PSYCHOLOGICAL BULLETIN
Vol. 53, No. 6, 1956

THE QUANTITATIVE STUDY OF SHAPE AND PATTERN PERCEPTION[1]

FRED ATTNEAVE AND MALCOLM D. ARNOULT
Skill Components Research Laboratory, Air Force Personnel and Training Research Center

The pre-eminent importance of formal or relational factors in perception has been abundantly demonstrated during some forty years of gestalt psychology. It seems extraordinary, therefore, that so little progress has been made (and, indeed, that so little effort has been expended) toward the systematizing and quantifying of such factors. Our most precise knowledge of perception is in those areas which have yielded to psychophysical analysis (e.g., the perception of size, color, and pitch), but there is virtually no psychophysics of shape or pattern.

Several difficulties may be pointed out at once: (*a*) Shape is a multidimensional variable, though it is often carelessly referred to as a "dimension," along with brightness, hue, area, and the like. (*b*) The number of dimensions necessary to describe a shape is not fixed or constant, but increases with the complexity of the shape. (*c*) Even if we know *how many* dimensions are necessary in a given case, the choice of particular descriptive terms (i.e., of reference-axes in the multidimensional space with which we are dealing) remains a problem; presumably some such terms have more psychological meaningfulness than others.

The need for an adequate psychophysical framework is most obvious in those studies (having to do with discrimination, for example, or with positive or negative transfer) in which it is necessary to manipulate shape or pattern as an independent variable. Unless some meaningful units of variation are specifiable, functional relationships cannot be obtained. It is somewhat less obvious, but nonetheless true, that a comparable need exists in experiments which seek to determine how form perception is influenced by extrinsic variables such as size, contrast, method and degree of familiarization, etc. In studies of this sort, the experimenter commonly uses some small, arbitrarily chosen set of stimuli: sometimes simple geometrical forms; sometimes a group of "nonsense" shapes which he draws in a more or less haphazard manner. If the results obtained are "significant" in the usual sense, we have some specifiable degree of confidence that they are generalizable to people other than those used as subjects, but the degree to which they are generalizable to new stimuli remains a matter of conjecture. Yet the latter kind of generalization is no less important than the former. Only in rare cases of applied research is the investigator really content with results which hold only for the particular stimulus objects employed experimentally.

Egon Brunswik (**9**, **10**, **11**) is perhaps the only psychologist who has ever given due weight to the importance of stimulus-sampling, or of

[1] This research was carried out at the Skill Components Research Laboratory, Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas, in support of Project 7706, Task 27001.

situation-sampling in general. Although the approach of this paper is somewhat different from Brunswik's, for reasons which are developed below, we wish to acknowledge freely Brunswik's influence upon our own thinking, and to commend his writings on this subject to any reader unacquainted with them. Brunswik takes the reasonable position that results with "ecological validity" may be obtained only by the use of experimental materials which are drawn from, and hence representative of, the real situations to which one wishes to generalize. Thus, in the study of shape perception, it would be desirable to experiment with the shapes of natural objects. Suppose, however, that we wish to investigate the learning and memory of shapes with which subjects are initially unfamiliar: the requirement of unfamiliarity will obviously preclude the experimental use of shapes which are commonly encountered. Is there any sensible procedure for choosing stimulus-materials in this sort of situation?

It is our belief, at this time, that the problem of generalizing from experimental stimuli may profitably be broken into two parts. First, there is the problem of specifying the *stimulus-domain*, i.e., the problem of drawing a sample of stimuli from a parent population characterized by certain determinate statistical parameters. The *stimulus-domain*, or parent population, includes all those stimuli to which the results may be generalized, and is defined by the statistical parameters which characterize it. In the following section we shall indicate a variety of particular methods for drawing "random" patterns and shapes from such clearly defined hypothetical populations, to which experimental results may then be generalized with measurable confidence.

The second problem, which is really a special case of the first, is that of drawing a sample which has "ecological validity." If our real aim is to generalize to natural forms, or to some subset thereof, it is necessary to estimate the psychologically important statistical parameters of these natural forms in order that experimental materials may be constructed to possess the same parameters. Thus, we are brought back to the acute need for a general psychophysics of form. In the final section we shall discuss the kinds of physical analysis and measurement which appear appropriate to such a psychophysics.

The Construction of Stimuli

All the methods described below for constructing nonsense shapes and patterns have in common the fact that the particular characteristics of each figure are randomly determined. Each method is, in effect, a set of rules by which points are plotted and connected in accordance with values obtained from a table of random numbers. Each method, or set of rules, thus determines a domain of stimuli. The stimuli actually constructed for use in a given experiment will, if they are all constructed according to the same rules, be a random sample of the *stimulus-domain* defined by the set of rules. The experimental results, consequently, may be generalized both to the entire stimulus-domain and to the appropriate subject population.[2]

[2] The kind of double-generalization proposed here would require an error term which included the variance due to subjects, the variance due to stimuli, and the interaction between them. In what is perhaps the most obvious analysis-of-variance design, the *subjects*×*stimuli*×*treatments* mean square would be the appropriate error term to use.

The experimenter who desires to use stimuli constructed in this manner must determine what set of rules will provide him with a stimulus population having the characteristics he wants. If one desires to generalize experimental results to the world of real objects (chairs, airplanes, people, etc.), it is necessary to have a stimulus sample possessing *ecological validity.* To construct nonsense stimuli of this sort one must know the pertinent parameters of the stimulus-domain of real objects and use these parameters in constructing the experimental stimuli. In the next section we shall discuss some of the problems inherent in this methodological requirement and some of the attempts which have been made to solve them.

In the present section, some general methods for constructing stimuli are described in sufficient detail that the reader, if he desires, may repeat the operations in order to develop additional stimuli belonging to the various stimulus-domains defined by the methods. It should be kept in mind, however, that these methods are described merely as examples and are not intended to constitute a comprehensive catalog of all possible methods. Descriptions will be given of methods for generating shapes having either closed or open contours, for generating various kinds of patterns, and for introducing systematic variations or transformations of shapes or patterns.

Closed Contours—Angular Shapes

Method 1. Starting with a sheet of graph paper—say 100×100—successive pairs of numbers between 1 and 100 are selected from a table of random numbers. Each pair will determine a point which can be plotted on the 100×100 matrix. The total number of such points to be plotted can be determined either randomly or arbitrarily.

When all the points have been plotted, a straightedge is used to connect the most peripheral points in such a way as to form a polygon having only convex angles. This operation will usually leave some unconnected points within the polygon (Fig. 1*a*). When a point falls within some small, arbitrarily chosen distance of the proper perimeter (e.g., the point between segments 7 and 8 in Fig. 1*a*) it is included even though it makes a slightly concave angle, since otherwise an indentation practically dividing the shape into two parts might later occur. The sides of the polygon are numbered, and the points remaining inside are assigned letters. The table of random numbers is then used to determine which of the

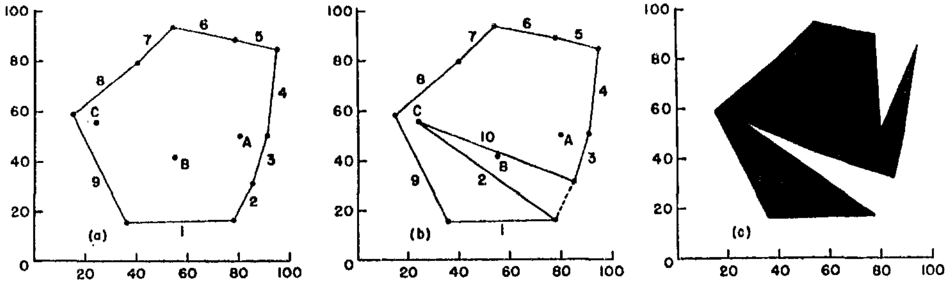

FIG. 1. SUCCESSIVE STAGES IN THE CONSTRUCTION OF A "RANDOM" FIGURE ACCORDING TO METHOD 1 (SEE TEXT)

central points is connected to which side. In the example given, Point *C* was connected to Side 2, forming in the process Side 10 (Fig. 1*b*). At this stage in the construction, the possibilities of connecting points have been changed. Point *A* may now be taken into Sides 3, 4, 5, 6, 7, 8, or 10, but not into Sides 1, 2, or 9. Point *B* may be connected only to Side 2 or Side 10. If Point *A* is connected to Side 5, forming new Side 11, there remains only the possibility of connecting Point *B* to Side 2 or Side 10 (see Fig. 1*b*). Connecting Point *B* to Side 10 completes the shape, which finally appears as shown in Fig. 1*c*.

It will be noted that every step in the procedure is determined either randomly or by the elimination of all other possibilities. Furthermore, every step is completely determinate and can be duplicated by anyone using the same rules and the same selections from the table of random numbers.

Method 2. This method of constructing random shapes is also started by plotting successive pairs of random numbers as coordinates on graph paper. As each point is plotted it is given a number so that eventually all are numbered serially. These points are then connected in the order in which their serial numbers first appear in a table of random numbers, except that numbers which violate certain rules of construction are rejected. The incomplete construction shown in Fig. 2*a* will provide examples of permitted and nonpermitted connections. The rules for connecting points are as follows:

a. No line may be drawn twice. Assume, in Fig. 2*a*, that the last line drawn was from Point 2 to Point 5. If the next number in the table were 2, it would be rejected since that connection has already been made.

b. No line may be drawn which completely encloses a point within the perimeter of the figure. From Point 5 it would not be permissible to draw a line to Point 6 or to Point 4, since either action would completely enclose Points 3 and 8.

c. No two points may be directly connected if they are already connected by a path which follows perimeter lines without passing through any other plotted points. For example, Point 5 may not be connected

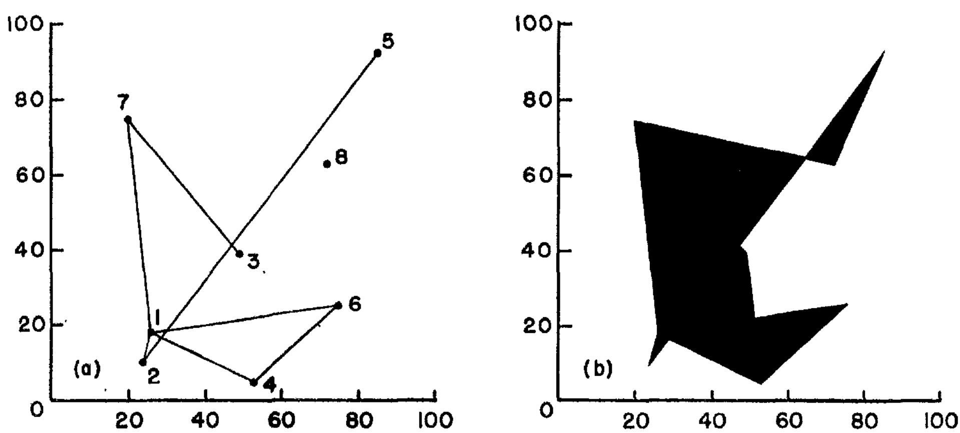

FIG. 2. EXAMPLE OF NONSENSE SHAPE CONSTRUCTED BY THE RULES OF METHOD 2: *a.* INCOMPLETE CONSTRUCTION DEMONSTRATING PERMISSIBLE AND NONPERMISSIBLE CONNECTIONS; *b.* THE COMPLETED SHAPE

to Points 3 or 7, Point 3 may not be connected to Points 5 or 6, and Point 2 may not be connected to Point 4.

d. The figure is complete when each point has been connected to at least two other points. It sometimes happens that the table of random numbers leads one to a point which already has all the other connections allowed it. In this case one of the other points is chosen randomly as a new origin and the regular process is continued. The incomplete shape of Fig. 2*a* is shown in a completed form in Fig. 2*b*.

As is the case with all the methods described in this paper, this method is completely objective. The resulting figure could be reproduced, if necessary, from a set of coded instructions consisting only of the numbers originally selected from the table.

Unlike Method 1, Method 2 usually generates shapes containing some angles in addition to all those at originally plotted points. This difference is emphasized by Rule *c* of Method 2.

In Method 1 there are no restrictions on the ways in which the plotted points may be connected except that (*a*) the figure must be closed, and (*b*) connecting lines may not cross, i.e., the completed figure may have angles only at the original points.

In Method 2, on the other hand, there may be "emergent" angles at places other than originally plotted points, and the figures produced tend to be characterized by "good continuation." Again, it is Rule *c* of Method 2 which causes many of the perimeter lines of the final figure to be continuations of other perimeter lines.

Comparing the two methods in terms of the informational content of the shapes produced shows that in Method 1 information (in addition to that required to locate the original points) is used only in connecting the interior points to the sides of the original perimeter, whereas in Method 2 information is used in making all connections between plotted points. For this reason a Method 2 shape composed of n original points and containing $n+k$ angles (k representing the number of "emergent" points) will contain more information than a Method 1 shape composed of n original (and final) points. Because of the good continuation introduced into the figure, however, the Method 2 shape having $n+k$ points will contain *less* information than would a Method 1 shape having $n+k$ original points.

Method 3. Fitts, Weinstein, Rappaport, Anderson, and Leonard (15) have developed a technique for constructing "metric" figures, the informational content of which may be easily and accurately determined. Starting with a somewhat smaller matrix—say, 8×8—the number of cells to be filled (from the bottom up) in each column of the matrix is randomly determined. This method produces shapes which belong to a relatively small stimulus-domain and which are equal in informational content. A variation of this method involves allowing each possible column-height to appear only once in each shape, with the order of appearance determined randomly. This second stimulus-domain contains members which are equal in area and, consequently, contain less information than the shapes first described. Still another variation may be introduced by reflecting each shape on one of its axes to produce a symmetrical shape containing no more information than its nonsymmetrical predecessor. Examples of these various classes of metric figures may be found in Reference 15.

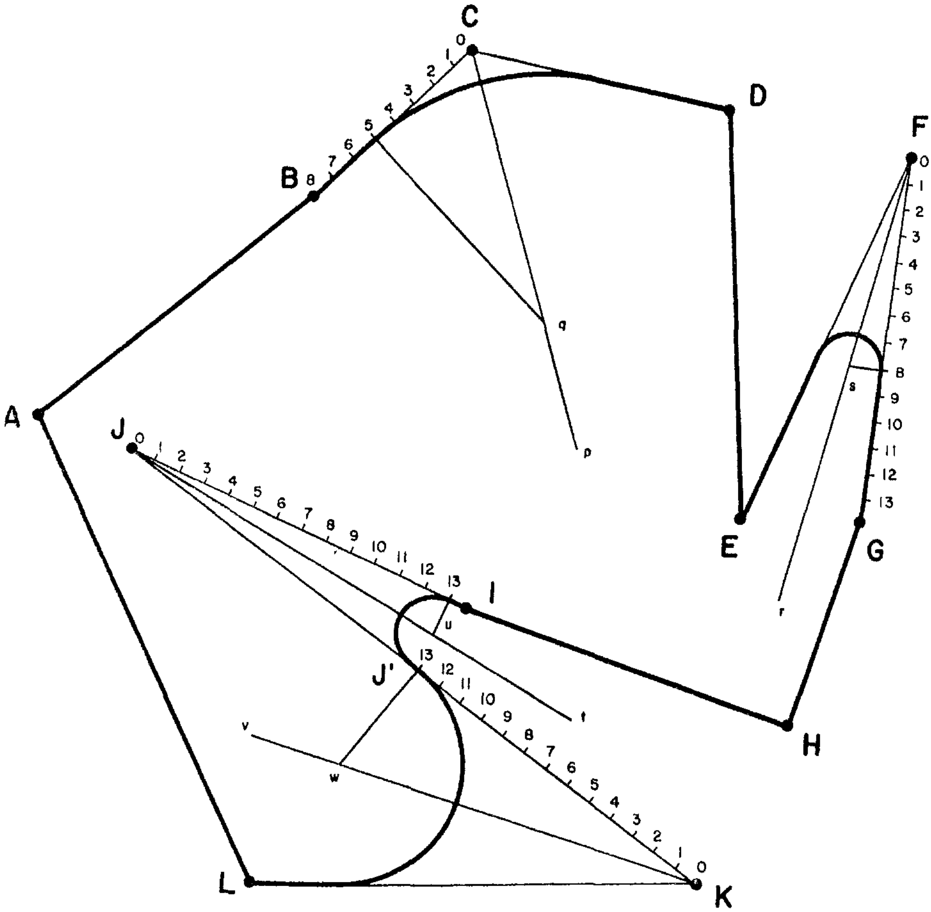

FIG. 3. METHOD FOR INTRODUCING "RANDOM" CURVES INTO AN ANGULAR NONSENSE SHAPE. THE ORIGINAL SHAPE IS THE SAME ONE WHICH APPEARED IN FIG. 1*c*

Closed Contours—Curved Shapes

Method 4. This method describes a procedure for making wholly or partially curved shapes from the angular shapes constructed by Method 1 or 2. This procedure may appear to be somewhat involved, but actually it requires more time to describe than to perform. Essentially, it consists merely of replacing angles with inscribed arcs, of curvature chosen randomly within limits imposed by the figure.

For purposes of demonstrating the method, let us start with the shape described and constructed under Method 1 (Figs. 1*a*–1*c*). It is decided (arbitrarily or randomly) that four of the twelve angles are to be curved. Let us suppose that Angles *C*, *F*, *J*, and *K* (Fig. 3) are chosen. (For convenience of exposition the angles have been assigned the letters *A* through *L*.) The first step in the process consists in constructing line *Cp*, which is the *bisector* of ∠*BCD*. Then, the shorter of the two arms of the angle (in this case, line *BC*) is divided into equal units. These units may be chosen for convenience. For example, Fig. 3 was constructed on a 100×100 matrix having matrix units equal to 0.20 in., and Line BC was arbitrarily divided into segments of 0.25 in. each. It should be noted that the divisions of the line are numbered in sequence, starting always from the *apex* of the angle.

One of these numbered points on line BC is now chosen at random and a perpendicular from Line Cp to it is constructed (Line 5-q). This line (5-q) now becomes the radius of an arc which is inscribed within $\angle BCD$. The arc is tangent to Line BC and Line CD at points equidistant from C. Thus, $\angle BCD$ has now been replaced by a curve (actually, two linear segments and an arc) going from B to D.

Point F has been curved by the same process. Angle EFG is bisected by line Fr, and Line FG is divided into equal segments. Division 8 having been chosen at random, line 8-s is constructed and used as a radius for inscribing a curve within $\angle EFG$.

The next two constructions demonstrate the complex curvature which may result when successive points are chosen to be curved. Point J is curved by the process described above, with Line 13-u being used as the radius of an arc inscribed within $\angle IJK$. However, in curving Point K it is necessary to inscribe an arc within $\angle J'KL$, not within $\angle JKL$. Point J' is the point at which the arc constructed with radius 13-u becomes tangent to line JK.

If it is so desired, all the points of an angular figure may be curved. It should be noted, however, that the *shorter* arm of every angle is divided into segments, and that its divisions are numbered *beginning with zero.* If the zero is the random choice, the resulting curve will have zero radius, i.e., that angle remains as originally drawn.

Method 5. Angular shapes can be changed into curved shapes by a process of photographic blurring. The figure is first photographed and then, with the help of an enlarger, is printed out-of-focus on high contrast paper. The resulting image has a contour which is curved, but which is also graded in density. A repetition of the process of photographing and printing, however, will eliminate the density gradient, producing a shape with contours which are rounded and well-defined. The amount of blur may, of course, be carefully controlled, and a graded series of curved shapes may be made from a single prototype shape.

Open Contours.

Method 6. There are many ways in which open-contour nonsense shapes may be constructed from a table of random numbers, but all that we have used have been variations on one basis method. Starting from the approximate center of a matrix of convenient size, a line is drawn to one of the eight intersections nearest the starting point. These eight intersections (or, more generally, directions) have been assigned numbers as shown in Fig. 4*a*. The intersection on the graph paper at which the first line terminates becomes the origin for the second line to be drawn, and so on. A difficulty with this method is that there is no intrinsic criterion for completeness in such a figure. One objective rule is to determine, before beginning the construction, the total number of digits to be selected from the table and to consider the figure complete when that number of lines has been drawn.

Many variations on this basic tech-

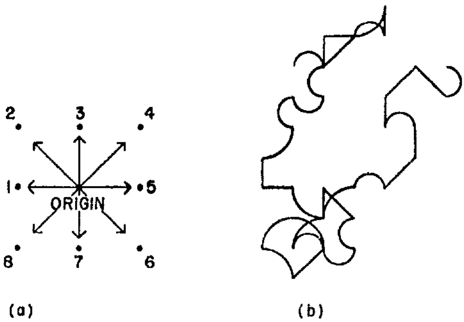

FIG. 4. CONSTRUCTION OF OPEN-CONTOUR "RANDOM" SHAPE: *a* NUMBERING OF POSSIBLE INTERSECTIONS; *b*. TYPICAL NONSENSE SHAPE

nique may be introduced. For example, for some purposes it may be desired to allow only four directions in which the contour may vary; also, the length of each line may be determined randomly as well as the direction. Partially or wholly curved contours may be produced by this method as follows: the radius of curvature of the arc drawn to connect successive intersections along the horizontal and vertical axes of the matrix is set as one-half the length of a matrix unit. To connect two intersections diagonally separated, the arc would have a radius equal to one matrix unit. Thus, for example, one might determine randomly for each line constructed: (*a*) which two intersections will be connected, (*b*) whether the connection is to be linear or curved, and (*c*) the direction of curvature. Figure 4*b* was drawn by this technique. Additional variations on these methods may be provided by using semi-log, log-log, or polar coordinate matrices on which to construct the nonsense contours.

Patterns

Method 7. Although the more obvious ways of generating random patterns have been used by a number of investigators, the possibilities of this approach to the construction of complex visual displays have never been adequately explored. In general the practice has been to construct a matrix of some given size and then to determine randomly which cells are to be filled. Patterns of dots were constructed in this fashion by Kaufmann, *et al.*, (**19**), French (**16**), and Klemmer and Frick (**20**), for example. Attneave used the same approach, including the introduction of a symmetry factor, in a study of the effect of redundancy on memory for patterns (**4**). In another slight variation Arnoult used random shapes as elements in constructing random patterns for use in a learning experiment (**2**). Patterns generated in this fashion are very attractive as stimuli because it is usually possible to compute fairly precisely the informational content of the display.

Systematic Variations

Frequently it is desired to construct "families" of shapes having known physical relationships among the individual members. Again, there are many possible techniques for accomplishing this end. The following two methods represent two kinds of systematic variations which have recently been used.

Method 8. A *prototype* shape is constructed by any of the methods so far described. Then, each point is moved to a new location and the connecting lines redrawn as before. In moving the points, any of the following parameters may either be held constant or varied randomly: (*a*) the number of points moved, (*b*) the particular points moved in making successive variations on the same prototype, (*c*) the distance through which a point is moved, and (*d*) the direction of movement. A number of variations made from a given prototype will form a distribution of shapes which "vary about" the prototype. Stimuli of this sort were used recently by Attneave in testing the hypothesis that knowledge of the prototype shape, or "schema," would facilitate discrimination of the variations in paired-associate learning (6), and by Arnoult in a study of the effect of predifferentiation training on recognition (1). A typical prototype shape and its variations are shown in Fig. 5.

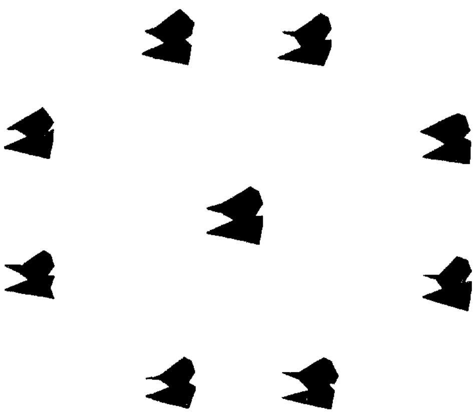

FIG. 5. A PROTOTYPE SHAPE AND "FAMILY" OF RANDOM VARIATIONS

Method 9. A somewhat different technique for creating "families" of shapes has been developed at Stanford by LaBerge and Lawrence (23). Initially, a random shape is constructed by a method essentially the same as those described in Method 1 and Method 2 (actually, LaBerge and Lawrence simply connected randomly chosen points into the polygon of minimal perimeter). Then, each point on the contour is assigned randomly chosen "*x*" and "*y*" increments to its coordinates, and these new coordinates are plotted and connected on a fresh matrix. These *same* increments are then added to the new coordinates and a third figure is constructed. This process may be continued until one has constructed a row of, say, six figures, each differing from its immediate neighbors by a constant amount of distortion as measured by the distance through which the points move. The next step is to label the former "*x*" increments as "*y*s" and the former "*y*" increments as "*x*s." These new increments are added to the coordinates of the points of all six of the figures already constructed, and the process of constructing successive shapes is repeated until there is a column of six shapes for each of the original six shapes. The final result is a matrix of 36 shapes in which any two adjacent shapes in a row or column are equally spaced in terms of the average distance the points have moved. Matrices of stimuli of this sort are currently being used by LaBerge and Lawrence in studies of transfer.

As has been emphasized a number of times in the preceding discussion, these methods for constructing "random" shapes are only a few which have been selected to show some of the *classes* of shapes which can be constructed. The number of different sets of rules which can be developed for plotting and connecting points taken from a table of random numbers is limited only by the fertility of the individual experimenter's imagination. It should be reiterated, however, that using stimuli constructed by these "random" methods does not insure that the generalizations resulting from the research will be pertinent to all other kinds of visual stimuli. It guarantees only that the results will be generalizable within a particular stimulus-domain, i.e., to any other stimuli constructed *by the same rules.*

Analysis of Natural Forms

Let us now return to a problem which the methods discussed in the previous section by no means obviate. We still need a technique, or a set of techniques, by means of which physical measurements of a psychologically relevant sort may be obtained for forms which we have not constructed ourselves. Any method of "random" construction must employ some set of rules, either arbitrary or otherwise, and these rules will strictly determine the class-characteristics, or statistical parameters, of the shapes constructed. We should like to be able to devise rules such that our synthetic shapes might possess the statistical characteristics (but not the familiarity) of natural shapes to which we wish to generalize. At present, we lack not only a factual knowledge of the values of these statistical parameters, but also a methodology to guide us in their determination. Likewise, when some experimental variation of form is found to produce a certain effect in the laboratory, it is necessary that the variable in question be identifiable and measurable outside the laboratory if the results are to be generalized. Unfortunately, however, it is much harder to measure form than to manipulate it.

Relatively few scientists have seriously applied themselves to the problems of analyzing and describing form; these problems seem to have fallen into the cracks between sciences, and no general quantitative morphonomy has ever developed. D'Arcy Thompson's *Growth and Form* (27) is virtually the only major work in the field: it is a fascinating and impressive book, but its contribution to the identification of psychophysical variables is limited. Rashevsky, whose work in mathematical biophysics is in some respects a continuation of Thompson's, has been more directly concerned with psychologically relevant measures of form.

Abstraction of contour. Considering that the first step in the analysis of a shape is the abstraction of its contour, Rashevsky (**25**, p. 449) devised a simple hypothetical nerve-net with this function. Suppose that the stimulation of the retina is projected to some central area as an activity of sharply localized excitatory fibers and of inhibitory fibers slightly more diffuse in their projection. If certain constants of the system have proper values, excitation from any area of uniform brightness will be suppressed, except at a contour where such an area is bounded by a darker one which provides less inhibition.

This nerve-net has a fairly close analogue in the following photographic process. A negative and a positive transparency, separated by a thin plastic sheet, are precisely superimposed so that they "cancel" each other when viewed from a right angle. A print is made by transmitting light from a diffuse source (e.g., the ground glass of a contact printer) through the superimposed positive and negative to a high-contrast paper placed in contact with the negative. In the case of a black object on a white ground, or vice versa, light can angle through both positive and negative only at the contour, and the resulting print is indistinguishable from an outline drawing of the object. In the case of more complex pictures, the abstraction of sharp brightness-gradients preserves texture, as well as contour: this is illustrated clearly in Fig. 6. A picture obtained in this way may be thought of as a differential (with respect to brightness) of the original, involving a "delta" of finite magnitude. If

FIG. 6. A DIFFERENTIAL PICTURE

The photograph of D'Arcy Thompson from which this was derived is by Björn Soldan; it appeared originally in *Isis* and was reproduced in the August, 1952, *Scientific American.* In the original, the lightest portions are Thompson's forehead and beard, and the darkest portion is the back of his coat. These have approximately equal brightness in the differential picture.

a smaller "delta" had been taken in the derivation of Fig. 6 (by reducing the space between the superimposed positive and negative), the iris and pupil of Thompson's eye, for example, would appear in outline instead of as a black dot.

In 1948 one of the authors (Attneave), in collaboration with John M. Stroud, attempted to develop this photographic technique to a degree of precision such that the total reflectance of the differential picture might serve as an index of the complexity of the original. That attempt was unsuccessful for several reasons, having to do chiefly with the unreliability of photographic operations: e.g., the initial step of making a positive and a negative which would adequately cancel always required considerable cut-and-try. It may be added that the process is a close relative of one which has long been used to produce a "bas-relief" effect, and that the Eastman Laboratories have recently employed a similar technique with color film to obtain photographs which look remarkably like paintings.

An electronic device lately developed by Kovasznay and Joseph at the National Bureau of Standards appears to accomplish much the same result as the photographic process described above, but in a manner subject to more precise control. The beam of a cathode ray tube, moving in a complex scan which covers the field in two orthogonal dimensions, transmits light through a photographic transparency to a photoelectric cell. The electrical signal thus generated is differentiated and squared electronically, and then fed into a receiving scope where it modulates a beam synchronized with the transmitting beam. Illustrations of the results, which are presented in the descriptive note of Kovasznay and Joseph (21), could be mistaken for the efforts of a somewhat naive artist.

A group of engineers in the Lincoln Laboratory of M.I.T., including Oliver G. Selfridge, Gerald P. Dinneen, and Marshall Freimer, are currently experimenting with the use of digital computers to perform operations relevant to object identification. They have been successful in programming a contour-abstracting operation; this is preceded by an averaging operation, which rids the figure of irrelevant detail, and followed by an operation which abstracts angles, or regions of high curvature, from the contour (26).

The mere abstraction of contour, whether by an objective process or with the aid of the experimenter's own perceptual machinery, does not in itself constitute quantification. It does, however, contribute to the isolation of that which is to be quantified: i.e., *form*. Whenever we speak of *form*, we are referring to a somewhat vague set of properties which are invariant under transformations of color and brightness, size, place, and orientation; our definition may or may not be extended to specify invariance under projective (or perspective) transformations. Contour is characterized by invariance under color and brightness transformations. Attneave (3) has previously pointed out the related (though not equivalent) fact that contours are regions of relatively high informational content.

Analysis of contour. There are various practical reasons for wishing to be able to describe a contour in terms which are independent of its size, place, and orientation. For example, subjects are often required to draw figures from memory: such drawings cannot be fairly evaluated by any simple method of superimposing a drawing upon the original and measuring deviations, because of differences in scale, etc. If both the original and the reproduction could be represented in terms descriptive of form alone, they could then be compared objectively.

Such a representation may take the form of a single function. If the reciprocal of the radius of curvature of a closed contour is plotted against distance along the contour, a periodic function results. This function may be *normalized* (i.e., rendered independent of the scale of the original figure) by assigning a value of unity to the perimeter of the figure and expressing radius of curvature in comparable terms, or by setting equal to unity the area under one period of the function. An angle is represented by a vertical line which rises (or falls, in the case of a concave angle) to infinity; a spike of this sort, of infinite height, infinitesimal width, and determinate area, is the so-called δ-function of Dirac, and is amenable to mathematical treatment.[3]

If one feels more comfortable dealing with finite ordinates, the following system may be used. Imagine a miniature tricycle, guided over a course such that a point midway between the rear wheels precisely follows the contour. The angle θ by which the front wheel deviates from a forward position may be plotted against distance travelled by the front wheel to give a periodic function descriptive of the contour. The front wheel will move in an arc concentric with the segment of the contour being followed. Wherever an angle occurs in the contour, the angle θ of the front wheel will be 90°; thus the function will always have some value between plus and minus 90°. Radius of curvature, r, is related to θ by the equation $r = L \cot \theta$, in which L is the distance between the front and rear wheels. Normalizing may be accomplished by giving the perimeter of the figure unit value, and setting L at some standard fractional value. If L is made to equal $1/2\pi$, regular polygons will be represented by square waves regularly alternating between 0 and 90°, a circle will become a horizontal line with an ordinate of 45°, and certain other regularities will be uniquely represented; this value of L is somewhat large for convenient use with more complex shapes, however. The interested reader will have little difficulty in working out further details of the system. It has the advantage of specifying an actual measuring device which is practical and simple to construct. Automatic recording of the function could be arranged with two pairs of selsyns: one translating the rotation of the front wheel into a movement of the recording paper; the other coupling the angular position of the front wheel with the position of a recording pen.

Both of the functions just described have a serious disadvantage. Suppose we wish to compare two shapes which have a part-to-part or part-to-whole similarity—say, the outline of a cow's head with the outline of a whole cow. The normalizing factors which will be employed on a basis of perimeter or area will obviously not be such as to give comparable representation to the similar portions of the outlines.

The method next to be described avoids this difficulty, though it is not without limitations of its own. Instead of describing the contour by

[3] This system of representation has been developed in considerable detail by Oliver Strauss (personal communication).

means of a continuous function, we may attempt to analyze it into parts which are individually homogeneous, and hence amenable to approximate description in terms of a few standardized dimensions. It is usually possible to construct a polygon about a figure made up of complex lines and curves, as in Fig. 7, by drawing tangents (*a*) at points of zero curvature (e.g., *CD*, *IJ*, etc.: whenever a curve changes from concave to convex, it must have an intermediate point of zero curvature), (*b*) at points of minimal curvature, where a decrease in curvature is followed by an increase (e.g., *FG*), and (*c*) at discontinuities of slope, or angles (e.g., *AB*, *GH*, etc.). The series of lines thus formed may be described simply by stating the slope and length of each line in succession, but this description is peculiar to a given orientation and size of the figure. It may be rendered orientation-free and scale-free by specifying instead, for each *pair* of adjacent segments, (*a*) *the change in direction* (*in degrees*), and (*b*) *the change in the logarithm of length*, as the contour is followed in a clockwise direction.[4] Curves are treated as "rounded-off" angles: i.e., a curve is approximated by an arc located tangent to two successive lines of the polygon we have been discussing. In most cases, the size of the arc will be limited by the length of the shorter of the two segments. Hence curvature is conveniently expressed by a third coordinate specifying (*c*) *the proportion of the distance between the apex of the angle and the end of the shorter segment at which the arc best approximating the curve is tangent.* This coordinate will usually have some value between 0 and 1.0, with 0 indicating an abrupt angle (radius of curvature equal to zero) and 1.0 indicating an arc which is tangent to the shorter segment at its end. In the case of Fig. 7, for example, (*c*) would have a value of 0 at *A* and *M*, a value of 1.0 at *G*, and a value of about .8 at *C*

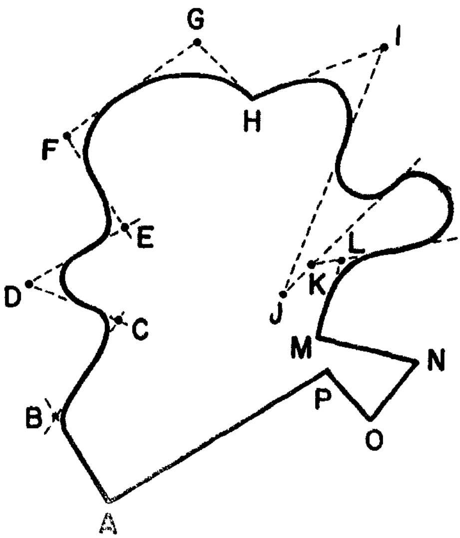

Fig. 7. Illustration of Method for Quantizing Irregular Contour

(note that the arc best approximating a curve will not necessarily have the same point of tangency as the original curve). When the arc approximating a curve turns through more than 180°, as in the case of the bulbous projection in the *JKL* region, the value of (*c*) will not remain between 0 and 1, since some of the points of tangency are on extensions

[4] Several other possible pairs of coordinates convey the same information. What is required, essentially, is to describe the shapes of successive segments of the polygon, taken in pairs. Measures of any two angles of such a triangle, or any two ratios of sides or differences between logarithms of sides, or any combination of an angle and a comparison of sides, is adequate to specify the shape of the triangle. The combination above is chosen for its intuitive appeal; also because errors of measurement have a more uniform effect on these coordinates than on certain others.

of segments of the polygon, rather than on the segments themselves. The values of (*c*) associated with *J*, *K*, and *L* would be about 3.8, 3.6, and .5, respectively.[5]

The reader will recognize this system of analysis as essentially the reversal of a method for constructing "random" shapes which was described in the previous section. The system has several advantages: (*a*) It yields a description which does not vary with size and orientation. (*b*) Since the use of a general normalizing factor is avoided, part-similarities between the contours of two objects are reflected in their numerical descriptions. Likewise, repetitious sequences of elements in the same contour (but not parallel lines) are reflected, and could be quantified by an autocorrelational technique. If two similar shapes (e.g., an original and a subject's reproduction from memory) were compared by cross correlation of their numerical descriptions, it would be desirable to calculate values for several "displacements" of one set of coordinates upon the other (as in autocorrelation), in order to allow for qualitative omissions or additions of elements. (*c*) There is reason to believe that the *number* of tricoordinates required to describe a shape constitutes a first order approximation of its psychological complexity (i.e., the number of psychologically discrete parts which it contains). Fehrer (**14**) used a similar measure (number of internally homogeneous lines) on her figures, and found that complexity, so measured, was closely related to difficulty in a reproduction-learning situation. Attneave (**5**) recently confirmed that the number of sides in a polygon is the primary determinant of its judged complexity. A better approximation would require some adjustment for repetitious sequences of elements, mentioned above (see Rashevsky, **25**, p. 486 ff.; also Attneave, **3**, **4**).

The major disadvantage of the system is that some figures (spirals are an obvious case) do not yield unique descriptions. This limitation arises inevitably from the approximation of all curves with straight lines and

[5] The two sets of numbers below, which are presented as a demonstration of the practicability of the system and as an amusement for the reader, describe recognizable profiles of the two authors. Successive tri-coordinates are given, thus: a_1, b_1, c_1; a_2, b_2, c_2; etc. In the actual reconstruction of a contour from such coordinates, a line of any desired length and slope is drawn to start. The first triad of coordinates gives the relationship of the second line of the contour to the arbitrarily drawn starting line, and so on. Cumulative error will be avoided if the values of *a* are cumulatively added to the slope, in degrees, of the starting line (with due regard for the circularity of the scale) to obtain the slope of each segment; likewise if the values of *b* are cumulatively added to the $\log_{10}$ length of the starting line to obtain the $\log_{10}$ length of each segment. Positive values of *a* denote clockwise turns; negative values counterclockwise (consistent reversal results in a mirror-image). In specifying values of (*c*), the symbol "<" is used to mean "less than .1," i.e., that the angle is rounded to a slight, practically unmeasurable, degree.

+61, −.56, .3; −84, +.08, <; +27, −.04, .1; +35, +.66, .2; +155, −.06, 1.0; −145, −.39, 0; +32, +.13, 1.0; +27, −.43, <; −56, +.61, .1; +107, −.38, .4; −69, −.12, .1; +37, −.62, <; −88, +.38, 0; +105, +.11, 1.0; −79, +.33, .1; +86, +.34, .5; −20, −.14, <; −25, −.76, .4; −66, +.23, 0; −55, −.15, 1.0; +30, +.09, 0; +113, −.02, 0; −56, +.15, .1.

+43, +.14, .4; +26, −.22, .4; −31, +.02, .3; +37, +.21, .4; +36, −.21, .6; +41, −.41, .3; −50, +.45, 0; −6, −.18, <; +29, +.03, .8; −43, +.26, <; +12, −.18, <; +88, −.11, .5; +23, −.27, .4; −124, +.31, .5; +90, −.12, .8; −29, −.28, <; +70, −.23, 1.0; −114, +.39, 0; +66, +.16, .3; −65, −.04, .4; +44, −.17, <; −43, +.77, .7; +121, −.28, .9; −26, +.04, <; +105, −.08, .7; −62, −.49, .6; +26, +.09, <; −113, +.13, .3; −25, +.25, .5; +44, −.26, .8; −83, −.24, 0; +19, +.42, 1.0.

arcs, and the ignoring of higher-order invariances. It is interesting to speculate that the system might be to some degree psychomimetic even in this limitation, and that objects for which it does not yield unique descriptions are less likely to evoke reliable perceptual responses, with the result that they may be perceived as "amorphous" or "unstable," and be difficult to remember.

Measuring operations like the foregoing, which involve following about a contour, are laborious to accomplish manually. It appears, however, that they are quite amenable to automation by electronic and mechanical means. For example, an electronic contour-follower, described by Beurle (7), has already been constructed. A point of light is moved rapidly through a very small circle; when its path crosses the contour, a signal is obtained. The phase relationship of this signal to the circular movement is used to guide the circle along the contour, i.e., to move the point of light about the contour in a cycloidal path. A record of the movement of the circle, taken from the servo control loop, constitutes a description of the shape which may further be transformed and analyzed by computer-type circuits.

Measurement of gestalt-variables. We have been considering analytical systems by means of which the formal properties of contours may be described in detail. Also of interest is another set of variables which do not provide a description from which the shape can be reconstructed, but which do abstract important properties of the shape as a whole. We shall refer to these as "gestalt-variables," or "gestalt-measures," even when they serve to summarize some quantizing or analytical process: e.g., the number of sides in a polygon is such a variable; so is the number of tri-coordinates necessary to describe a shape by the system discussed above. Likewise, the mean value of the c-coordinate in that system might be taken as a crude measure of overall curvedness-vs.-angularity. It should be clear that the "statistical parameters" of populations of shapes, referred to earlier, necessarily pertain to distributions of certain gestalt-measures.

The more restricted notion of a gestalt as a system in which every part is affected by every other part has been incorporated by Rashevsky (25, p. 451 ff.) into a hypothetical nerve-net. Suppose that the contour of an object is projected to some sheet of neurons in the cortex as an isomorphic excitation (Rashevsky's mechanism for contour-abstraction has already been described). Suppose further a distribution of inhibitory fibers such that, in the next higher projection area, every point on the contour (i.e., every excited neuron) receives inhibition from every other point in an amount which varies as a function (presumably decreasing) of the distance between the points. At this level, the various neurons to which the contour is projected will retain more or less residual excitation, depending upon the degree to which each is isolated from the others. A given contour will be characterized (though not uniquely) by some distribution of residual excitations which will be invariant with respect to its place and orientation in the field (but not with respect to its size). The integral, or mean, of this distribution would constitute a measure of the "simplicity" or compactness of the figure (provided size were held constant, or corrected for); e.g., a circle would have the highest possible value, since its points are as far from one another as is possible in a closed contour, and jagged or sinuous shapes would have low values (see Householder, **18**). The neurological terms in which this model is presented need not be taken too seriously; Rashevsky's basic idea might equally well be applied to the programming of a man-made computer, or to a series of photographic operations.

Deutsch (**12**) has recently suggested a model for shape perception which is somewhat akin to Rashevsky's. Since it may be described very simply in terms of geometrical concepts, we shall ignore the neural mechanisms which Deutsch proposes as its basis. Suppose that a perpendicular is drawn to a closed contour at every point along its length. Each such perpendicular will contain a segment which lies inside, and is bounded by, the contour. The lengths of these segments will have some distribution which will depend upon the shape of the contour; this distribution may be rendered size-invariant by expressing the length of each segment as a proportion of the length of the contour. In the case of a circle, a square, or any other regular polygon with an even number of sides, the distribution will consist of a single spike, since all the segments will be of equal length. Deutsch suggests that the most primitive mechanism of form-discrimination may abstract a distribution of this sort; at the human or primate level it would obviously need supplementing with some finer mechanism, perhaps one involving contour-following. He points out that rats have more difficulty discriminating a square from a circle than from a triangle, and predicts further that regular polygons with even numbers of sides should be more difficult to discriminate from one another than from odd-sided polygons.

Merely to order shapes along a compactness-dispersion continuum requires nothing so elaborate as the Rashevsky model outlined above. The relationship of the perimeter of a shape to its area provides an attractively simple means of measuring this characteristic. The quotient P/A, which has been employed by some investigators (**8**, **17**), is unsatisfactory from our standpoint because it varies with size as well as with shape, but either P^2/A or $P/\sqrt{A}$ is size-invariant. These ratios may be transformed in various ways to suit the user's convenience; e.g., the measure

$$D = 1 - \frac{2\sqrt{\pi A}}{P}$$

expresses dispersion as some number between zero and one, assigning zero value to the most compact figure possible, the circle. Dispersion (as measured by any such relationship of perimeter to area) is not the same as complexity (in the sense of number of parts). Although a deeply convoluted or jagged figure will indeed tend to have a high dispersion value, so will a very thin rectangle or ellipse.

Bitterman, Krauskopf, and Hochberg (**8**, **22**) have found that under conditions of low illumination or short exposure, shapes are perceived in much the same way as if they were physically diffused, or blurred. These experimenters created a physical diffusion model by cutting filter paper into various shapes and impregnating it with an inhibitor of bacterial growth. This inhibitor was then allowed to diffuse from the paper into bacterial cultures. The shapes which most resembled each other after diffusion were those most often confused under adverse viewing conditions. Likewise, identification of impoverished stimuli was most impaired in the case of shapes characterized by relatively small detail, which would be averaged out in a diffusion process.

These findings are interesting and important, but the clumsy and somewhat bizarre bacterial model does not lend itself to quantitative prediction. There is no apparent reason why it might not be replaced with a model employing optical blur, in which case diffusion would be measured by the radius of the blur circle. An image may readily be blurred to a measurable degree in an ordinary photographic enlarger, and then resharpened by means of high-contrast paper or film (cf. Method 5 under "The Construction of Stimuli"). This resharpening process introduces another parameter, that of the black-white threshold to be used in printing. It is easiest photographically simply to employ long exposure and development, with the result that a white-on-black figure will diffuse outward into the field to the full extent of the radius of the blur circle. If it is desired that concavities and convexities be affected symmetrically, however (note that a psychological question requiring an empirical answer is thus raised), it is necessary to resharpen the image into black and white about some intermediate gray such that a linear contour between black and white fields will be restored to its original position.[6] This may be accomplished with the aid of a suitable test-figure.

[6] Dinneen (13) has succeeded in programming a digital computer to perform averaging-and-resharpening operations of almost exactly this sort. His paper, which contains copious illustrations of the effect of varying resharpening threshold, is recommended to the reader who finds the above discussion insufficiently informative.

Over a wide range of values on the resharpening threshold parameter, the process of blurring and resharpening will decrease the dispersion (P^2/A, or D) of any shape except a circle, which is already the most compact shape possible. For any such value, dispersion will tend to decrease as amount of blur increases, but the form of this function—which we shall call a *blur-response function*—will vary with the shape involved and will describe certain important characteristics of the shape. Since the decrease in the function is associated with the "washing out" of progressively larger detail as the blur circle increases in size, any sharp drop indicates that the shape contains considerable detail of a magnitude indicated by the blur circle at that point. The blur-response function (or, perhaps better, its derivative) is thus a potential aid in the statistical evaluation of "magnitude of critical detail," which Bitterman, *et al.*, found to be of primary importance in determining the identifiability of an impoverished shape (8). A full exploration of the properties of such functions (particularly in the case of shapes characterized by certain types of regularity, or redundancy) is beyond the scope of this paper; our purpose here is merely to suggest their feasibility and possible usefulness. One further point should be made, however; neither the blur-response function nor any other gestalt measure can possibly predict the relative identifiability of shapes except in a limited, statistical way. The kinds and degrees of similarity which an impoverished shape bears to all the other shapes with which it might be confused will clearly affect the difficulty with which it is identified (quite apart from any intrinsic properties it may have), and these similarities may be evaluated, if at all, only by recourse to analytical measures. A particular detail in a shape may or may not be *critical* to identification, depending upon the specific discriminations which identification requires.

Gestalt measures, as defined earlier, all involve a reduction in the dimensionality of figures (sometimes, though not necessarily, to a single dimension) with a concomitant discarding of information. The number of operations by means of which a shape may be "collapsed" to lower dimensionality is indefinitely large, as Selfridge (**26**) has recently pointed out. At the simplest level, for example, we may literally collapse a shape upon any spatial axis by plotting, as a function of distance along that axis, the thickness of the shape in the orthogonal dimension (**26**, Fig. 3). The axis involved need not even be linear; e.g., it might be a circle about the center of gravity of the shape (cf. Pitts and McCulloch, **24**).

Of all the conceivable physical measures of shape, analytical as well as gestalt, there are undoubtedly many that have little or no value from a psychophysical point of view. On the other hand, it appears unlikely that any single system of physical measurement can be optimal for all psychophysical situations: in other words, we are suggesting that form perception involves a number of different psychological mechanisms which function in a complementary, and to some degree overlapping, manner. Unfortunately, there is no quick and easy way to determine which physical measurements have greatest psychological relevance; only experimentation can answer this question. The preceding discussion and review may at least serve, however, to alleviate somewhat the paucity of hypotheses which in the past has characterized this research area.

REFERENCES

1. ARNOULT, M. D. Recognition of shapes following paired-associate pretraining. Paper read by title at USAF—NRC Sci. Sympos., Washington, D. C., November, 1955.
2. ARNOULT. M. D. A comparison of training methods in the recognition of spatial patterns. *USAF Personnel Train. Res. Center, Res. Rep.*, No. AFPTRC-TN-56-27.
3. ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–193.
4. ATTNEAVE, F. Symmetry, information, and memory for patterns. *Amer. J. Psychol.*, 1955, **68**, 209–222.
5. ATTNEAVE, F. Physical determinants of the judged complexity of shapes. *J. exp. Psychol.* (in press).
6. ATTNEAVE, F. Effect of familiarization with a class-prototype on identification-learning of shapes. *Amer. Psychologist*, 1955, **10**, 400–401. (Abstract)
7. BEURLE, R. L. Discussion in W. Jackson (Ed.), *Communication theory.* New York: Academic Press, 1953. Pp. 323–326.
8. BITTERMAN, M. E., KRAUSKOPF, J., & HOCHBERG, J. E. Threshold for visual form: a diffusion model. *Amer. J. Psychol.*, 1954, **67**, 205–219.
9. BRUNSWIK, E. *Systematic and representative design of psychological experiments: with results in physical and social perception.* Berkeley: Univer. of California Press, 1947.
10. BRUNSWIK, E. *The conceptual framework of psychology.* Chicago: Univer. of Chicago Press, 1952.
11. BRUNSWIK, E., & KAMIYA, J. Ecological cue-validity of "proximity" and of other Gestalt factors. *Amer. J. Psychol.*, 1953, **66**, 20–32.
12. DEUTSCH, J. A. A theory of shape recognition. *Brit. J. Psychol.*, 1955, **46**, 30–37.
13. DINNEEN, G. P. *Programming pattern recognition.* Cambridge: Massachusetts Inst. of Technology, Lincoln Lab., 1955.
14. FEHRER, E. V. An investigation of the

learning of visually perceived forms. *Amer. J. Psychol.*, 1935, **47**, 187–221.

15. Fitts, P. M., Weinstein, M., Rappaport, M., Anderson, N., & Leonard, A. J. Stimulus correlates of visual pattern recognition. *J. exp. Psychol.*, 1956, **51**, 1–11.
16. French, R. S. The discrimination of dot patterns as a function of number and average separation of dots. *J. exp. Psychol.*, 1953, **46**, 1–9.
17. Hochberg, J. E., Gleitman, H., & MacBride, P. D. Visual thresholds as a function of simplicity of form. *Amer. Psychol.*, 1948, **3**, 341–342. (Abstract)
18. Householder, A. S. Concerning Rashevsky's theory of the "Gestalt." *Bull. Math. Biophys.*, 1939, **1**, 63–73.
19. Kaufmann, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. The discrimination of visual number. *Amer. J. Psychol.*, 1949, **62**, 498–525.
20. Klemmer, E. T., & Frick, F. C. Assimilation of information from dot and matrix patterns. *J. exp. Psychol.*, 1953, **45**, 15–19.
21. Kovasznay, L. S. G., & Joseph, H. M. Processing of two-dimensional patterns by scanning techniques. *Science*, 1953, **118**, 475–477.
22. Krauskopf, J., Duryea, R. A., & Bitterman, M. E. Threshold for visual form: further experiments. *Amer. J. Psychol.*, 1954, **67**, 427–440.
23. LaBerge, D. L., & Lawrence, D. H. A method of generating visual forms of graded similarity. *Amer. Psychologist*, 1955, **10**, 401. (Abstract)
24. Pitts, W., & McCulloch, W. S. How we know universals. The perception of auditory and visual forms. *Bull. Math. Biophys.*, 1947, **9**, 127–147.
25. Rashevsky, N. *Mathematical biophysics.* Chicago: Univer. of Chicago Press, 1948.
26. Selfridge, O. G. *Pattern recognition and learning.* Cambridge: Massachusetts Inst. of Technology, Lincoln Lab., 1955.
27. Thompson, D. W. *Growth and form.* New York: MacMillan, 1942.

Received March 19, 1956.