Journal of Experimental Psychology: Human Perception and Performance
1975, Vol. 1, No. 3, 214–220

Visual Transformation of Size

Claus Bundesen and Axel Larsen
Copenhagen University, Denmark

To investigate human visual identification of different-sized objects as identically shaped, matching reaction times were measured for pairs of simultaneously presented random figures. In three experiments, reaction time for correct reactions to test pairs of figures of the same shape and orientation consistently increased approximately linearly as a function of the linear size ratio of the figures. In the second experiment, where this ratio was defined for control pairs as well as for test pairs, reaction time for correct reactions to control pairs showed a similar increase as a function of size ratio. The results suggest that the task was performed by a gradual process of mental size transformation of one of the members of each pair of figures to the format of the other one.

The way we visually identify form independently of size has long been considered a problem (Ehrenfels, 1890; Hebb, 1949; Köhler, 1929; Lashley, 1942; Pitts & McCulloch, 1947). An aspect of this problem concerns our general ability to identify objects of different sizes as identically shaped.

One type of explanation (Type 1) of this ability invokes the fact that we seem capable of gradually transforming a mental image of an object of a given size to an image of an object of the same shape, but of any other size. An example of an explanation of this type runs as follows: Given a visual impression of an object of a specific shape and size and a visual impression of another object of the same shape and a specific size, we can identify the two objects as identically shaped by encoding one of the impressions as a mental image; by transforming the image so that the represented size changes to that of the other visual impression; and then matching this impression with the transformed mental image. It is not implied that the mental image is a detailed representation (i.e., a holistic template); in most cases a quite schematic image (i.e., a partial representation) might suffice. Other explanations of this type, based on mental processes of size transformation, are also possible (e.g., Posner, 1969).

According to another type of explanation (Type 2), visual identification of form is based on direct extraction of size-invariant features from the sensory input—without processes of size transformation (e.g., Dodwell, 1970; Gibson, 1969). Visual identification of identically shaped objects as such should take place, then, by comparison of different sets of extracted size-invariant features.

Type 1 explanations have specific implications concerning processing time. A mental size transformation is a gradual process so that, for instance, a mental dilation of a given object takes more time, the greater the corresponding factor of geometric multiplication. The particular example of a Type 1 explanation given above implies, accordingly, that the time taken to perceive two objects as having the same shape increases monotonically with their linear size ratio, for constant distance and orientation, provided that one of the objects is kept constant. On the other hand, a simple variation in processing time for varying size ratios is not to be expected with a Type 2 explanation.

The purpose of Experiments 1 and 2 was to test these two types of explanations by measurement of reaction time.

Requests for reprints should be sent to Claus Bundesen, Psychological Laboratory, Copenhagen University, Njalsgade 90, DK-2300 Copenhagen S., Denmark.

Experiment 1

Method

Subjects. Seven subjects participated, including the authors. Four of the subjects were in their twenties, two in their thirties, and one in her fifties. All had normal or corrected vision.

Stimuli. A total of 224 slides of black line drawings on white backgrounds were used. Half of the 224 slides were test slides, each of which showed two figures of the same shape; the rest were control slides showing pairs of figures of different shapes (see Figure 1). The test slides were photographed from drawings made by a Calcomp plotter using the following three-step procedure:

1. Twelve numbers, $x_1, x_2, \ldots, x_6$, and $y_1, y_2, \ldots, y_6$, were randomly drawn from a uniform distribution of rational numbers between -2 and 2. The numbers specified a figure that could be generated in a coordinate system by connecting (x_1, y_1) with (x_2, y_2), (x_2, y_2) with $(x_3, y_3), \ldots,$ and (x_5, y_5) with (x_6, y_6), all connections being made by straight lines.

2. The "center of gravity" (x_0, y_0) of the specified figure was computed by

$$x_0 = \sum_{i=1}^{5} u_i d_i / \sum_{i=1}^{5} d_i$$

and (1)

$$y_0 = \sum_{i=1}^{5} v_i d_i / \sum_{i=1}^{5} d_i,$$

where (u_i, v_i) was the midpoint of the line segment between (x_i, y_i) and (x_{i+1}, y_{i+1}), and d_i the length of this segment. The figure was plotted following a parallel displacement so that its center of gravity was positioned at (0, 0).

3. Geometric multiplication by a factor M (either 1, $\frac{3}{4}$, $\frac{3}{5}$, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, or $\frac{1}{5}$) specified a new figure. This figure was plotted following a parallel displacement so that its center of gravity was positioned at (6.4, 0).

The control slides were photographed from drawings made by a procedure consisting of subprocedures 1 and 2, followed by subprocedures 1 and 3. The seven different values of M were used with equal frequency for test drawings as well as for control drawings. From each of the 112 drawings, two identical slides were photographed. In each case, the duplicate was rotated π rad in the plane of the picture.

Procedure. Each subject was tested individually in two experimental sessions. In each experimental session the 224 slides were presented in random order. The first experimental session was preceded by a practice session lasting approximately 30 min, in which similar stimulus material was employed to familiarize the subject with the apparatus and procedure.

The subject was seated about 3 m in front of a screen on which the projections of the slides spanned approximately .3 rad horizontally and .2 rad vertically. For each pair of figures presented, the centers of gravity were positioned symmetrically around the midpoint of the screen with the same vertical height and a horizontal separation of about .1 rad. During projection of a slide, the pupils of the subject received an illuminance of approximately 10 lx from the stimulus field and 3 lx from the surrounding field.

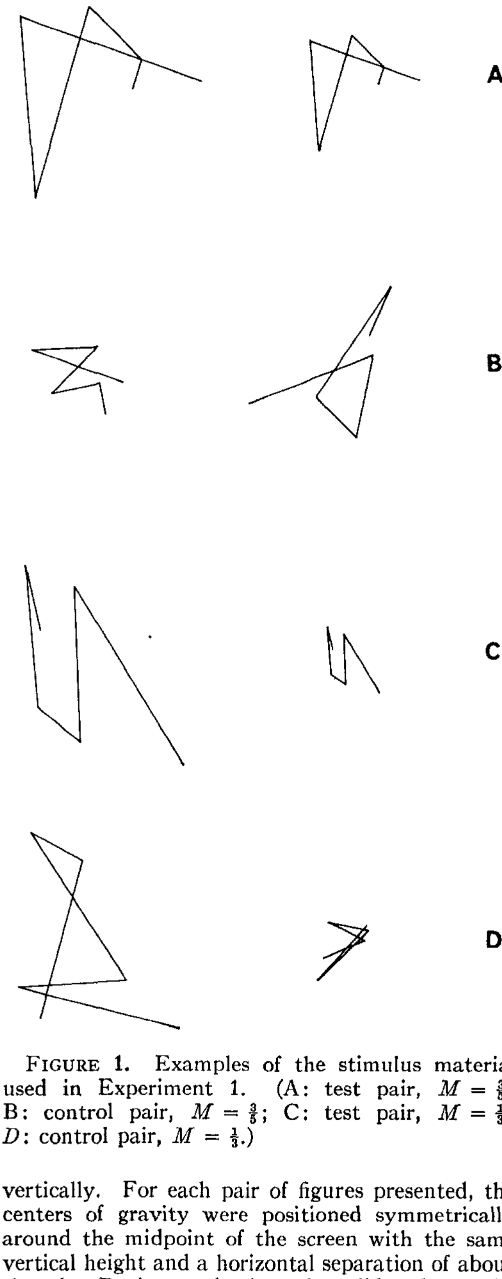

Figure 1. Examples of the stimulus material used in Experiment 1. (A: test pair, $M = \frac{3}{5}$; B: control pair, $M = \frac{3}{5}$; C: test pair, $M = \frac{1}{3}$; *D*: control pair, $M = \frac{1}{3}$.)

The subject was told that each slide showed two distinct figures, and that his task was to decide "as quickly as possible" whether the two figures were identical except for a parallel displacement and a change of size. If the figures were identical, he pressed a button on his right; if they were not, he pressed a button on his left. The exposure stopped immediately when one of these buttons was pressed. A new slide was projected with a latency of 2 sec when the subject signaled "ready" by pressing a third button. The experiment was run by a laboratory computer with a crystal clock, which measured reaction time to the nearest 10 msec from the onset of the stimulus field.

Results

The size ratio of a test pair was defined as the inverse of that value M ($M \leqslant 1$) by which it was constructed. For control pairs, no values of the size ratio were defined. Only correct reactions to test pairs had immediate interest. The subjects' error rates were between 2% and 11% with means of 4% for test pairs and 5% for control pairs.

For correct reactions to test pairs, mean reaction time across subjects and sessions increased approximately linearly as a function of the size ratio. The function is shown in Figure 2 with a straight line fitted by the method of least squares.

For each Subject × Session × Value of Size Ratio, the median reaction time for correct reactions to test pairs was computed. Across subjects and sessions, the pattern of means of medians was very similar to the one shown in Figure 2. To avoid parametric assumptions, the further analysis was based on the medians.

For each Subject × Session a least squares line was computed for median reaction time as a function of size ratio. In each case the slope was positive. Goodness of fit was evaluated by test of the hypothesis that for each Subject × Session × Value of Size Ratio, the probability that reaction time fell above the fitted line was .5. The fit proved to be satisfactory, $\chi^2(70) = 70.6$, $p = .46$.

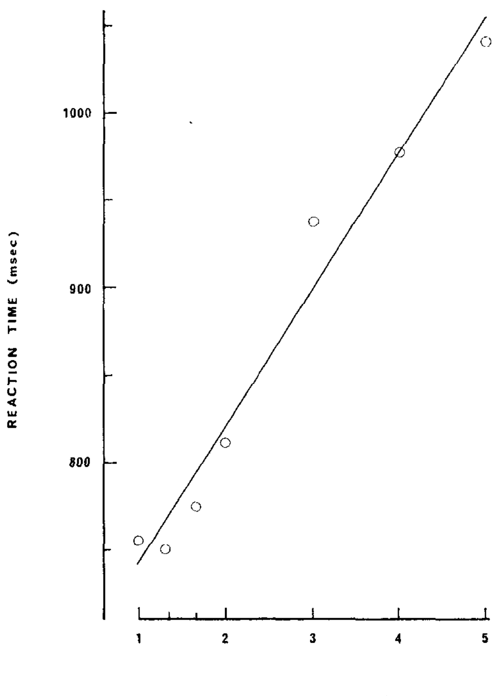

FIGURE 2. Mean reaction time for test pairs as a function of linear size ratio in Experiment 1.

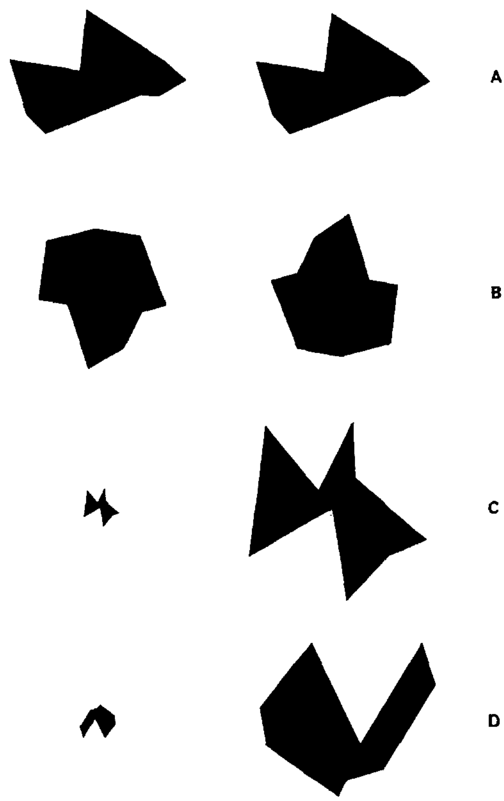

FIGURE 3. Examples of the stimulus material used in Experiment 2. (A: test pair, $M = 1$; B: control pair, $M = 1$; C: test pair, $M = \frac{1}{5}$; D: control pair, $M = \frac{1}{5}$.)

EXPERIMENT 2

In order to test the generality of the pattern of reaction times obtained in Experiment 1, a second experiment was performed where the stimulus material consisted of solid shapes instead of line figures, and where the size ratio was defined for control pairs as well as for test pairs.

Method

Stimuli. A total of 200 slides of black solid shapes on white backgrounds were used. The slides were photographed from closed outline drawings filled in with india ink (see Figure 3).

The outline drawings were made by a Calcomp plotter using subprocedures 1, 2, 3a, and 4 of the program below for the test drawings and subprocedures 1, 2, 3b, and 4 for the control drawings.

1. Eighteen numbers, $x_1, x_2, \ldots, x_9$, and $y_1, y_2, \ldots, y_9$, were randomly drawn from a uniform distribution of rational numbers between -2 and 2, so that $x_1 < x_2 < \ldots < x_9$, and $y_1 < y_2$, $y_2 > y_3$, $y_3 < y_4$, $y_4 > y_5, \ldots, y_7 < y_8$, $y_8 > y_9$. The numbers specified a closed figure that could be generated in a coordinate system by connecting (x_1, y_1) with (x_2, y_2); (x_1, y_1) with (x_3, y_3), and (x_2, y_2) with (x_4, y_4); (x_3, y_3) with (x_5, y_5), and (x_4, y_4) with (x_6, y_6); (x_5, y_5) with (x_7, y_7), and (x_6, y_6) with (x_8, y_8); and finally (x_7, y_7) and (x_8, y_8) with (x_9, y_9); all connections were made by straight lines.

2. The center of gravity (x_0, y_0), defined as the first moment of area, was computed for the specified figure. The figure was plotted following a parallel displacement so that its center of gravity was positioned at (0, 0).

3a. Geometric multiplication by a factor M (either 1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, or $\frac{1}{5}$) specified a new figure.

3b. Geometric multiplication by M (1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, or $\frac{1}{5}$) followed by a plane rotation of π rad specified a new figure.

4. The new figure was plotted following a parallel displacement so that its center of gravity was positioned at (6.4, 0).

The different values of M were used with equal frequency for the 50 test drawings as well as for the 50 control drawings. From each drawing, two identical slides were photographed; the duplicate was rotated π rad in the plane of the picture.

Subjects and procedure. The subjects and the procedure were the same as in the previous experiment.

Results

The size ratio of a stimulus pair was defined as the inverse of that value M ($M \leqslant 1$) by which it was constructed. Only correct reactions were analyzed. Error rates were between 1% and 12% with means of 5% for test pairs and 7% for control pairs.

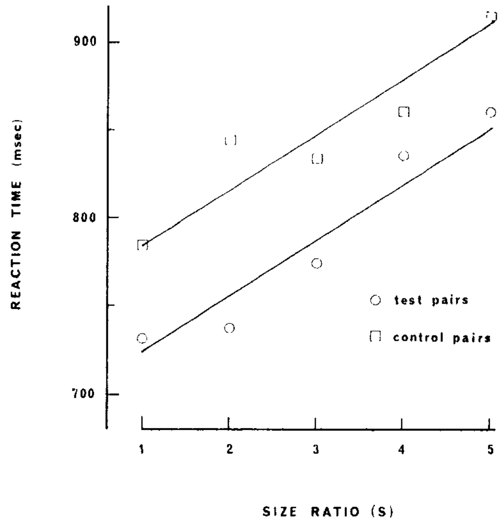

FIGURE 4. Mean reaction time as a function of linear size ratio for test pairs and control pairs in Experiment 2.

For correct reactions to control pairs as well as for correct reactions to test pairs, mean reaction time across subjects and sessions increased as a function of size ratio, the relations being approximately linear and parallel. The functions are shown in Figure 4, fitted by a least squares pair of parallel lines.

For each Subject × Session × Value of Size Ratio, median reaction times for correct reactions to test pairs and to control pairs were computed. Across subjects and sessions, the pattern of means of medians was very similar to the one shown in Figure 4.

For each Subject × Session two separate least squares lines were fitted to median reaction time as a function of size ratio: one for correct reactions to test pairs, one for correct reactions to control pairs. Goodness of fit was evaluated by test of the hypothesis that for each Subject × Session × Value of Size Ratio and for each type of stimulus pair (test vs. control), the probability that reaction time fell above the fitted line was .5. The fit was satisfactory; for test pairs, $\chi^2(42) = 46.5$, $p = .30$; for control pairs, $\chi^2(42) = 41.0$, $p = .52$; in total, $\chi^2(84) = 87.5$, $p = .39$. For each Subject × Session, then, a least squares pair of parallel lines was fitted to the medians. The fit was acceptable, $\chi^2(98) = 112.8$, $p = .14$.

Discussion of Experiments 1 and 2

The pattern of reaction times obtained in Experiments 1 and 2 is not easily understood by Type 2 explanations of direct extraction of size-invariant features. If the results for test pairs were considered separately, some sort of test for general similarity within the pairs of figures presented might possibly be invoked to explain the increase in reaction time with size ratio; the explanation would be based on the fact that when similarity of size decreased, so did the general similarity. With that interpretation, however, reaction time for control pairs should decrease with increasing dissimilarity (increasing size ratio), which runs counter to the results of Experiment 2.

On the other hand, the direct relationship between reaction time and size ratio follows readily from Type 1 explanations. For correct reactions to test pairs (Experiments 1 and 2), the direct relations between reaction time and size ratio can be accounted for by an assumption that the reactions were based on mental size transformation followed by successful match. For correct reactions to control pairs (Experiment 2), the direct relation between reaction time and size ratio can be explained on a related assumption that these reactions were based on mental size transformation followed by mismatch. Furthermore, these assumptions can explain the fact that the relations between reaction time and size ratio could be approximated by parallel curves for the two types of stimulus pairs (Experiment 2). Thus, the pattern of reaction times obtained in the two experiments supports the supposition that the task was performed by a process of mental size transformation.

There are, however, many possible Type 1 explanations for the results. Two models are especially tempting. One of them (Model T), previously outlined, implies that the task was accomplished by mental size transformation of one of the members of each pair of figures to the format of the other one. The other model (Model N) implies that the task was accomplished by "normalization" (cf. Minsky, 1961) of all figures to a common standard format. Model T explains immediately the direct relations between reaction time and size ratio, as does Model N, if the standard format is assumed to have been larger than the largest experimental figure.

According to Model T, reaction time was determined by the size relation between the presented figures; whereas according to Model N, reaction time was determined by the absolute size of the figures. That the two experiments did not decide between these models was due to the fact that the size ratio between simultaneously presented figures was closely correlated to the size of the smaller figure, the size of the larger figure being independent of the size ratio. For the same reason it was impossible to evaluate different versions of Model T on the basis of whether reaction time depends on size ratios or size differences between figures.

Experiment 3

In order to restrict the set of possible Type 1 explanations, a third experiment was performed where reaction time's dependence on relative size was contrasted with its dependence on absolute size.

Method

Subjects. Seven subjects participated, none of whom had served in the previous experiments. Because one of the seven subjects served in only half of the experiment, his data were discarded. All subjects were between 25 and 35 years old and had normal or corrected vision.

Stimuli and procedure. The stimulus material consisted of the 224 slides from Experiment 1 supplemented by 160 new slides in a random sequence. The new slides were constructed by the same method as the old ones, except that rational numbers between −1 and 1 were drawn in subprocedure 1 of Experiment 1 and M took on the values $\frac{4}{3}$, $\frac{5}{3}$, 3, 4, or 5 in subprocedure 3. (For $M = 2$ the modified procedure was equivalent to the earlier procedure for $M = \frac{1}{2}$.) In all other ways, stimuli, apparatus, and procedure were the same as in the previous experiments.

Results

The size ratio of a test pair was defined as the value of M, if $M > 1$, and as the value of $1/M$, if $M \leqslant 1$. Two overlapping series of test pairs were defined: Series A contained those pairs that were constructed by $M \leqslant 1$, whereas Series B contained those constructed by either $M > 1$ or $M = \frac{1}{2}$. For control pairs, no values of size ratio were defined. Only correct reactions to test pairs had immediate interest. Error rates were between 1% and 30% with means of 7% for test pairs and 9% for control pairs.

For each series of test pairs, mean reaction time for correct reactions across subjects and sessions increased approximately linearly as a function of size ratio. The functions are shown in Figure 5, fitted by a least squares line.

Median reaction times were computed for each Subject × Session × Series × Value of Size Ratio. The pattern of means of medians across subjects and sessions was very similar to the one shown in Figure 5.

The results from Series A replicated the results from Experiment 1. For each Subject × Session, median reaction time as a function of size ratio was well approximated by a straight line, $\chi^2(60) = 61.1$, $p = .44$. The results from Series B were similar to the results from Series A, $\chi^2(48) = 46.6$, $p = .54$. Moreover, for each value of size ratio the results from the two series corresponded so that for each Subject × Session they could be approximated satisfactorily by a common straight line, $\chi^2(120) = 100.8$, $p = .90$.

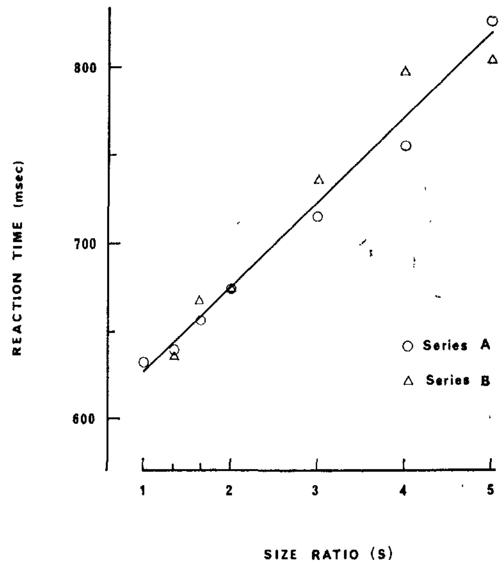

FIGURE 5. Mean reaction time as a function of linear size ratio for Series A and Series B in Experiment 3.

Discussion

The results agree with the expectations of Model T: Reaction time should increase with size ratio for both series. Model N can explain an increasing relation for Series A, if a sufficiently large standard format is assumed; and for Series B, if a sufficiently small standard format is assumed. However, Model N cannot explain increasing relations for both series at the same time. In the model's explanation for Series A, the reaction times for pairs with, for example, size ratio equal to 2 must be assumed to express the durations of normalization of the smaller members of pairs of figures, and not of the larger ones as demanded by the model's explanation for Series B.

The results indicate, in fact, that the processing time increased linearly as a function of the size ratio of the presented figures, so that absolute sizes and size differences were unimportant per se. Consequently, the results support a version of Model T in which the time taken to perceive two objects as having the same shape increases linearly with the objects' size ratio when distance and orientation are kept constant.

GENERAL DISCUSSION

It is obviously possible to identify objects of different sizes as identically shaped without the use of mental processes of size transformation. In principle, the experimental task could be accomplished without errors by an intellectual process of comparing verbal descriptions of the presented figures; or it could be accomplished with a few errors by tests for parallel lines disregarding size and figural organization. It is difficult to explain the obtained pattern of reaction times by strategies that do not involve mental processes of size transformation. But it is possible that the subjects sometimes used such strategies, which may have decreased the rate of increase of reaction time as a function of size ratio. On the other hand, it is also possible that the subjects sometimes repeated a test, performing several size transformations for the same pair of figures, which would add to the rate of increase of reaction time. In consequence, it is difficult to ascertain how the obtained slopes reflect the absolute rate of mental transformation of size.

It should be noted that presentations in which the left-hand figure was smaller than the right-hand figure alternated randomly with presentations in which the reverse was the case. Accordingly, the explanation by mental size transformation implies that the process of transformation was determined by prior computation of, at least, required direction of transformation. It is reassuring that this kind of preprocessing is comparatively simple.

The present findings on size transformation are related to the results of Cooper and Shepard (1973) and Shepard and Metzler (1971) on visual rotation. However, the present results are incompatible with a simple generalization of the interpretation by Shepard and his associates, which may point to important functional differences between mental size transformation and mental rotation.

In Shepard's interpretation, the degree of mental rotation in a matching task may well be determined by prior computation of the degree required, that is, of the orientational difference involved. This corresponds to the results on size transformation. However, predetermination of orientation of figures is assumed to depend essentially on predetermination of their identities. On the contrary, the present experiments suggest that predetermination of a required size transformation does not imply determination of identity. As members of the same control pair were generally rather dissimilar to each other (in Experiments 1 and 3), prior determination of identity would actually have eliminated the need for a subsequent mental size transformation.

REFERENCES

Cooper, L. A., & Shepard, R. N. Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press, 1973.

Dodwell, P. C. *Visual pattern recognition.* New York: Holt, 1970.

Ehrenfels, C. von. Über Gestaltqualitäten. *Vierteljahrsschrift für wissenschaftliche Philosophie,* 1890, *14,* 249–292.

Gibson, E. J. *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts, 1969.

Hebb, D. O. *Organization of behavior.* New York: Wiley, 1949.

Köhler, W. *Gestalt psychology.* New York: Liveright, 1929.

Lashley, K. S. The problem of cerebral organization in vision. In H. Klüver (Ed.), *Biological symposia* (Vol. 7). *Visual mechanisms.* Lancaster, Pa.: Jaques Cattell Press, 1942.

Minsky, M. Steps toward artificial intelligence. *Proceedings of the Institute of Radio Engineers,* 1961, *49,* 8–30.

Pitts, W., & McCulloch, W. S. How we know universals: The perception of auditory and visual forms. *Bulletin of Mathematical Biophysics,* 1947, *9,* 127–147.

Posner, M. I. Abstraction and the process of recognition. In J. P. Spence & G. H. Bower (Eds.), *The psychology of learning and motivation* (Vol. 3). New York: Academic Press, 1969.

Shepard, R. N., & Metzler, J. Mental rotation of three-dimensional objects. *Science,* 1971, *171,* 701–703.

(Received February 26, 1975)