a warning tone for 500 msec, followed, after a delay of 1 sec, by exposure of the test sentence or picture. In the visualization study, the display remained on the screen for 500 msec. In the construction study, display duration was set at 1 sec. A Venner millisecond stopclock started at onset of the display. In the visualization study, the clock stopped when the S closed a switch. In the construction study, a contact between a ballpoint pen and a sheet of electrically conductive paper closed a circuit that stopped the timer. A mirror located above the S enabled the E to monitor the S's drawings.

Displays

The stimuli consisted of two pictures, showing a circle inside a square or a square inside a circle, and a set of 12 sentences. There were 4 example sentences for each of the three sentence types, in which the subject might be "square" or "circle" and the preposition "inside" or "outside." If the subject was "square," the object was "circle," and vice versa. The pictures were drawn on cards in black ink, such that the outside shape was a 5 x 5 cm square or a circle of 5 cm diam and the inside shape was a 2.5 x 2.5 cm square or a circle of 2.5 cm diam. The sentences were printed in lowercase Letraset (Univers 55, 36-point) as a horizontal array of words about 19 cm long. These displays were then photographed and printed as a white image on a dark ground as slides for projection by the apparatus. The projected display reproduced approximately the dimensions of the original. The viewing distance was about 100 cm.

Procedure

Ss taking part in the visualization study were tested under two instructions. Under a reading instruction, they were told to press the response key as soon as the sentence or picture had been scanned. The intention here was to obtain some indication of the difference in scanning time for sentences and pictures and also to determine whether sentence type affected the latency in a situation that did not necessarily involve comprehension of the sentence. Under a visualization instruction, the Ss were told to press the key as soon as they could visualize, or were otherwise aware of, the spatial arrangement of the shapes. For pictures, this meant that the contours and their relative locations should have been identified, or explicitly considered in some way. These instructions appear somewhat obscure, although the Ss did not in practice experience difficulty in understanding or applying them. Reading and visualization were tested in separate blocks of trials, half the Ss being tested on reading first and half on visualization first. Within each block, each of the 12 sentences was presented three times and each of the two pictures six times, in random orders which were independently determined for each S.

In the construction study, the Ss were instructed to draw a picture of a square-circle pattern in response to the sentence or picture stimuli. The drawings were done on a sheet of conductive paper that had been divided into 120 squares, each about 2.5 x 2.5 cm. Ss were shown examples of the two possible patterns and were told that the outer shape must be drawn so as almost to fill the square panel and that the inner shape must be considerably smaller. This was to ensure that the task could not be completed by drawing the first shape irrespective of the stimulus event and by then filling in the second shape in the inner or outer location. Care was taken, during a preliminary practice period and during the main experiment, to ensure that this instruction was obeyed and that the pattern was drawn straight off, without an appreciable pause between the two shapes. Each of the 12 sentences was presented 6 times and each of the two pictures 12 times, in random sequences independently determined for each S.

Both studies were carried out in a quiet darkened laboratory, during a test session lasting about 30-45 min. The E took a record of the response latency on each trial and, in the construction task, of the order in which the two shapes were drawn.

Results and Discussion

The data for reading and visualization will be considered first. A preliminary analysis of variance was run to establish that the order of testing on the two tasks did not affect the response latencies or interact with the main variables under study. The results for the two S groups were then combined, and have been summarized in Table 1. It is clear that latencies for both scanning and visualizing were much shorter when the stimulus was a picture than when it was a sentence. There were differences between pictures and sentences of Type I of about 250 msec for reading and of about 477 msec for visualizing. The differences were significant in both cases [$F(1,8) = 21.64$ and 25.51, $p < .01$ and .001]. There was also a significant effect of sentence type on both reading and visualization latencies [$F(2,16) = 6.48$, $p < .01$]. This effect did not interact with tasks, and follow-up tests confirmed that it was significant under both instructions. The difference in latency between reading and visualizing was 729 msec [$F(1,8) = 61.06$, $p < .001$]. The latency difference between reading and visualizing was about 450 msec for picture stimuli. This effect appears smaller than the one for sentences, but was also significant [$F(1,8) = 51.28$, $p < .001$].

These data should not be overinterpreted, in view of the lack of experimental control over the criterion adopted by the S for closure of the switch under the two instructions. However, if the latencies are treated as a valid indicator of completion of scanning and onset of visualization, the following conclusions may be drawn: (1) Access to awareness of contour identities and spatial relations is faster, by nearly half a second, for picture stimuli than for sentence stimuli. (2) Access to visualization is affected by sentence structure, principally because sentences of Type III are associated with a substantial delay of response relative to sentences of Type I. (3) Both of these effects appear to be partly a matter of differences in reading or scanning time, since sentences are scanned less rapidly than pictures, and sentence structure affects reading time. However, it appears unlikely that the effects on visualization may be reduced to differences in scanning time, since the picture-sentence difference is about twice as great for visualization as for reading and because the difference in means between sentences of Types I and III is about three times as great for visualization as for reading.

Table 2 summarizes the latency data for the construction task. The drawing response was started about 585 msec sooner when the stimulus was a picture than when it was a sentence of Type I [$F(1,5) = 87.62$, $p < .001$]. There was also a significant effect for sentence type [$F(2,10) = 6.98$, $p < .025$], which was primarily attributable to a difference of nearly 200 msec between the means for Type I and Type II sentences. The preposition used in the sentence ("inside" or "outside") did not affect the construction latency, except in the case of sentences of Type I, where "outside" sentences were responded to about 144 msec faster than "inside" sentences [$F(1,5) = 26.55$, $p < .01$].

A record was also taken of the order in which the two shapes were drawn. Table 3 shows the frequencies with which drawings were started with the inside or outside shape for "inside" and "outside" sentences of the three types and also for the pictures. The order of construction in response to pictures showed marked intra-individual consistency: three Ss always started with the outside shape, two with the inside shape, and one S used both starting positions about equally often. It seems clear from Table 3 that sentences of Type I constrained Ss to draw the shape occupying the location specified by the preposition used in the sentence, since only 5 out of a total of 144 drawings went against this tendency. The difference in frequency of compliant and noncompliant drawings was highly significant for Type I sentences [$F(1,5) = 547.44$, $p < .001$], but not for sentences of Types II and III [$F(1,5) = 1$ and 5.94].

The data for the construction task are similar to those for the visualization task in showing a marked latency difference between picture and sentence stimuli and an effect for sentence type. If we define the availability of the pictorial code in terms of capacity to make a correct drawing, the data may be taken as confirming that this code is accessed more rapidly by a picture than by a sentence and more rapidly by a sentence of Type I than by a sentence of Type II or III.

Table 2
Summary of Mean Latencies (Milliseconds) to Start Construction of a Drawing in Response to a Sentence of Type I, II, or III or a Picture (Experiment II)

Preposition	Sentence Type I	II	III	Picture
"Inside"	1336	1452	1422	679
"Outside"	1192	1467	1490	
Mean	1264	1460	1456	679

The constructional response to a sentence might be achieved in a number of ways. The sentence might be converted to a pictorial code, which could be translated into a plan for drawing. Generation of the pictorial code would depend on prior interpretation of the sentence, and the time to complete this interpretation might be expected to vary with transformational complexity. Alternatively, the constructional plan might derive directly from a "deep structure" representation of the propositional content of the sentence. According to both hypotheses, the distinction between sentences of Type I and sentences of Types II and III arises during the process of interpretation and ceases to be important once the abstract propositional or pictorial representation has been established. If the pictorial representation of a sentence is quite like the pictorial description of a picture, as is implied by the dual coding models of Paivio (1971) and Bower (1972), and the drawing depends on a pictorial representation, we would expect order of construction to be similar for sentence and picture stimuli. This prediction is not supported by the data, since order of construction was consistent but idiosyncratic for pictures, constrained by the preposition for sentences of Type I, and generally inconsistent for sentences of Types II and III. These results are also difficult to reconcile with the view that the sentence is eventually represented as a proposition of the form: "X inside Y" or "X outside Y" (cf. Clark & Chase, 1972). If the drawing was an externalization of one of these two basic assertions and the relational term specified the

Table 3
Summary of Frequencies of Starting Construction at Inside or Outside Location (I or O) for Pictures and "Inside" and "Outside" Sentences of Types I, II, and III (Experiment II)

	Sentence I				Sentence II				Sentence III				Picture	
	Inside		Outside		Inside		Outside		Inside		Outside			
	I	O	I	O	I	O	I	O	I	O	I	O	I	O
S1	12	0	1	11	10	2	3	9	12	0	0	12	24	0
S2	12	0	0	12	3	9	0	12	3	9	1	11	0	24
S3	12	0	0	12	2	10	9	3	6	6	5	7	12	12
S4	9	3	0	12	1	11	7	5	8	4	3	9	0	24
S5	11	1	0	12	11	1	0	12	12	0	1	11	23	1
S6	12	0	0	12	2	10	8	4	6	6	6	6	0	24
Total	68	4	1	71	29	43	27	45	47	25	16	56	59	85

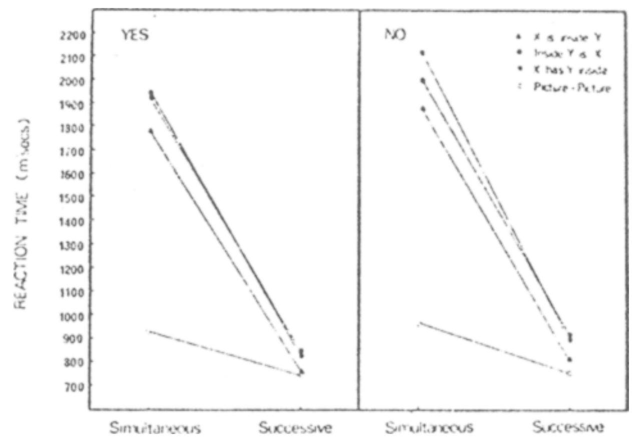

Fig. 1. Summary of mean latencies (milliseconds) of "Yes" and "No" responses to simultaneous and successive picture-picture and sentence-picture combinations (Experiment III).

order of construction, the preposition would be expected to control order for all three sentence types.

EXPERIMENT III: VERIFICATION

Posner, Boies, Eichelman, and Taylor (1969) introduced an experimental methodology that proved useful in distinguishing between nominal and visual memory codes for letters. Using a "same"-"different" judgmental task, they demonstrated that a difference in reaction time for classification of simultaneously displayed nominally or physically identical letters might be eliminated when the letters were successively presented. If the S was certain about the case of the second letter, the function of reaction time against interstimulus interval was suggestive of *retention* of visual information when the first event was a visual letter and of *generation* of a visual code when the first event was a spoken letter name. The methodology used by Posner may be adapted to test some implications of the hypothesis that a sentence may be converted to a picture-like representation. Under a simultaneous presentation condition, a physical identity match between two adjacent pictures will have a shorter reaction time than a semantic identity match between a sentence and a picture. Since sentence structure affects conversion time, sentence-picture comparisons should be faster for sentences of Type I than for sentences of Types II or III. Under conditions of successive presentation, when the first event may be either a sentence or a picture and the second is always a test picture, S may retain a pictorial memory representation of the picture and generate such a representation from a sentence. In this case, there should be no difference in reaction time between picture-picture and sentence-picture comparisons and no effect of sentence type on sentence-picture comparisons. These predictions of the pictorial coding hypothesis were tested in Experiment III.

Method

Subjects

The Ss were 10 volunteers from senior classes at Dundee secondary schools.

Apparatus

This was the same as for Experiments I and II, apart from the inclusion of a second projector. Under the simultaneous condition the shutters of both projectors were opened simultaneously for 2 sec and the timer was started 1 sec after offset of the warning tone. Under the successive condition, the first projector shutter was opened for 1 sec and there was then a delay of 1 sec, at the end of which the timer started and the shutter of the second projector opened for 1 sec. The timer stopped when the S's vocal "Yes" or "No" response closed the relay of a voice key.

Displays

The sentence and picture stimuli were the same as those used in Experiments I and II.

Procedure

The experiment was run as two blocks of 96 observations. These were held on different days, separated by about 1 week. The simultaneous and successive conditions were examined on different days, half the Ss receiving the one first and half the other. Within each session, the 12 sentences occurred three times with a matching picture and three times with a mismatched picture. Each of the four possible picture-picture combinations occurred six times. The order of presentation was randomized independently for each S under each condition, apart from the constraint that each sentence-picture and picture-picture pair should occur with equal frequencies in each of three blocks of 32 trials per session. Ss were instructed to respond by calling out "Yes" to equivalent picture-picture or sentence-picture combinations and "No" for different combinations, and were given some practice on the task before testing started. When errors occurred, or an observation was lost on account of failure of the voiceswitch, the trial was repeated later in the random sequence.

Results and Discussion

Errors were infrequent in this experiment and occurred on only 1.8% of trials.

A preliminary analysis of variance of the latency data was carried out to determine the effects of order of scheduling of the simultaneous and successive conditions. Separate analyses of the "Yes" and "No" response times produced no significant differences between the S groups and no significant interactions with the main variables under consideration. The data for the two groups were then combined, and are presented in summary form in Fig. 1.

A main prediction of the pictorial coding hypothesis about sentence comprehension is that sentence structure should affect conversion time but not verification time. Hence, an effect of sentence type is expected under the simultaneous condition, but not under the successive condition. This prediction was not supported by the data. There were significant sentence effects on both "Yes" and "No" response times [$F(2,16) = 15.75$ and 10.78, $p < .001$ and .01], but the interactions of these effects with simultaneous vs successive conditions were not significant [$F(2,16) = 1.63$ and 2.49]. As a check on this conclusion, separate analyses of the simultaneous and successive conditions were carried out, in which response and sentence type were factors. The sentence effect was significant in both cases [$F(2,16) = 9.94$ and 8.61, $p < .01$], and there was also a significant response effect [$F(1,8) = 7.95$ and 28.31, $p < .025$ and .001], although the Sentence by Response interactions were not significant. The sentence effect occurred because sentences of Type I were verified about 170 msec faster than sentences of Types II and III under the simultaneous condition and about 80 msec faster under the successive condition. The response effect occurred because "Yes" responses were faster than "No" responses by 130 msec under the simultaneous condition and by 72 msec under the successive condition.

A second prediction of the pictorial coding hypothesis is that sentence-picture and picture-picture comparison times should differ under the simultaneous condition but should be equivalent under the successive condition. An assumption here is that the successive picture-picture condition involves retention of a pictorial representation of the first picture. Although this point is difficult to establish, it is given some support by the observation that picture-picture comparisons were faster by 184 msec under the successive condition than under the simultaneous condition [$F(1,8) = 30.76$, $p < .001$], suggesting that there was no loss or decay of pictorial information. It is also worth noting that the "Yes"/"No" difference of about 20 msec observed for simultaneous and successive picture-picture comparisons was not significant [$F(1,8) = 1.36$]. The picture-picture data were then compared with the data for sentences of Type I, in separate analyses of "Yes" and "No" responses. In both cases, there were significant differences between the simultaneous and successive conditions [$F(1,8) = 158.20$ and 190.38, $p < .001$] and between picture-picture and sentence-picture combinations [$F(1,8) = 52.02$ and 82.29, $p < .001$], and a significant interaction between these factors [$F(1,8) = 37.06$ and 62.21, $p < .001$]. It is obvious from Fig. 1 that these effects and the interactions occurred because the difference between picture-picture and sentence-picture comparisons was much greater under the simultaneous than under the successive condition. Subsidiary analyses of the data for the successive condition confirmed that the differences between pictures and sentences of Type I were not significant for "Yes" or "No" responses [$F(1,8) < 1$ and 3.78]. On Posner's criteria, this equivalence of picture-picture and sentence-picture comparison times suggests conversion of the sentence to a generated pictorial code.

The outcome of this study may be compared with the findings of earlier experiments reported by Gough (1965, 1966), Cohen (1969), Trabasso, Rollins, and Shaughnessy (1971), and Seymour (1973). All of these authors have reported effects of sentence form on successive sentence-picture comparisons, at offset-onset intervals ranging from 0 to 5 sec. The present study confirms this result for the 1-sec delay interval. It also indicates that an effect of sentence type may coexist with the elimination of a difference between sentence-picture and picture-picture comparison times. Thus, a period of 2 sec from sentence onset is sufficient to eliminate a 900-msec difference between simultaneous picture-picture and sentence-picture comparisons, but not a small effect of sentence structure on verification time.

EXPERIMENT IV: VERIFICATION AND CONSTRUCTION

The results of Experiments I and II showed that Ss required about 1,500 msec to report visualization of the arrangement defined by a sentence and about 1,400 msec to start a drawing of that arrangement. This suggests that conversion to pictorial format may be achieved within 1,500 msec or less of sentence onset. Under the successive condition of Experiment II, a period of 2,000 msec elapsed before presentation of the test picture. This period should be sufficient for conversion to a pictorial representation, and the similarity of the sentence-picture and picture-picture comparison times suggests that this may have occurred. However, the residual effect of sentence type is difficult to reconcile with the view that the different sentence types are converted to a common pictorial or linguistic representation. It is, of course, possible that effective elimination of the sentence effect would occur at a slightly longer interval, although the studies of Gough (1966) and Cohen (1969) provide evidence counter to this view. In a pilot study, in which the delay between sentence offset and picture onset was set at 100, 750, or 1,500 msec, the effect of sentence form was not significant at the longest interval. It was considered worthwhile, therefore, to replicate the successive condition of Experiment III using a 1,500-msec delay.

In order to obtain more information about the nature of the representation of the sentence at this delay, the construction and verification tasks were combined in Experiment IV. The S was shown a sentence of Type I, II, or III or a picture for 1 sec, and was tested after a delay of 1,500 msec by a picture or by an instruction to make a drawing. If the test event was a square-circle picture, he made a vocal "Yes" or "No" response. If a dot appeared on the screen, he made a drawing of a square-circle picture. According to the pictorial coding hypothesis, latency of response on both tasks should be the same for sentences and for pictures, and there should be no effect of sentence type.

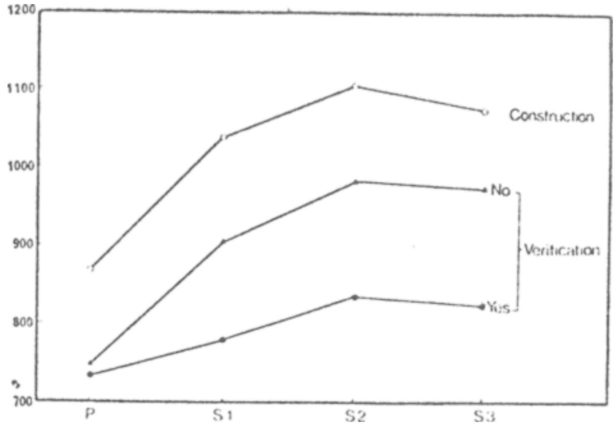

Fig. 2. Summary of mean latencies (milliseconds) on the delayed construction and verification tests of Experiment IV.

Method

Subjects

The Ss were eight volunteers, mostly research workers and others connected with the Department of Psychology at the University of Dundee.

Apparatus

The two Gaf projectors were set up so that, following a warning tone and a 1-sec delay, the shutter of the first projector was opened for 1 sec. There was then a delay of 1,500 msec, at the end of which the shutter of the second projector was opened for 1 sec and the Venner stopclock was started. The clock was stopped either by closure of the voice key relay or by a contact between the ballpoint pen and an electrically conductive sheet of paper.

Displays

The sentence and picture displays were the same as those used in the earlier experiments. A small white dot on a dark ground was used as a signal to the S that he should respond by drawing a picture.

Procedure

All Ss were instructed that they would see a sentence or picture, followed after a short interval by a test picture or a dot. If a picture was presented, they were to report "Yes" if it matched the earlier sentence or picture and "No" if it did not. If the dot was presented, they should draw a square and circle to correspond to the earlier picture or sentence. The instructions about the manner in which the drawing should be made were the same as for Experiment I. Ninety-six observations were made on each S, in which picture tests and drawing tests were randomly intermixed. Thirty-two trials required a drawing response and involved two observations on each of 12 sentences and four observations on each of two pictures. The remaining 64 trials involved picture tests. Each sentence occurred four times, twice with a matching picture and twice with a mismatched picture. There were two matched and two mismatched picture-picture combinations, and each of these occurred four times. The sequences of presentation were randomizations, which were independently determined for each S. Error responses in the verification task were discarded and replaced.

Results and Discussion

Figure 2 presents a summary of the mean latencies, averaged over Ss, to initiate construction of the square-circle pictures or to report "Yes" or "No" in the verification task. The results for construction will be considered first. A preliminary analysis was carried out to test for differences among picture stimuli and the three sentence types. This showed a significant effect [$F(3,21) = 8.15$, $p < .001$], which appears on inspection to be principally attributable to a difference between sentence and picture stimuli. A subsequent analysis indicated that differences among the sentence types were not significant [$F(2,14) = 1.01$], whereas the difference of about 170 msec between pictures and sentences of Type I was significant [$F(1,7) = 11.71$, $p < .025$].

For pictures, the results for order of construction were quite similar to those of Experiment I. Four Ss always drew the outside shape first, and two Ss always drew the inside shape first. The remaining two Ss drew the inside shape first on five and six out of eight trials. Table 4 shows the proportion of occasions on which construction order in response to a sentence agreed with the preposition used in the sentence. For sentences of Type I there was a significant tendency for order of construction to follow the preposition [$F(1,7) = 8.79$, $p < .025$], although the effect was not significant for sentences of Types II or III [$F(1,7) = 4.43$ and 0.22].

The verification data were submitted to an analysis of variance in which "Yes" vs "No" responses and types of memory stimulus were factors. There was a significant response effect [$F(1,7) = 21.38$, $p < .01$], and a significant effect for types of stimulus [$F(3,21) = 14.82$, $p < .001$]. Figure 2 suggests an interaction of these effects, but this was not in fact significant [$F(3,21) = 2.42$]. A test for differences among sentence types showed a significant effect [$F(2,14) = 6.34$, $p < .025$], principally because sentences of Type I led to faster classification times than did sentences of Types II and III. The difference of 140 msec between "Yes" and "No" responses was significant [$F(1,7) = 16.49$, $p < .01$], and independent of the sentences effect [$F(2,14) < 1$]. For "Yes" responses there was a nonsignificant difference of 40 msec between pictures and sentences of Type I [$F(1,7) < 1$]. This difference was about 160 msec for "No" responses, and was significant [$F(1,7) = 14.96$, $p < .01$].

The results of this experiment appear contrary to predictions derived from the assumption that the pictorial code for a sentence is essentially equivalent to the code for a picture. Construction of a square-circle picture was faster if a picture was in memory than if a sentence was in memory. The form of the sentence did not affect time to start construction, but did influence the order of construction. "Yes" responses to test pictures were about equivalent when a picture or a sentence of Type I was in memory, but were slightly

Table 4
Total Number of Drawings Started at Inside and Outside Location (I or O) for Pictures and "Inside" and "Outside" Sentences on the Delayed Construction Test of Experiment IV

Sentence I				Sentence II				Sentence III				Picture	
Inside		Outside		Inside		Outside		Inside		Outside			
I	O	I	O	I	O	I	O	I	O	I	O	I	O
21	11	7	25	8	24	19	13	9	23	12	20	27	37

delayed for sentences of Types II and III. "No" responses to test pictures were faster if the test was matched against a picture memory than if it was matched against a sentence memory, and were also affected by the form of the sentence. Finally, in verification, the size of the "Yes"-"No" difference again appeared greater for sentence-picture than for picture-picture comparisons.

GENERAL DISCUSSION

The experiments were conducted with the aim of testing some implications of the hypothesis that comprehension of a locative sentence may involve a conversion to a picture-like representation. The object of Experiments I and II was to demonstrate a difference between sentences and pictures and among sentences of three surface forms, as stimuli for the responses of visualization and construction. The results for the simultaneous condition of Experiment III showed that similar differences were obtained in the verification task. The delayed verification and construction tests of Experiments III and IV were then used to test the hypothesis that the effects observed in the immediate test conditions might be eliminated when S was given a short delay in which to recode the sentence stimulus.

The strong prediction of the pictorial coding hypothesis is that the sentence-picture difference, and the effect of sentence type, will be eliminated under conditions of delayed testing. This outcome would be expected if performance on the various tasks utilized a representation of pictorial properties which could be accessed directly by a picture and, indirectly, by spatially synonymous sentences of different surface form. The data are not consistent with this hypothesis, because sentences and pictures differed in their effects on order and latency of construction in the immediate and delayed tests of Experiments II and IV and because there were effects of sentence type on the delayed verification tests of Experiments III and IV. Sentence-picture and picture-picture comparisons were also distinguished in terms of magnitude of the "Yes"/"No" difference under both simultaneous and successive presentation conditions.

An alternative position, which would be consistent with the proposals of Bower (1972) and Paivio (1971), would be that pictures are retained in pictorial format, whereas sentences are retained in a verbal/semantic memory. The different sentence types might be held in a form which preserved the surface order of their components or might converge on one or other of the locative assertions, "X inside Y," "X outside Y." On the latter view, one would expect a difference between sentences and pictures, but no effect for sentence type in the delayed test situations. A result of this kind was obtained for the delayed construction tests in Experiment IV, but not for the delayed verification tests in Experiments III and IV. Both experiments showed effects of sentence type on "Yes" and "No" reaction times. Further, there was no difference between pictures and sentences of Type I in Experiment III, although there was an effect on "No" responses in Experiment IV. In addition, sentences of Type I constrained order of construction in both the immediate and delayed test situations. These results suggest that the sentence memory preserves the surface ordering of the original and cast doubt on the view that sentences and pictures are stored in memories that differ in accessibility with respect to a test picture presented in a verification task. The dual memory position might be salvaged by arguing that sentences of Type I are recoded pictorially, whereas sentences of Types II and III are held in the verbal memory. However, the difference between pictures and Type I sentences obtained in the delayed construction task is not consistent with this view.

A third position may be stated, in which the distinction between a verbal and pictorial memory is dissolved, and replaced by the notion of a general purpose system for construction and storage of plans of action (Miller, Galanter, & Pribram, 1960). The S's interpretation of the experimental task and instruction may be viewed as a TOTE unit which is held in readiness throughout the experimental session. The procedure starts when a test for presence of the display is positive and halts when a response is made. The operate phase of the TOTE may be analyzed into subprocedures for acquisition of information from the display, conversion to an output plan, and initiation of the execution of that plan. Most of the results of the present set of experiments can be accommodated by assuming that S has standardized procedures for scanning the sentence and picture displays and that these serial acquisition procedures interact with the spatial tarrangement of the information on the display to determine the precise form of the conversion procedure and output plan.

In the case of construction, onset of the display will result in the start of an acquisition procedure that involves a left-to-right scan if the display is a sentence and an inside-out or outside-in scan if it is a picture. The picture-scan procedure is probably standardized within an individual S, but may differ from one S to another. The assumption of left-right scanning of the sentence is no doubt an oversimplification, but is adopted here for convenience in presenting the argument. Construction can go ahead, in this particular task context, as soon as one of two possible contours has been correlated with one of two possible locations. This presents little difficulty when the display is a picture, since S always checks a given location first and always starts his drawings at that location. Further, the relevant contour is directly specified as a line or curve on the display. Thus, pictorial information is readily available for entry into an output subprocedure in which the location at which the construction is to start has been determined in advance. For these reasons, copying of a two-shape picture is a fast and efficient operation, characterized by consistency of order of construction.

Constructional responses to sentences of Type I involve the conversion of an array of symbols, of the form: "The X is inside/outside the Y" to a sequentially achieved two-shape drawing. Under immediate test conditions, the latency on this task will be longer than copying of a picture because the sentence takes longer to scan, contour properties are not directly specified but must be retrieved from memory, and the form of the output plan is not standardized but must be established during analysis of the sentence. Hence, performance is characterized by a delay of response relative to copying and by an order of construction which is controlled by the preposition. S therefore derives contour information from the subject of the sentence, and locative information from the preposition, and enters these values into the first stage of a two-step output plan. Since responses in the delayed construction test of Experiment IV were slower when a Type I sentence was in memory than when a picture was in memory, it seems likely that contour information is retrieved at onset of the test and not during the delay. However, the delay of 1,500 msec from sentence offset is probably sufficient for S to sort out the identity of the shape he is to draw and its location. The data suggest that this information is held in a relatively abstract form, and is translated into a more concrete or pictorial form only when the constructional plan is initiated. Presumably a picture can be retained pictorially, so that contour information is available for immediate entry into the constructional plan.

Sentences of Types II and III caused delays in response on the immediate construction test of Experiment II, but not on the delayed test of Experiment IV. The data on reading latencies obtained in Experiment I suggest that this might be partly a matter of scanning time. However, sentences of Type II differ from sentences of Type I in two important respects: (1) the more complex form involves a separation of the subject of the sentence and the locative preposition, and (2) the noun adjacent to the preposition does not occupy the location specified by the preposition. In transforming "Inside the Y is the X" S must opt for one of two possible procedures. He may accept the location specified by the preposition as defining the place at which construction will start, and then proceed through the sentence, rejecting Y but accepting X as the shape to occupy that location. Alternatively, he may focus on the constituent "Inside the Y," opt to construct Y first, and then reverse the specification of location. Similar ambiguities are inherent in sentences of Type III. The array "The X has the Y inside" may be transformed by opting to draw X first, in which case the locative must be altered, or by accepting the location and establishing that Y, and not X, must be drawn first.

These comments suggest that a locative sentence may be represented as a constructional plan or TOTE and that there is a procedural ambiguity in the representation of a sentence of Type II or III, but not of a sentence of Type I. A Type I sentence may be represented as a two-stage TOTE, of the form: [*First* (Draw X at L_1) *Then* (Draw Y at L_2)]. This form of representation results in consistency of output order on both immediate and delayed construction tests. Sentences of Types II and III may be represented as a disjunction: [*First* (Draw X at L_1) *Or* (Draw Y at L_2)] *Then* [*If* X at L_1 (Draw Y at L_2)] *Or* [*If* Y at L_2 (Draw X at L_1)]. This type of representation is associated with variability of output order, and some delay of response on the immediate construction tests, which may be an indication of the time required to resolve the ambiguity. The same line of explanation is, in principle, applicable to the visualization data of Experiment I, if visualization is thought of as internalized execution of a constructional plan.

In the case of verification, S must set up a TOTE for scanning and testing of the contours of the test picture. The results of Experiment III suggest that this procedure may be set up very much more rapidly when the test is against a picture than when it is against a sentence, but that TOTEs derived from pictures and sentences of Type I are, once established, quite similar. That this is so is suggested by the equivalence of reaction times for picture-picture and sentence-picture comparisons found in Experiments III and IV, and in a comparable experiment by Seymour (1973). Thus, one might think of the representations of pictures and Type I sentences as both being of theform: [*First* (Test for contour X at L_1) *Then* (Convert outcome to response)]. If the test operated on relatively abstract representations of locative and configurational properties there might be no advantage in having concrete information about contours in memory. The effect of sentence type on verification time may then depend on the procedural ambiguity of sentences of Types II and III, assuming that these are represented in memory as a disjunction of alternative TOTEs: [*First* (Test for X at L_1) *Or* (Test for Y at L_2) *Then* (Convert outcome to response)]. The effect of sentence type on verification time, under both simultaneous and successive conditions, is a reflection of the time required to resolve the ambiguity. This might be a matter of deciding which of two search procedures to apply, or which of two components in the sentence representation to use in testing input from the test picture.

The data of Experiments III and IV indicated that the magnitude of the "Yes"/"No" effect was greater when the stimulus in memory was a sentence than when it was a picture, although this effect was independent of sentence type. This result suggests that the threshold for selection of a "No" response is raised if the item in memory is a sentence. Thus, for verification, sentences and pictures are distinguished in terms of their effects on a response selection stage of the process, rather than in terms of the nature of the memory representation which may be used in making the comparison.

With respect to the notions of retention and generation of a pictorial code, the following conclusions appear warranted: (1) Pictures of simple designs may be retained over a short interval both as an image-like representation of contour and as a more abstract description of identities and relative locations. The explicit representation of contour is useful in setting up a constructional plan, but less so in verification, which depends on the more abstract description. (2) A locative sentence is retained as an abstract description of identities and relative locations, and not as a generated image of contours. The logical structure of the memory representation is related to the surface form of the sentence from which it was derived. This representation is at an appropriate level of abstraction for verification, but must be converted to a more concrete form for construction.

REFERENCES

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

Clark, H. H., & Chase, W. G. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3, 472-517.

Cohen, G. Pattern recognition: Differences between matching patterns to patterns and matching descriptions to patterns. Journal of Experimental Psychology, 1969, 82, 427-434.

Fillmore, C. J. The case for case. In E. Bach and R. Harms (Eds.), *Universals in linguistic theory.* New York: Holt, Rinehart & Winston, 1968.

Gough, P. B. Grammatical transformations and speed of understanding. Journal of Verbal Learning & Verbal Behavior, 1965, 4, 107-111.

Gough, P. B. The verification of sentences: The effects of delay of evidence and sentence length. Journal of Verbal Learning & Verbal Behavior, 1966, 5, 492-496.

Huttenlocher, J., & Strauss, S. Comprehension and a statement's relation to the situation it describes. Journal of Verbal Learning & Verbal Behavior, 1968, 7, 300-304.

Miller, G. A., Galanter, E., & Pribram, K. H. *Plans and the structure of behavior.* New York: Holt, Rinehart & Winston, 1960.

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, 1969, 10, 408-412.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. Retention of visual and name codes of single letters. Journal of Experimental Psychology Monographs, 1969, 79, 1-16.

Searle, J. R. *Speech acts.* Cambridge: Cambridge University Press, 1969.

Seymour, P. H. K. Pictorial coding of verbal descriptions. Quarterly Journal of Experimental Psychology, 1973, in press.

Trabasso, T., Rollins, H., & Shaughnessy, E. Storage and verification stages in processing concepts. Cognitive Psychology, 1971, 2, 239-289.

(Received for publication July 13, 1973;
accepted August 17, 1973.)