PHYSICAL DETERMINANTS OF THE JUDGED COMPLEXITY OF SHAPES [1]

FRED ATTNEAVE

Operator Laboratory, Air Force Personnel and Training Research Center

Many psychological tasks vary in difficulty with the complexity of the stimulus objects involved. Complex visual objects are not only harder to reproduce from memory than simple ones (4, 6), but also harder to learn by name (4, 7) and to match (11). *Complexity* is an ill-defined variable, however. No two of the experiments referenced above employ exactly the same operations of physical measurement, and an essential communality between them is not easy to specify in objective terms.

In the study reported here, ratings of the complexity of nonrepresentational shapes were obtained from a large number of Ss and related to measurable physical characteristics of the shapes. The results not only have interest in their own right, but also serve to indicate the physical variables most likely to be relevant to other tasks, like those initially mentioned, on which data are typically

harder to get and less precise. The relationship of judged complexity to informational content, "degrees of freedom," compactness, and certain other variables will be considered.

METHOD

Materials.—The stimuli were 72 "random" shapes, each constructed by the following general method. In a $k \times k$ matrix, n random points were plotted, i.e., all coordinates of the points were random numbers between 1 and k. The points were then connected randomly into a polygon of n sides. This connecting process involved two steps. First, peripheral points were connected into a convex polygon which enclosed all the points not included in its contour (as if a pin were stuck into each point and a rubber band snapped around the whole cluster). Second, the unconnected points were given a random order and each in turn was "taken into" a randomly chosen segment of the surrounding polygon (as if by hooking that segment of the rubber band over the interior pin). Since lines connecting the points were not permitted to cross, the number of alternative segments into which a given point might be taken could either increase or decrease as the process continued; hence the assignment of a random sequence to the unconnected points. Connections which placed certain points outside the polygon were permitted, though in consequence the ways in which such points could themselves be connected were restricted.

Some of the polygons thus constructed were further developed into curved figures by replacing angles with inscribed arcs of random

curvature. If the shorter of the two segments forming an angle was between d (an integer) and $d + 1$ matrix units in length, a random number between O and d was taken as the distance from the vertex of the angle to the point of tangency of the arc to be inscribed. [An illustrated and more detailed description of the construction methods outlined above is presented in a methodological paper by Attneave and Arnoult (5).]

A class of symmetrical figures was also constructed by the following operations. An asymmetrical polygon was reflected about a vertical axis passed through the point farthest to the right in the original polygon. This reflection resulted in a figure consisting of two parts: the original and its mirror-image, touching (ordinarily) at a single point—the point of reflection. In order to tie the two halves together into a single unified shape, a part of the area between them was filled in—specifically, the area between that pair of touching symmetrical segments which had the greater vertical component. Reflecting a shape resulted in a corresponding increase in area.

THE 48 ASYMMETRICAL SHAPES:
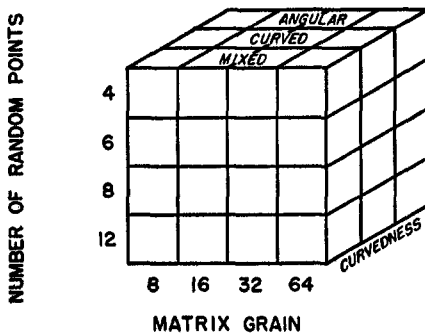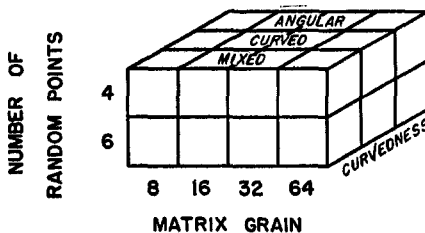

THE 24 SYMMETRICAL SHAPES:

FIG. 1. Plan by which the shapes were systematically varied.

In the construction of curved symmetrical shapes, the figure was first reflected and then curved—all curves being repeated symmetrically—except in the unique case of the arc associated with the point of reflection.

The following parameters of the shapes were varied in a quasi-factorial design (see Fig. 1):

1. *Matrix grain.* Four different matrix grains (8×8, 16×16, 32×32, and 64×64) were used in plotting the random points from which the shapes were constructed. Eighteen shapes were made with each grain. This variation in grain did not involve a concomitant variation in size. Moreover, the matrix lines did not appear when the shapes were displayed to *Ss*. The amount of information required to locate each point in the matrix (i.e., the amount of information taken from the random number table) was very nearly equal to 6, 8, 10, or 12 bits, depending upon the fineness of grain (these values are a trifle too great, since the selection of each point reduced by one the number of alternative positions which the next point might take).

2. *Curvedness.* One-third of the shapes, or 24, were entirely angular, one-third were entirely curved (i.e., all angles were replaced with inscribed arcs), and the remaining third were mixed. For each mixed shape, both the number and the identity of angles to be replaced with curves were randomly determined.

3. *Symmetry.* Two thirds of the shapes, or 48, were asymmetrical; the remaining third were symmetrical. The reason for this unequal division will be discussed in connection with the next variable.

4. *Number of turns.* Either 4, 6, 8, or 12 random points were initially plotted for the construction of each shape. The asymmetrical shapes were evenly distributed over these four classes with 12 in each class. The 24 symmetrical shapes, however, were equally divided between the 4- and 6-point classes. They were thus equivalent to the 4- and 6-point asymmetrical shapes with respect to number of *independent* sides or contour turns, and approximately equivalent to the 8- and 12-point asymmetrical shapes with respect to *total* number of sides or turns. (In referring to the present variable, we shall hereafter use the term *turns* in preference to "points," "angles," or "sides," since it applies equally well to curved and angular figures.) In the process of joining points into polygons, a point was occasionally "lost" when the two lines meeting at the point themselves formed a straight line; hence it was possible for a 12-point figure, for example, to have only 11 or even fewer turns in its contour. The actual number of turns, rather than the original number of points, was considered to be the

important variable in the subsequent analysis of data.

The four variables just enumerated were orthogonally related, as indicated in Fig. 1, and each possible combination of values on these variables was represented in one and only one of the 72 shapes, except as otherwise noted in connection with Symmetry and Number of Turns. Six of the shapes are reproduced in Fig. 2 to give the reader a general idea of their appearance. It should be understood that every shape was constructed from a completely independent selection of random points; e.g., the same shape never appeared in an angular and a curved version, nor in an asymmetrical and a symmetrical version.

In addition to these systematically varied parameters, the following two physical measures were taken after the shapes were constructed:

5. P^2/A. The square of perimeter, divided by area, was obtained for each shape. This is a size-invariant measure of dispersion, or non-compactness; see Attneave and Arnoult (5).

6. *Angular variability*. This measure represents the average *difference* between adjacent angles in a polygon. More precisely, it is the arithmetic mean (sign ignored) of the algebraic differences in degrees of slope-change (sign observed) between all successive or adjacent angles taken in overlapping pairs about the contour, convex angles being considered positive and concave ones negative. For example: suppose that the slope-changes associated with the angles of a four-sided figure are 150, 130, − 90 (a concave angle), and 170 degrees. The differences between adjacent angles will be 20, 220, 260, and 20 degrees. Angular Variability is the mean of these four differences, or 130. In the case of curved figures, measurement was made upon the polygons from which the curved shapes were constructed. Something like "good continuation" is quantified in this measure; a value of zero means that all angles of the polygon are equal, whereas high values are associated with jagged, irregular figures. As one might expect, this is correlated with the preceding measure, P^2/A: $r = .48$ over the 72 shapes.

Subjects.—One hundred and sixty-eight airmen basic trainees at Lackland Air Force Base rated the shapes for complexity.

Procedure.—The shapes were displayed, in a white-on-black figure-ground relationship, by projecting them individually to the front wall of the room in which Ss were seated. Each S was provided with a seven-category rating form containing a line for each shape and columns headed "Extremely Simple," "Very Simple," "Simple," "Medium," "Complex," "Very Complex," and "Extremely Complex." The instructions given contained no definition, either
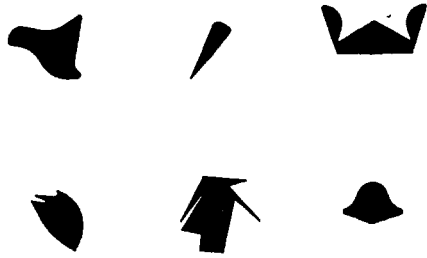

FIG. 2. Six of the 72 shapes used in the study.

explicit or implicit, of the terms "Simple" and "Complex."

Before actually making their ratings, the Ss were shown all 72 shapes in rapid succession (2-sec. exposure each) in order that they might, from the beginning, adapt their rating behavior to the range of stimuli with which they were to deal. In displaying the shapes for rating, each was exposed for 10 sec. with a negligible interval between stimuli.

The Ss served in four approximately equal groups which differed only in the ordering or sequence of stimulus presentation. The first sequence was a random permutation of the 72 stimuli, the second was an "inside-out" reordering of the first, and the third and fourth were reversals of the first two.

RESULTS

Ratings were scaled by the method of graded dichotomies (1)[2] in order to render them comparable to paired-comparison scale values. It turned out that mean ratings might as well have been used, since a plot of means-versus-scaled ratings showed little or none of the usual curvilinearity. The linearity of this relationship implies that the "widths" of the various scale categories were very nearly equal in terms of the dispersion of response distributions, though Ss had no special instructions to regard them as equal. Mean ratings ranged from .82 to 5.24

[2] This method, or the method of successive intervals, has been independently discovered and rediscovered, with minor variations, by Urban, Guilford, Thurstone, and me [see Guilford (8, Ch. 10)]. I continue to use my own version simply because it is the one most familiar to me.

(with "Extremely Simple" scored as zero and "Extremely Complex" as 6). Scaled ratings covered a range of 4.17 SD units with an arbitrary zero point.

The effect upon judged complexity of each of the six physical variables described in the previous section will now be considered.

Matrix grain.—This variable had no measurable effect on the ratings. In the original data there was a slight, nonsignificant suggestion of a trend toward greater complexity for finer grains, but even this failed to appear in the residuals from subsequent variables (*vide infra*).

Curvedness.—Likewise, whether the shapes were angular, curved, or mixed made no significant difference in judged complexity. After the removal of variance attributable to the important variables Symmetry, Number of Turns, and Angular Variability, a simple analysis of variance for effects of Curvedness was performed on the residuals. This yielded an F of 1.73 with $df = 2/69$, which is far short of the 3.13 required for the 5% level of significance. The mean rating obtained for curved shapes was lower by about .15 scale unit (i.e., SD units) than the means for angular and mixed shapes, which were almost exactly equal. Even if this difference were reliable, it would account for less than half of one per cent of the total variance of the ratings.

Symmetry.—Symmetrical shapes were judged *more* complex than asymmetrical shapes with the same number of *independent* turns by a mean of about .41 unit ($P = .001$, $t = 3.99$, $df = 69$). They were judged *less* complex than asymmetrical shapes with the same *total* number of turns by a somewhat greater amount, i.e., about .9 or 1.0 units. Reflecting a shape symmetrically is equivalent to an increase of about 19% in number

of independent turns, in terms of effect on judged complexity (see Equation 1, below).

Number of turns.—This was by far the most important variable. Plots of rated complexity vs. number of independent turns showed appreciable curvilinearity, which was rectified when the logarithm of the number of independent turns was used instead. Regression lines relating complexity to log turns were found, by the least-squares method, for symmetrical and asymmetrical shapes separately. Associated with the former was a correlation coefficient of .75; with the latter, a coefficient of .93. This difference is attributable chiefly to the restricted range of the symmetrical shapes. It was possible to adjust the lines to equality of slope, by the use of a weighted average, with negligible impairment of goodness of fit; the resulting parallel lines were separated by the .41 units mentioned above.

The relationships thus far were combined into the following equation:

$$J = 5.46 \log_{10} T + .41 S - 2.30 \quad (1)$$

in which J is judged complexity (on a scale in which the simplest figure is arbitrarily assigned a value of zero), T is number of independent turns, and S is a variable on which symmetry has a value of 1 and asymmetry a value of 0. This equation accounts for 82.5% of the total variance of the complexity ratings (see Table 1). For purposes of further analysis, a table of actual residuals from Equation 1 was made up by subtracting the predicted rating of each shape from the obtained rating.

P^2/A and Angular Variability.—It will be recalled that these two physical measures were intercorrelated to the extent of .48. Each was now correlated with the residuals from Equation 1. Angular Variability showed

TABLE 1

CONTRIBUTIONS OF THE INDIVIDUAL PHYSICAL
VARIABLES TO JUDGED COMPLEXITY

Physical Variable	Percentage Variance of Complexity Ratings Explained
Matrix Grain	. . .
Curvedness	. . .
Symmetry	3.8
Number of Turns	78.7
P²/A	0.9*
Angular Variability	7.1
Total	90.5

* P²/A actually explains 4.5% of the variance of the ratings, but of this 3.6% is shared with Angular Variability. The shared 3.6% is arbitrarily included in the 7.1% attributed to Angular Variability.

the higher relationship, $r = .64$, as compared with $r = .51$ for P²/A. One may question whether P²/A accounts for any variance not also accounted for by Angular Variability. The answer is probably affirmative, since the partial correlation between P²/A and the residuals, with Angular Variability held constant, is equal to .30, which is on the borderline of significance at the 5% level. Bearing in mind that the residuals from Equation 1 themselves represent only 17.5% of the total variance of the ratings, we may calculate that 3.5% of the total variance is uniquely predicted by Angular Variability, that .9% is uniquely predicted by P²/A, and that a further 3.6% is predictable from either Angular Variability or P²/A, or common to the two predictor variables.

The relationship between Angular Variability and the residuals from Equation 1 was also considered for symmetrical and asymmetrical shapes separately. The correlation coefficient was .77 for the former and .57 for the latter. After equating the regression lines in slope, the line for asymmetrical shapes was about .13 unit higher than that for symmetrical

shapes (this constant is accordingly added to the coefficient of S in Equation 2 below). However, the gain in predictive efficiency which results from treating symmetrical and asymmetrical shapes separately at this stage is scarcely worth bothering about; it amounts to only about .6% of the total variance of the ratings.

Adding Angular Variability as a predictor variable:

$$J = 5.46 \log_{10} T + .54 S + .005 AV - 2.91. \quad (2)$$

A second table of residuals from this equation was prepared and examined for possible minute effects of Matrix Grain and Curvedness, with the negative results mentioned earlier. It may be noted that although Curvedness per se has no demonstrable effect on the ratings, Curvedness is nevertheless represented in the variable P²/A (i.e., the curving of any shape by the method used necessarily decreases its P²/A value), and the component of P²/A which is affected by Curvedness does attain a marginal level of significance (see above).

DISCUSSION

The lack of effect of Matrix Grain on the judgments, though not intuitively surprising, does serve to demonstrate how the amount of information contained in a stimulus (from the experimenter's point of view) may be varied greatly without changing the apparent complexity of the stimulus. The amount of information gained by S in viewing a shape may very well be independent, within broad limits, of the amount of information used in this way by E in constructing the shape.

More interesting is the discovery that curved shapes are judged no more complex than angular shapes. One might reasonably expect that apparent complexity would depend upon the number of dimensions or degrees of freedom

associated with a stimulus, i.e., upon the number of numbers necessary to describe it in some coordinate system [this is essentially what MacKay (9) has called "logon content," and Pollack and Klemmer (10) have called "coordinality"]. This hypothesis correctly predicts the effect of Number of Turns upon judged complexity, since the number of coordinates per turn is a constant for either angular or curved shapes, considered separately. But whereas only two numbers are necessary to describe each turn in an angular shape (i.e., the x and y coordinates of the turning point), an additional number (i.e., radius of curvature) is necessary at each turn in a curved shape. Thus, curved shapes not only contain more information, but also require more dimensions for their specification; nevertheless they appear no more complex than angular ones. Another hypothesis which would predict higher values for the curved shapes is that complexity is a function of the total number of individually homogeneous parts, e.g., lines of constant slope, or curves of constant curvature, in the contour [cf. the measure of complexity used by Fehrer (6)].

The finding that symmetrical figures are intermediate between asymmetrical figures with the same number of independent turns and those with the same total number of turns is consistent with results previously obtained in a group of learning studies (4) in which dot patterns were used as stimuli.

The Angular Variability measure, no less than Symmetry, has to do with repetitiveness or similarity between parts of a contour; e.g., to state the limiting case, any polygon which is regular in the sense that all angles are equal will have a value of zero on this variable.

The Number of Turns variable, which accounted for nearly four-fifths of the variance of the judgments, may be described as the number of maxima (regardless of sign) in one cycle of the function relating curvature to distance along the contour. This function is a series of spikes for any angular shape, and a step-function for any curved shape

constructed by the present method (5). In the case of the latter function, information is concentrated in the vertical transition lines; i.e., in points on the contour at which arcs are tangent to straight lines. This is true from the point of view of a hypothetical observer who is in some sense capable of extrapolating curvature. From the point of view of an observer incapable of extrapolating curvature, however, but capable of extrapolating slope, information is associated simply with the maxima of the curvature function; i.e., information, or uncertainty, is concentrated in the arcs of the contour, but uniformly distributed within any given arc (though by some "averaging" process the observer might associate the information more specifically with the midpoint of the arc).

In an earlier study (2), Ss tended to choose maxima of curvature when asked to select the points most representative of a given contour. It may be that an observer's primary impression of complexity depends upon the number of such maxima in a shape and that this first estimate is then reduced more or less by the presence of certain types of redundancy or repetitiveness to which he is sensitive. The possibility of devising a single physical measure which would predict apparent complexity as efficiently as the combination of predictor variables which we have considered is not to be dismissed, but appears unlikely.

SUMMARY

Judgments of the complexity of 72 shapes were obtained from 168 Ss. The shapes were constructed by a method in which certain physical characteristics were systematically varied and the remainder randomly determined. About 90% of the variance of ratings was explained by (a) the number of independent turns (angles or curves) in the contour, (b) symmetry (symmetrical shapes were judged more complex than asymmetrical with number of *independent* turns constant, but less complex with *total* number of turns constant), and (c) the arithmetic mean of algebraic differences, in degrees, between successive turns in the contour. Angular and curved shapes were judged about equally complex, though the latter involved additional degrees of freedom (radii of curvature). Also

immaterial, within broad limits, was the grain of the matrix from which critical points were chosen to construct the shapes.

REFERENCES

1. ATTNEAVE, F. A method of graded dichotomies for the scaling of judgments. *Psychol. Rev.*, 1949, **56**, 334–340.
2. ATTNEAVE, F. *The relative importance of the parts of a contour*. San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, November 1951. (*Res. Note* P&MS 51-8.)
3. ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–193. (Distributed separately as *Res. Rep.* AFPTRC-TR-54-33, August 1954).
4. ATTNEAVE, F. Symmetry, information, and memory for patterns. *Amer. J. Psychol.*, 1955, **68**, 209–222. (Distributed separately as *Res. Rep.* AFPTRC-TN-55-29, October 1955).
5. ATTNEAVE, F., & ARNOULT, M. D. The quantitative study of shape and pattern perception. *Psychol. Bull.*, 1956, **53**, 452–471.
6. FEHRER, E. V. An investigation of the learning of visually perceived forms. *Amer. J. Psychol.*, 1935, **47**, 187–221.
7. FRENCH, R. S. Identification of dot patterns from memory as a function of complexity. *J. exp. Psychol.*, 1954, **47**, 22–26.
8. GUILFORD, J. P. *Psychometric methods.* New York: McGraw-Hill, 1954.
9. MACKAY, D. M. The nomenclature of information theory. In H. von Foerster (Ed.), (*Transactions of the eighth conference on*) *Cybernetics; circular causal and feedback mechanisms in biological and social systems.* New York: Josiah Macy, Jr. Foundation, 1952. Pp. 222–235.
10. POLLACK, I., & KLEMMER, E. T. *The assimilation of visual information from linear dot patterns.* Cambridge, Mass.: Air Force Cambridge Research Center, July 1954. *Tech. Rep.* (AFCRC-TR-54-16.)
11. WEINSTEIN, M. Stimulus complexity and the recognition of visual patterns. Unpublished doctor's dissertation, The Ohio State Univer., 1955.