

Structural descriptions and the limitations of visual images*

STEPHEN K. REED

Case Western Reserve University, Cleveland, Ohio 44106

Two experiments are reported that require Ss to judge whether or not the second of two sequentially presented patterns is a part of the first pattern. The results suggest that Ss code the pattern as a structural description and find it difficult to recognize a part of the pattern which does not match the description. It is proposed that a structural description is a combination of visual and verbal codes and that visual images lack detail when not supported by verbal concepts.

Visual imagery has been revived as a topic of current interest but attention has focused more on the usefulness of imagery than on its limitations. There is, indeed, overwhelming evidence that imagery provides an effective mnemonic code (Paivio, 1971), but it is less clear to what extent an image preserves the details of a stimulus pattern. For example, Bahrick and Boucher (1968) have argued that visual images can help Ss recall the names of objects even if the images lack many of the details of the actual objects. The implication of Bahrick and Boucher's argument is important because it suggests that visual images would not have to be very accurate in order to serve as mnemonic aids in many memory experiments.

A second reason why it is difficult to study the accuracy of visual images is that language can also provide an effective memory code. If I perceive the pattern in Fig. 1 as two overlapping triangles, words such as "triangle," "inverted," and "bisect" can help me to remember the pattern. According to the verbal loop hypothesis proposed by Glanzer and Clark (1964), Ss in a perceptual task translate input information into words, store the words, and utilize them in responding. One consequence of the hypothesis is that the complexity of a stimulus pattern should be determined by the number of words needed to describe it. By using overt descriptions to estimate the length of verbalizations, Glanzer and Clark were able to account for two measures of complexity of line drawings. The accuracy of reproducing a pattern following a brief exposure resulted in a correlation of 0.81 between difficulty and length of verbalization. A correlation of 0.87 was obtained between length of verbalization and judged complexity. Although these correlations are of interest, they give little direct evidence for the relative use of visual and verbal memory codes.

*This research was supported by NIMH Grant MH-21115. Preliminary work was begun while the author was a NIMH postdoctoral fellow at the University of Sussex. I would like to thank Austin Abramson and Mark Westerhaus for their assistance in testing Ss. Requests for reprints should be sent to Stephen K. Reed, Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106.

STRUCTURAL DESCRIPTIONS

The notion of a pattern description has been maintained in many of the more recent theories, but some theorists have been neutral with respect to the modality of the description. Simon (1972) has recently pointed out that the principal theories that explain performance in tasks involving sequences of symbols agree in proposing that Ss form pattern descriptions from the sequences. Simon argued that, although the descriptions incorporate the relations between symbols, the repetition of subpatterns, and hierarchic phrase structure, there is nothing in the data to suggest that the descriptions are verbal. He recommended taking a more neutral stance toward the modality of the description (Simon, 1972, p. 380).

Sutherland (1968) has proposed a theory of visual pattern recognition in which structural descriptions play a predominant role. The theory has three parts: (1) a processor that extracts local features from the input picture, preserving information about spatial relationships between the features; (2) a mechanism that produces an abstract description of the output from the processor; (3) a store in which such descriptions are held. When a picture is memorized, a rule is written into a store describing the output from the processor in highly abstract language. Such a language would express how the features are interrelated and would allow for the possibility of segmenting a picture in different ways. Recognition of a picture occurs by matching the output from the processor to a stored description. What we see is determined by the descriptive rule selected to describe the pictures, so that a person is unable to respond to details of a picture not represented in the rule. Like Simon, Sutherland emphasized the formal nature of the descriptive rule, rather than its modality. Since the basic theory was developed to apply to animals and man, the description would presumably be a nonverbal one in the former case.

Other theorists (Hochberg, 1970; Noton, 1970) have advocated structural descriptions as a means for directing visual scanning. Hochberg has argued that, as the structure of linguistic expectancies provides the basis

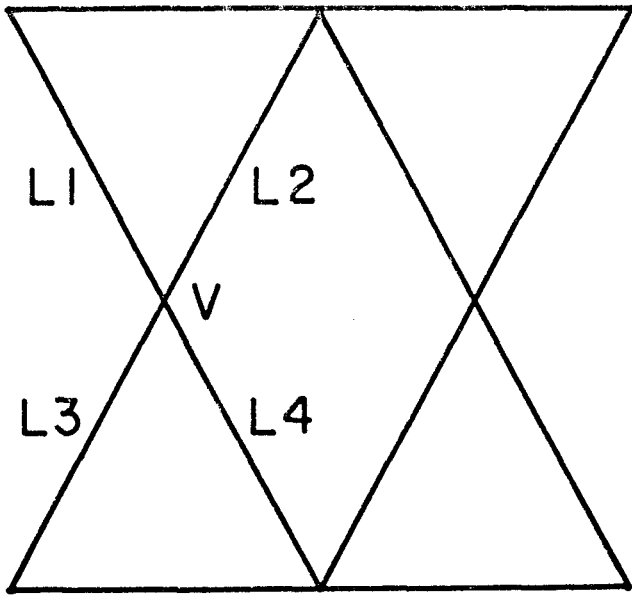


Fig. 1. An ambiguous pattern.

for selective perception of speech, the organization of visual-motor expectancies provides the basis for the selective perception of visual form. Noton (1970) proposed a similar theory, emphasizing that patterns are stored in memory in the form of a feature network recording the attention shifts that are required to pass from one feature to another. According to Noton, the order of scanning a picture should reveal the structure of the feature network, since the theory predicts that Ss use a consistent scanpath when becoming familiar with a pattern. Once again, the modality of the feature network—whether visual, verbal, or kinesthetic—was not emphasized in the theory.

The present formulation continues in this tradition by emphasizing the formal aspects of structural descriptions rather than their modality. A theory of *structural descriptions* can be formalized by specifying a pictorial language or grammar to describe the patterns. According to Narasimhan (1969), a pictorial language can be defined as $G = G(P, A, R, C, T)$, where G is specified by a set of primitives, P ; a set of attributes, A ; a set of relations, R ; a set of compositional rules, C ; and a set of transformations, T . I have attempted elsewhere (Reed, 1973) to develop a grammar to describe the patterns shown in Figs. 2 and 3, but the discussion here will be limited to a few main points concerning the ambiguous pattern shown in Fig. 1. The pattern can be seen in a number of different ways; for example, as two overlapping triangles, two overlapping parallelograms, two adjacent hour-glasses, or a diamond contained in a larger hour-glass. Which description the viewer forms is determined by the decisions he makes regarding how lines should be joined. Perceiving a triangle requires joining Lines L2 and L3, whereas perceiving a diamond requires joining Lines L2 and L4.

RECOGNIZING PARTS

A method for testing what types of structural descriptions are formed by Ss would be to require them to judge whether the second of two sequentially presented patterns is a part of the first pattern. Figure 2 shows a variety of geometric patterns and their parts, i.e., Patterns 1A-1E are considered parts of Pattern 1. If Pattern 1 was encoded as two overlapping triangles, it should be easier for Ss to recognize a triangle as a part of Pattern 1 than a parallelogram. The relative difficulty of recognizing alternative parts of a pattern should provide information regarding the kind of structural descriptions stored in memory.

Although the analysis of a pattern into its various subpatterns is a likely candidate for the memory code, it is not the only possible candidate, since the pattern could be stored as a visual template. Smith (1968) has pointed out that the concept of a template has never been very adequately described by psychologists, but the term usually refers to some unanalyzed whole, usually in the form of a visual image so that a match can be determined by a point-to-point isomorphism between

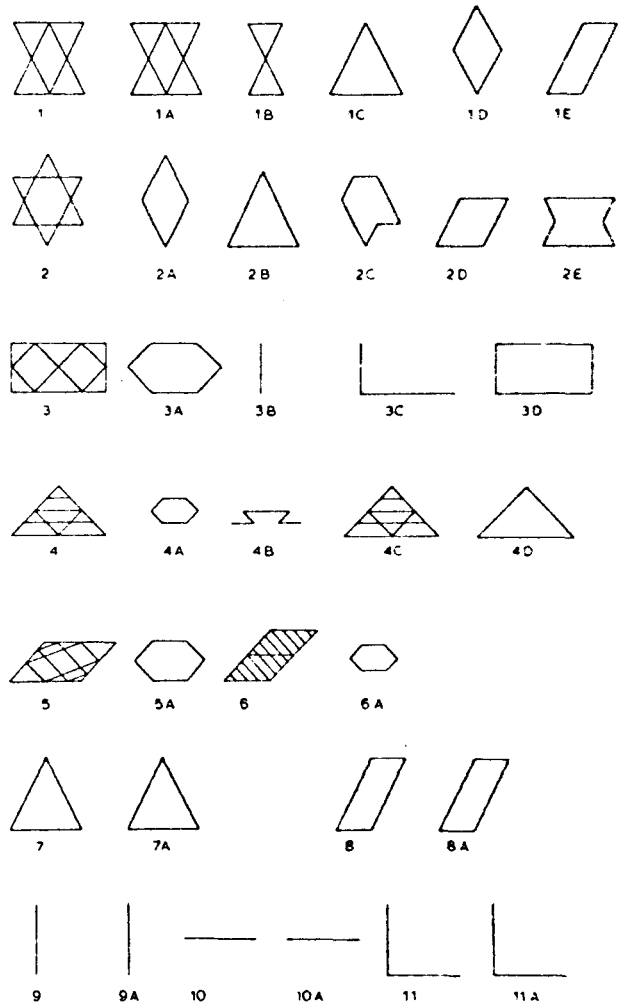


Fig. 2. Positive pairs of patterns used in Experiment I.

stimulus and template. It would, therefore, seem impossible to describe a template, since a description usually implies some degree of analysis (in terms of features or primitives) which serves as the basis for the description.

It is my own bias that, whatever is stored in memory requires a description, but there is one way in which we can define a *template description* for the patterns shown in Fig. 2. A template description requires the specification of features or primitives (although feature and template theories are usually contrasted, cf. Gibson, 1963) and in fact utilizes the grammar *G* defined earlier. For our purposes, the difference between a template description and a structural description is that a structural description describes a pattern in terms of subpatterns (requiring a hierarchical grammar in the formulation of Reed, 1973), whereas a template description does not. More specifically, a template description would describe Lines L1, L2, L3, and L4 as joining at vertex V in Fig. 1 without specifying which lines are joined to form subpatterns. Both a template description and a structural description use the same primitives and the relationship "join," but the template description is uncommitted to a particular interpretation of the pattern.

There may in fact be some advantage in trying to remember the pattern as a whole. Ss storing the pattern as a structural description would have an advantage in recognizing parts that were represented in the structural description but would be at a disadvantage in recognizing parts that were not represented in the structural description. Ss trying to maintain an accurate visual image of the whole pattern might be able to scan the image for any possible part to determine whether it contained that part. Their success would, of course, depend upon their ability to maintain a complete and accurate visual image.

The literature on the decay rates of visual images is not easy to interpret. Recognition accuracy for pictures is very good after long retention intervals (Shepard, 1967; Standing, Conezio, & Haber, 1970) and it is likely that visual imagery is involved, at least in part. Other investigators (Posner, Boise, Eichelman, & Taylor, 1969; Smith & Nielsen, 1970; Phillips & Baddeley, 1971) have reported a decay in visual images over an interval of a few seconds. These latter investigations were particularly concerned with preservation of detailed images and used techniques that distinguished between visual and verbal coding. The retention interval is, therefore, a variable of interest since a decline in performance over a short, unfilled interval would suggest the use of visual imagery as a memory code.

GENERATING VISUAL TEMPLATES

There is a second possible means of forming a visual template and that is to generate it from a structural description. Instead of trying to maintain an accurate

visual image over the retention interval, Ss might try to maintain a structural description in the form of a verbal code or a highly schematic image emphasizing specific subpatterns. If a part failed to match the structural description, Ss would then have the option of generating a visual template to scan for the part. Their performance, in this case, would depend upon their ability to form an accurate image of the entire pattern.

A prerequisite for being able to generate a visual template is a structural description that completely describes the pattern. If Ss had perceived Pattern 1 as two triangles but did not remember how the triangles overlapped, they would not be able to generate a visual template from their incomplete description. The formation of incomplete descriptions is illustrated in an experiment by Rock, Halper, and Clayton (1972). The stimuli used in this study were unfamiliar figures composed of either curved or jagged lines forming an outer contour and inner configuration. Recognition accuracy for the more complex figures was very low even when tested immediately. The Es concluded that the results indicated a failure of perception since all aspects of the figure were not adequately described. I will argue, in contrast, that the patterns shown in Fig. 2 are familiar enough that a complete description (whether structural or template) is initially formed. We will return to this issue later.

A final issue regarding the distinction between structural descriptions and template descriptions concerns their relation to the modality of the memory code. A structural description is based upon specific subpatterns, and the subpatterns composing the patterns in Fig. 2 have names—"triangle," "diamond," "hour-glass," "rectangle," etc. A structural description, even if represented in part by a schematic image, has access to the verbal code. A structural description cannot be based on pure imagery when language is available, since whatever parts are perceived by the S are associated with verbal labels. In contrast, a S cannot store a verbal label for a part that he did not see. If he encoded Pattern 1 as two triangles and did not notice that it contained a parallelogram, he could not rely on his verbal code to determine that Pattern 1 contained a parallelogram. He would then have to scan a visual template of the pattern to determine if a parallelogram was a possible part. Decisions regarding parts that were perceived by Ss can be based on a verbal code, but decisions regarding parts that were not perceived by Ss depend more upon their ability to form an accurate visual image of the complete pattern.

In order to determine what kind of descriptions are formed, Ss in Experiment I saw a sequence of patterns. Each trial consisted of the presentation of a pattern for 1 sec, an unfilled retention interval, and the presentation of a second pattern. The task was to decide whether the second pattern was a part of the first pattern. Reaction times were recorded in an attempt to distinguish between decisions based on previous perception of the

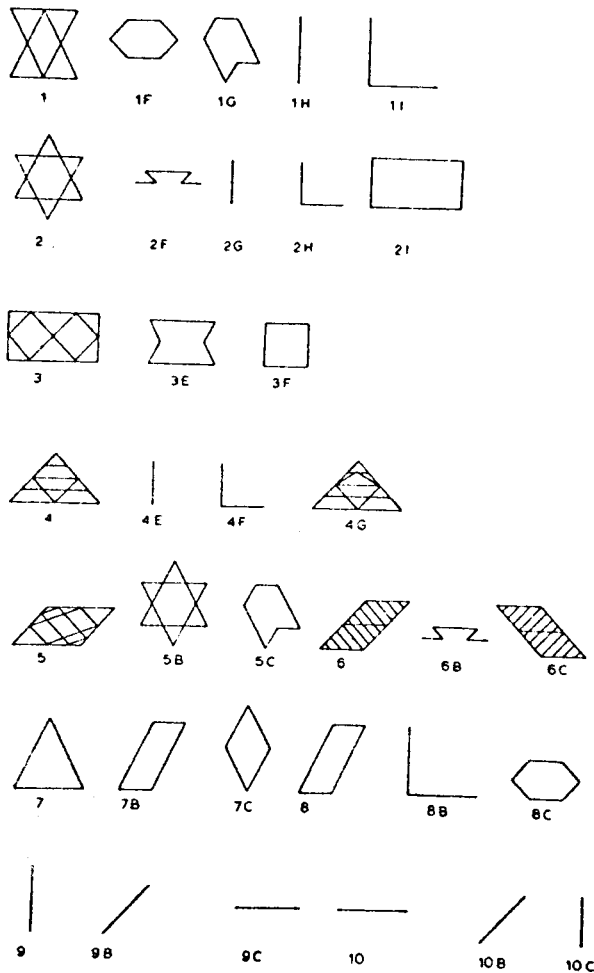


Fig. 3. Negative pairs of patterns used in Experiment I.

part and decisions based on scanning a visual image for the possible presence of the part.

EXPERIMENT I

Method

Subjects

The Ss were 80 undergraduates enrolled in an introductory psychology course at Case Western Reserve University. They received points that counted toward their grade in the course.

Procedure

Each trial consisted of the presentation of a pattern for 1 sec, a retention interval, and the presentation of a second pattern. The retention interval was 1.5 sec for half of the Ss and 5.5 sec for the other half of the Ss. Their task was to decide whether the second pattern was or was not a part of the first pattern, and they were told that a pattern identical to the first pattern should also be considered a part of the first pattern. They were further informed that a part of a pattern would always be the same size and occur in the same orientation, but its position on the screen might change, since all stimuli would be presented on the center of the screen. The instructions indicated that Ss should respond as quickly as possible but utilize enough time so as to avoid making errors.

A Kodak Carousel projector was used to rear-project the stimuli onto a translucent screen at a visual angle of 6 deg. Ss

responded positively by pressing a switch held in the right hand and negatively by pressing a switch held in the left hand. Reaction times were recorded to the nearest hundredth of a second by a Lafayette stop clock. A warning tone occurred 0.5 sec before the presentation of the second slide, and a Gerbrands electronic tachistoscopic shutter was used to present a rapid onset of the second slide, which was terminated by S's response. S indicated during the intertrial interval whether he was very certain, fairly certain, or uncertain regarding his decision. There were 7 sec between trials.

The experiment consisted of 10 practice trials followed by 50 experimental trials. The second pattern was a part of the first pattern on 50% of the trials. Figure 2 shows the 25 positive pairs and Fig. 3 shows the 25 negative pairs. Pattern 1, for example, was presented nine times during the experiments and was followed by an actual part on five trials (Patterns 1A-1E) and by a nonpart on four trials (Patterns 1F-1I). Patterns were presented in a random order under the constraint that at least two trials intervened between a repetition of the same first pattern and that runs of positive pairs or runs of negative pairs could not be greater than four in length. There were six different presentation orders used throughout the experiment.

Results

Table 1 shows for each positive pair (a) the number of correct and incorrect responses, (b) the mean reaction time, and (c) standard deviation for correct and incorrect responses. Table 2 shows the corresponding data for negative pairs. A few response times were lost as a result of E or equipment error, so the total responses do not always sum to 80 for each pattern. In addition, response times over 5 sec were not included in the data analysis.

The results indicate that Ss are very accurate in

Table 1
Reaction Times for Positive Pairs

Pattern	Cor- rect	RT	SD	Incor- rect	RT	SD
1A	78	1.20	0.48	2	1.86	1.54
1B	63	1.48	0.85	14	2.41	1.15
1C	52	1.68	0.90	25	1.65	0.78
1D	66	1.40	0.74	13	1.63	0.75
1E	11	1.39	0.78	69	1.55	0.87
2A	30	1.61	0.72	47	1.82	0.88
2B	71	1.59	0.90	9	1.85	1.15
2C	10	1.76	0.71	65	1.48	0.63
2D	11	1.78	0.87	67	1.58	0.88
2E	24	2.17	1.07	54	1.46	0.67
3A	28	2.08	0.88	47	1.70	0.90
3B	63	1.62	0.64	15	1.46	0.66
3C	60	1.62	0.65	19	1.66	0.67
3D	63	1.47	0.74	16	1.89	1.19
4A	14	2.04	1.13	65	1.53	0.74
4B	31	2.01	0.99	48	1.61	0.82
4C	73	1.15	0.46	3	1.21	0.19
4D	61	1.42	0.74	19	1.79	1.02
5A	16	1.66	0.75	62	1.48	0.69
6A	9	1.63	1.06	64	1.39	0.65
7A	77	0.95	0.39	2	1.49	0.40
8A	76	0.96	0.37	3	2.57	2.98
9A	77	0.85	0.29	2	2.08	2.37
10A	77	0.87	0.26	2	0.43	0.04
11A	77	0.95	0.57	0	-	-

recognizing a repetition of a pattern. For example, 78 of 80 Ss responded correctly to 1A and 73 of 76 Ss responded correctly to 4C. Although these results suggest that Ss have stored a complete description of the pattern, they do not enable us to conclude whether the pattern is remembered as a structural description or as a template description. The differential results for identifying the various parts of the pattern suggest that Ss were storing the patterns as structural descriptions. Although Ss were fairly accurate at identifying 1B, 1C, and 1D as parts of Pattern 1, only 11 of the 80 Ss correctly identified 1E (a parallelogram) as a part of Pattern 1. Ss were very accurate at identifying 2B (a triangle) as a part of Pattern 2 but generally did very poorly on the other parts of Pattern 2. Parts of other patterns, such as 3A and 4B, were also often responded to incorrectly.

The length of the retention interval did not have a significant effect on the percentage of correct responses. For the 1-sec interval, 67.3% of the responses were correct, compared to 68.6% for the 5-sec interval ($z = 0.87$, $p > .38$). Surprisingly, the majority of Ss thought they would be more accurate after the longer retention interval. Twenty of the Ss tested on the 1-sec delay thought their performance would improve with a 5-sec delay, eight thought their performance would remain the same, and 10 thought their performance would be worse. Twenty-three of the Ss tested on the 5-sec delay thought their performance would be worse with a 1-sec delay, three thought their performance would stay the same, and 13 thought their performance would improve.

Ss may have preferred a longer retention interval in an

Table 2
Reaction Times for Negative Pairs

Pattern	Cor- rect	RT	SD	Incor- rect	RT	SD
1F	68	1.52	0.67	9	1.61	0.85
1G	72	1.58	0.81	6	1.96	0.48
1H	50	1.84	0.81	29	1.67	0.92
1I	52	1.60	0.75	23	1.44	0.72
2F	67	1.66	0.80	8	2.12	0.90
2G	56	1.72	0.83	23	1.51	0.83
2H	72	1.33	0.44	4	1.70	0.89
2I	61	1.59	0.75	18	1.77	0.75
3E	48	1.73	0.85	29	1.90	0.79
3F	52	1.77	0.84	27	1.83	0.85
4E	54	1.72	0.76	25	1.54	0.73
4F	66	1.61	0.79	13	1.42	0.58
4G	49	1.99	0.75	28	1.51	0.75
5B	70	1.41	0.56	8	1.44	0.29
5C	75	1.58	0.84	4	1.62	0.87
6B	71	1.60	0.80	7	2.36	1.27
6C	45	1.52	0.64	32	1.63	0.71
7B	70	1.28	0.62	8	1.18	0.65
7C	63	1.38	0.51	16	1.55	0.77
8B	67	1.44	0.61	12	1.46	0.44
8C	73	1.15	0.42	6	1.16	0.56
9B	56	1.10	0.46	22	1.45	0.77
9C	59	1.15	0.51	20	1.24	0.72
10B	52	1.14	0.63	24	0.98	0.38
10C	51	1.11	0.54	26	1.32	0.56

Table 3
Percent Correct Responses for Patterns Used in
Both Experiment I and Experiment II

Positive Pairs			Negative Pairs		
Pattern	Experiment		Pattern	Experiment	
	I	II		I	II
1B	84	76	1F	85	84
1C	56	60	1H	56	70
1E	17	25	1I	67	76
2A	39	54	2F	86	70
2C	17	25	2G	70	76
2D	13	30	3E	60	34
3A	34	70	3F	58	64
3C	76	76	4G	58	56
4A	30	25	5C	94	76
4B	35	50	6B	100	84
5A	23	17	6C	48	82

attempt to use a visual image to identify additional subpatterns. There are two aspects of the data that question how successful they were at doing this. First, their performance as a whole was limited, since they failed to recognize many of the parts. Second, their performance after a 5-sec retention interval was not significantly better than their performance after a 1-sec retention interval. One possible reason for their limited accuracy is that they may have attempted to respond too rapidly without attempting to scan a visual image. The second experiment was conducted as an attempt to maximize performance by presenting the second pattern after a 1-sec retention interval and leaving the second pattern on the screen for 5 sec before requiring a response.

EXPERIMENT II

Method

Subjects

The Ss were 40 undergraduates enrolled in an introductory psychology course at Case Western Reserve University. They received points that counted toward their grade in the course.

Procedure

Each trial consisted of the presentation of a pattern for 1 sec, a 1-sec retention interval, and the presentation of a second pattern. The second pattern was terminated by a S's response, but Ss were instructed to respond only after hearing a tone. They were told that they would have 5 sec to make their decision in order to give them sufficient time to answer correctly.

The experiment consisted of 10 practice trials followed by 22 experimental trials. There were fewer experimental trials in Experiment II because only the more difficult pairs were used in an attempt to discover whether performance would improve on these patterns. Table 3 lists the 11 positive pairs and 11 negative pairs used in the experiment.

Results

Table 3 shows the percent correct responses for each

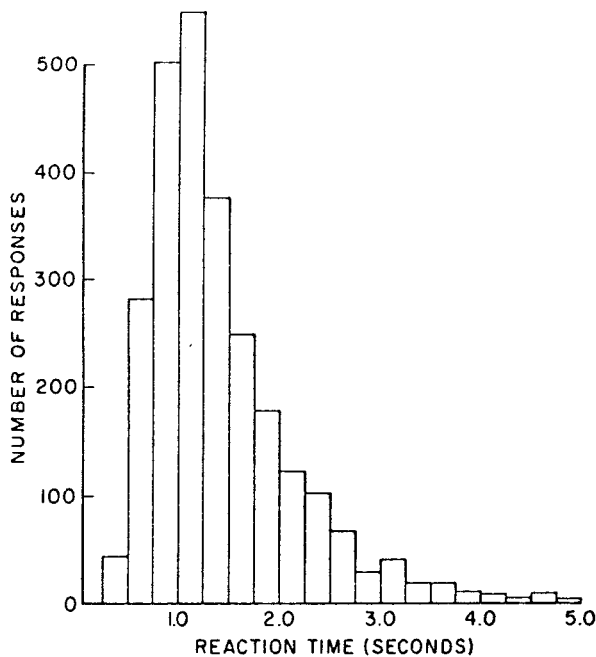


Fig. 4. Reaction time distribution for all correct responses in Experiment I.

pair in Experiment I and Experiment II. The Experiment I results are only from the group that had the 1.5-sec retention interval. The 1.5-sec retention interval included the 0.5-sec warning tone, which was not used in Experiment II since reaction times were not recorded. The principal difference between the two groups was that Ss in Experiment II had 5 sec to make a decision, whereas Ss in Experiment I were encouraged to respond as soon as they thought they were correct. Table 3 shows that in some cases the extra decision time in Experiment II resulted in a substantial improvement. However, the overall improvement was not significant. Ss in Experiment I were correct 55.2% of the time for the 22 pairs and Ss in Experiment II were correct 58.6% of the time for the 22 pairs ($z = 1.43, p > .15$).

A TENTATIVE MODEL

Although further research is necessary to identify the operations used in this task, I would like to offer the following tentative model as providing the best account of the data reported in the previous two experiments. The S encodes the first pattern as a complete structural description representing all the features and their interrelations. A description of Pattern 1 might be: "An equilateral triangle superimposed with an inverted equilateral triangle of equal size such that the apex of each triangle bisects the base of the other triangle and the sides of each triangle bisect the sides of the other triangle." The description is stored as a complex interaction of both visual and verbal memory codes (see Paivio, 1971, for a discussion of the involvement of both memory codes). In addition, if there is sufficient time, the S may also notice the presence of other subpatterns,

such as a diamond. When the second pattern is presented, the S encodes the pattern and uses a sequential memory scan to determine whether it matches any part of the stored description. If a match is obtained, the S responds positively. If a match is not obtained, the S has the option of responding negatively, attempting to scan a stored visual template of the first pattern, or attempting to generate a visual template from a structural description.

The fact that Ss had difficulty in correctly recognizing many parts of a pattern provides evidence that they store patterns as structural descriptions and recognize only what is a part of their description. This aspect of the data is consistent with the general theory of pattern recognition proposed by Sutherland (1968). It is likely that Ss were able to form a complete structural description, since they had no difficulty in recognizing a repetition of the first pattern. Pattern 4G was included in the experiment in order to determine whether Ss could correctly reject a pattern that was similar to the first pattern. Although 36% of the Ss responded positively to Pattern 4G, this percentage is substantially less than the 96% correct responses to Pattern 4C ($z = 8.22, p < .001$). Other evidence which suggests that incorrect responses were not the result of incomplete structural descriptions is that the majority of Ss were incorrect on four of the five parts for Pattern 2 even though Pattern 2 is a familiar pattern. Both Pattern 1 and Pattern 2 were presented nine times in the experiment which should also facilitate the formation of complete structural descriptions. The results, therefore, differ from the findings of Rock et al (1972), which indicated that the patterns in their experiment were so complex and unfamiliar that Ss could not form an accurate description.

I have taken the view that a pattern is stored as a structural description and that it is difficult to identify a part of a pattern that is not directly coded into the structural description. This is not to say, however, that it is totally impossible to correctly identify a part that was not previously perceived. Although we do not know for sure whether a correct response is based on prior perception or a visual image, it is likely that the faster decisions are based on prior perception and the slower decisions are based on inferences from visual images. If a part is presented that was previously perceived, the S should be fairly fast in responding, whereas if a part is identified through inference, the decision should take longer. Ideally, one would like to find a criterion on the reaction time distribution such that responses less than the criterion would be primarily due to previous perception and responses greater than the criterion would be primarily due to inference.

Figure 4 shows the reaction time distribution for correct responses and Fig. 5 shows the reaction time distribution for incorrect responses. Both distributions represent a pooling of responses from Experiment I over (1) positive and negative pairs and (2) 1- and 5-sec

retention intervals. The length of the retention interval had almost no effect on the distribution of reaction times, since the distributions are very similar when plotted separately for each of the two intervals. As can be seen by comparing Figs. 4 and 5, the distributions for correct and incorrect reaction times are also very similar. Both are positively skewed with a mode at about 1 sec.

The shape of the reaction time distribution does not by itself suggest where one should place a criterion to separate responses that are primarily based on previous perception from responses that are primarily based on inference. A theory of memory scanning developed by Sternberg (1967) may be a useful source for estimating such a criterion. The first assumption is that Ss in this experiment store in memory one complete structural description of a pattern and, usually, a number of separate parts. For example, a S may encode Pattern 1 as two triangles overlapping in a specified manner and he may also store the fact that Pattern 1 contains a diamond and oblique lines. When the second pattern is presented, S encodes the pattern and attempts to match it to previously perceived parts of the first pattern. Since attempted matches are made serially, decision time should be a function of the number of parts stored in memory.

A second determinant of decision time is the complexity of the patterns stored in memory. Cavanaugh (1972) reports a correlation greater than 0.99 between the time required to match a single item in memory and the number of items that can be stored in short-term memory (memory span). This observation suggests that memory matches are slower for more complex patterns. The present results are consistent with this observation. Table 1 shows that mean reaction times were 0.85 (9A) and 0.87 (10A) for comparing lines; 0.95 (7A), 0.96 (8A), and 0.95 (11A) for comparing simple geometric forms; and 1.20 (1A) and 1.15 (4C) for comparing complex patterns. It should be noted, however, that these reaction times include both encoding and matching of the second pattern, so part of this difference may be due to differences in encoding time.

The data in Table 1 can be combined with the data reported by Cavanaugh to form a rough estimate of the time required to encode the second pattern, scan the descriptions stored in memory, and respond. Since it takes 1.2 sec to respond to a repetition of Pattern 1 and approximately 50 msec/item to scan memory for additional geometric forms, it seems reasonable to conclude that decisions based on previously perceived parts should be completed by 1.5 sec and certainly by 2.0 sec. If the second pattern fails to match any part of the structural description, S has the option of responding negatively, scanning a remembered visual image for the presence of the part, or using a structural description to generate a visual image for scanning. It is likely that at least some of the correct response times over 2 sec involve a successful attempt to scan a visual

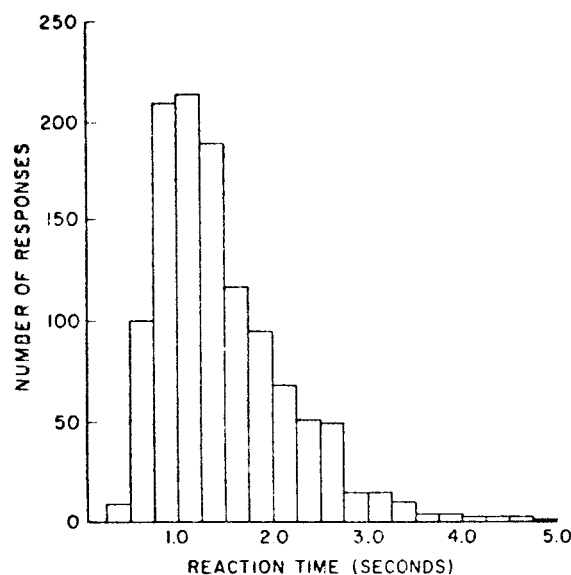


Fig. 5. Reaction time distribution for all incorrect responses in Experiment I.

image, and there may be differences in the quality of images across both Ss and patterns. But the fact that Ss' overall performance was mediocre and the fact that performance did not significantly improve in Experiment II suggests that the clarity of images is limited when verbal concepts are not available to assist performance.

THEORETICAL IMPLICATIONS

The current study is one of a number of recent studies that have sought to identify the kinds of operations that can be applied to images. Piaget and Inhelder (1969, p. 87) distinguish between two aspects of cognition: the figurative aspect, which appears as approximate copies of objects, and the operative aspect, which consists of modifying or transforming an object. The two aspects become complementary when we can imagine the transformations figuratively. An important question is to what extent do images preserve the detail of objects and what kind of operations can be applied to images.

One example is a task devised by Shepard to determine whether Ss could rotate a mental image in order to decide whether two patterns were identical except for differences in orientation. Cooper and Shepard (1972) summarize a number of sophisticated experiments which show that Ss are capable of performing such rotations. They suggest that images are highly abstract and schematic in comparison to a concrete perceptual image but preserve enough of the essential structure to permit accurate comparisons in their experiments.

The present results and interpretation differ from the previous article in placing more emphasis on the limitation of images. Since different patterns and operations are used, the view adopted here does not

necessarily conflict with the view expressed by Cooper and Shepard. The present study requires Ss to scan an image to retrieve new information—whether or not a previously unnoticed part is contained in the image. Kosslyn (1973) reports that Ss can scan an image to retrieve known information and that retrieval time is related to where Ss are instructed to focus on the image. The difference between scanning for old (previously perceived) information and scanning for new information corresponds to the distinction made by Piaget and Inhelder (1969) between reproductive and anticipatory images, the latter being the more complex.

The model proposed to account for these results is a tentative one and awaits further confirmation or disconfirmation and modification. A technique that may provide additional information is to give some of the more difficult patterns in a problem-solving task with an unlimited amount of time. Verbal protocols could be collected from Ss in an attempt to identify the specific operations used in the task. This approach was successfully applied by Baylor (1973) to Guilford's block visualization test. An interesting question would be whether identical operations are used in trying to find an embedded figure in a visual image as in a perceptual pattern that is physically present. Reed and Angaran (1972) used many of the patterns shown in Fig. 2 in a standard embedded-figures task in an attempt to discover the stimulus variables that make it difficult to find a figure. Although this study used children as Ss, the present paradigm could be used for adults by reversing the order of the two patterns in each pair. Ss would then have to scan the second pattern to determine whether it contained the first pattern.

REFERENCES

- Bahrick, H. P., & Boucher, B. Retention of visual and verbal codes of the same stimuli. *Journal of Experimental Psychology*, 1968, 78, 417-422.
- Baylor, G. W. Modelling the mind's eye. In A. Elithorn and D. Jones (Eds.), *Artificial and human thinking*. San Francisco: Josey-Bass, 1973.
- Cavanaugh, J. P. Relation between the immediate memory span and memory search rate. *Psychological Review*, 1972, 79, 525-530.
- Cooper, L. A., & Shepard, R. N. Chronometric studies of the rotation of mental images. Paper presented at the Symposium on Visual Information Processing, Carnegie-Mellon University, May 18-19, 1972.
- Gibson, E. Perceptual learning. *Annual Review of Psychology*, 1963, 14, 29-56.
- Glanzer, M., & Clark, W. H. The verbal-loop hypothesis: Conventional figures. *American Journal of Psychology*, 1964, 77, 621-626.
- Hochberg, J. Attention, organization, and consciousness. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts, 1970.
- Kosslyn, S. M. Scanning visual images: Some structural implications. *Perception & Psychophysics*, 1973, 14, 90-94.
- Narasimhan, R. On the description, generation, and recognition of classes of pictures. In A. Grasselli (Ed.), *Automatic interpretation and classification of images*. New York: Academic Press, 1969.
- Noton, D. A theory of visual pattern perception. *IEEE Transactions on Systems Science & Cybernetics*, 1970, 6, 349-357.
- Paivio, A. *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston, 1971.
- Phillips, W. A., & Baddeley, A. D. Reaction time and short-term visual memory. *Psychonomic Science*, 1971, 22, 73-74.
- Piaget, J., & Inhelder, B. Mental images. In P. Friaese and J. Piaget (Eds.), *Experimental psychology: VII. Intelligence*. London: Routledge & Kegan Paul, 1969.
- Posner, M. I., Bois, S. J., Eichelman, W. H., & Taylor, R. L. Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monograph*, 1969, 80, 1-13.
- Reed, S. K. *Psychological processes in pattern recognition*. New York: Academic Press, 1973.
- Reed, S. K., & Angaran, A. J. Structural models and embedded figure difficulty for normal and retarded children. *Perceptual & Motor Skills*, 1972, 35, 155-164.
- Rock, I., Halper, F., & Clayton, T. The perception and recognition of complex figures. *Cognitive Psychology*, 1972, 3, 655-673.
- Shepard, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning & Verbal Behavior*, 1967, 6, 156-163.
- Simon, H. A. Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 1972, 79, 369-382.
- Smith, E. E. Choice reaction time: An analysis of major theoretical positions. *Psychological Bulletin*, 1968, 69, 77-110.
- Smith, E. E., & Nielsen, G. D. Representations and retrieval processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology*, 1970, 85, 397-405.
- Standing, L., Conezio, J., & Haber, R. N. Perception and memory for pictures: Single-trial learning of 2560 visual stimuli. *Psychonomic Science*, 1970, 19, 73-74.
- Sternberg, S. Two operations in character recognition: Some evidence from reaction time measurements. *Perception & Psychophysics*, 1967, 2, 45-53.
- Sutherland, N. S. Outlines of a theory of visual pattern recognition in animals and man. *Proceedings of the Royal Society*, 1968, 171, 297-317.

(Received for publication June 11, 1973;
accepted August 27, 1973.)