The physical factor structure of random forms and their judged complexity¹

HERBERT H. STENSON

BEHAVIOR RESEARCH LABORATORY, ANTIOCH COLLEGE

This study relates the perceived complexity of 20 random forms to their physical factor structure. Ten principal axes, accounting for 94 per cent of the total variance of 24 physical measures, were rotated using the Varimax criterion. Factor scores for each form were correlated with the complexity ratings of the forms by each of 11 Ss. A single factor accounted for most of the variance in the complexity ratings. This factor was best described by four physical measures: the number of turns in the form, the length of the perimeter, the perimeter squared to area ratio, and the variance of the internal angles of the form.

A persistent problem in the study of the psychophysics of form has been the lack of a complete and reliable measure of the physical stimulus. The various approaches taken to resolve this problem have been reviewed by Michels and Zusne (1965) and by Hake (1957). Attneave (1957), Arnoult (1960), and Vanderplas and Garvin (1959) have successfully related psychological variables to certain isolated physical measures made on "nonsense" forms, and Vanderplas, Sanderson, and Vanderplas (1965) have provided values for a list of physical variables describing certain characteristics of a large number of such forms. But the problem of the measurement of form in general is still unresolved.

In this study, factor analysis was used in an attempt to define an orthogonal set of physical variables that would describe a set of forms in detail as a basis for psychophysical studies. A single psychological variable, judged complexity, was then related to the observed physical factor structure of the forms.

Attneave (1957) and Arnoult (1960), using multiple regression techniques, found that the complexity ratings of forms were related to the number of turns in the form, the ratio of the perimeter squared to the area, and a measure of the variability of the angles contained in the perimeter of a form. Using group complexity data they found that these variables accounted for about 80 percent of the variance of the complexity ratings. Because of this result these three physical variables were included in the factor analysis done in the present study.

METHOD

Subjects

The Ss were 11 undergraduates at Antioch College who were paid to participate in the experiment.

Stimuli

Twenty outline forms were constructed using Method 4 of Attneave and Arnoult (1956). A random number

between 5 and 20 was chosen for each form to determine the number of random points that would be plotted on a 100 mm x 100 mm matrix to initiate the construction of that form. Once a form was at the stage of construction given by Method 1 of Attneave and Arnoult (1956), that is, being entirely composed of straight lines, another random number was used to determine the number of angles to be replaced by arcs according to Method 4. This number was allowed to vary between zero and N, where N was the total number of points plotted to construct the form. Thus, it was possible that none or all of the angles would be rounded. The resulting forms are shown in Fig. 1.

Physical Measures of Stimuli

Twenty-four physical measurements were made on each stimulus form. They are defined as follows:

1. Turns. The number of points plotted to construct the form defines the number of turns in the perimeter of the form.

Fig. 1. Twenty random forms generated by Method 4 of Attneave and Arnoult (1956).

2. Proportion Arcs. The ratio of the number of arcs to the number of turns.
3. Proportion Convex Arcs. The ratio of the number of arcs whose radius is interior to the form to the total number of arcs.
4. Proportion Convex Angles. The ratio of the number of angles whose measurement interior to the form is less than 180° to the total number of angles.
5. Perimeter. The length of the perimeter (in cm) as measured with a map-distance measurer.
6. Area. The area encompassed by the perimeter as measured (in sq. cm) with a polar planimeter.
7. P^2/A . The ratio of the squared length of perimeter to the area. This measure may be considered a measure of "compactness." It is at its minimum when the form is a circle.
8. Proportion Straight Perimeter. The ratio of the sum of the lengths of straight segments in the perimeter to the total length of the perimeter.
9. Orientation. A straight line (d_1) was drawn connecting the two points on the perimeter having the greatest distance between them. The angle formed by d_1 and the abscissa of the graph paper on which the form was constructed was defined as the orientation of the form.
10. Elongation Index. A straight line (d_2) was drawn to (a) be perpendicular to d_1 , and (b) connect the two points on the perimeter most distant from each other along a perpendicular to d_1 . Elongation was defined as $\pm(d_1^2 - d_2^2)/(d_1^2 + d_2^2)$. The sign was made positive if the orientation was less than 90° and negative otherwise. This measure is modeled after the method of calculating a correlation coefficient from the squares of the axes of an equal-probability ellipse. If the form being measured were an ellipse the Elongation Index would be identical with the correlation coefficient implied by the ellipse. As the shape of the ellipse approaches a circle the absolute value of the index approaches zero, and as the shape approaches a line the absolute value of the index approaches unity.
11. Areal Distribution Index. Each form was divided into four quadrants by first defining the center of gravity of the form as the origin. Then an axis (X') was drawn through this origin parallel to the abscissa of the graph paper on which the form was constructed, and another axis (Y') was drawn perpendicular to X' and passing through the origin. If a_i is the area of the form contained in quadrant i then the Areal Distribution Index =

$$[(a_1 + a_3)^2 - (a_2 + a_4)^2] / [(a_1 + a_3)^2 + (a_2 + a_4)^2].$$

This measure is similar to the Elongation Index, but deals with the distribution of area rather than the distribution of the perimeter as the Elongation Index does.

12. Variance Index. Each form was treated as a discrete bivariate frequency distribution on X and Y , the

abscissa and ordinate of the graph paper on which the form was constructed. Each 1 mm division on X and on Y was treated as a class interval. The frequency function was taken to be unity for each X, Y class within the form and zero for each X, Y class outside the form. The variances of this distribution along the X axis (S_x^2) and along the Y axis (S_y^2) were then calculated. The Variance Index = $(S_y^2 - S_x^2)/(S_y^2 + S_x^2)$. This is a measure of elongation in a strictly vertical or horizontal direction as opposed to the Elongation and Areal Distribution Indices which measured elongation in non-vertical and non-horizontal directions.

13. Mean Straight Length. The mean of the distribution of the lengths of all straight segments of the perimeter.

14. Variance of Straight Lengths. The variance of the distribution of the lengths of all straight segments of the perimeter.

15. Skewness of Straight Lengths. The third central moment of the distribution of the lengths of all straight segments of the perimeter.

16. Mean Arc Length. The mean of the distribution of the lengths of all arcs on the perimeter.

17. Variance of Arc Lengths. The variance of the distribution of the lengths of all arcs on the perimeter.

18. Skewness of Arc Lengths. The third central moment of the distribution of the lengths of all arcs on the perimeter.

19. Variance of Internal Angles. The variance of the distribution of all internal angles (in degrees) including those angles that were replaced by arcs.

20. Skewness of Internal Angles. The third central moment of the distribution of all internal angles (in degrees) including those angles that were replaced by arcs.

21. Variance on X . The variance along the X axis as defined earlier under "Variance Index."

22. Skewness on X . Each form was treated as a bivariate distribution on X and Y as discussed earlier under "Variance Index." The mean (M_x) and the median (MD_x) were found on X , and Skewness on X was defined as $(M_x - MD_x) / S_x$, where S_x is the standard deviation along X .

23. Variance on Y . Defined in the same manner as "Variance on X ."

24. Skewness on Y . Defined in the same manner as "Skewness on X ."

Procedure

Physical Analysis. The matrix of correlations between physical measures on the forms was subjected to a principal axes factor analysis with unity in the diagonal of the correlation matrix. Rotations of the retained axes were performed using the Varimax criterion and factor scores were computed for each form on each rotated axis. The method of obtaining factor scores was the solution of the matrix equation, $S = ZA(A'A)^{-1}$,

where S is the factor score matrix, Z is the standardized data matrix, and A is the factor loading matrix resulting from the Varimax rotation.

Complexity Ratings. Each S rated each form five times on a complexity scale, and the average of his five ratings was used as his complexity rating for a form. These ratings were made after the S had already used these forms in a similarity-rating task in five previous sessions that were part of another experiment. Thus, the stimulus context was already well-known at the beginning of the session on complexity.

Each S was first instructed to look at an array of all of the forms and to make a careful selection of the form that he would choose as the most complex and the form that he would choose as the least complex. These forms were then placed on either side of a sheet of paper on which were printed twenty 100 mm lines with the words "Least Complex" and "Most Complex" printed at either end of the lines. Each S was told to consider a line to be a ruler measuring complexity with the ends of the line defined by the degree of complexity in the forms that he chose as least and most complex. Each of the remaining forms was then presented to the S in a random order. He was instructed to make a check mark on the scale line corresponding to the form he was presented with, such that the location of the check indicated the degree of complexity for that form. The deck of cards containing

the forms was reshuffled and each card was presented again until the deck had been gone through five times. Each time that a form was presented a new scale line was used for the rating. Thus, the S could not easily compare his present rating with previous ratings of the same form. The complexity measure for a single form was defined as the mean distance of the five checkmarks for that form from the "least complex" end of the scale.

RESULTS

Physical Analysis

Table 1 shows the product-moment intercorrelations of the physical variables measured on the forms. It is this matrix that was factor analyzed.

The principal axes analysis yielded 19 factors since there were 24 tests but only 20 forms. The first 10 of these factors accounted for 94 percent of the total variance, and these 10 were retained for rotation. Table 2 is the factor loading matrix resulting from the Varimax rotation of these factors. The percentage of the total variance accounted for by each factor is also shown. Factors D and H will be the factors of major interest when discussing the physical predictors of rated complexity. It should be noted that Turns, Perimeter, P^2/A , and Variance of Internal Angles specify factor D, while Proportion Convex Angles alone specifies factor H.

Table 1. Coefficients of Correlation between All Pairs of 24 Physical Variables Measured on 20 Forms
(In Units of 0.01)

Variable ^a	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.
1.	100	02	01	15	58	15	47	13	-03	24	-13	27	-23	-24	19	-10	-23	-31	38	-06	-09	01	33	-06
2.		100	04	-40	-50	-39	-44	-88	37	-28	17	-18	-79	-42	31	18	26	40	-15	-14	11	29	-36	23
3.			100	-19	-02	-10	-05	-05	39	-07	-28	-16	-19	09	-17	14	09	13	-06	36	-03	-18	-18	14
4.				100	33	21	28	55	-04	01	-17	11	33	40	16	-04	-19	-36	45	18	09	-22	14	11
5.					100	44	87	67	-09	16	-25	09	54	32	08	-19	-15	-27	58	05	19	-10	44	-07
6.						100	-02	43	11	-33	-44	22	46	28	24	-49	-40	-31	-08	-06	06	-22	66	12
7.							100	58	-28	42	03	-01	49	28	-04	03	05	-18	65	03	17	-03	17	-14
8.								100	-29	17	-16	15	80	54	-10	-31	-31	-43	31	16	-03	-19	38	-06
9.									100	-75	-34	-37	-44	-34	21	38	36	49	20	05	18	23	-37	01
10.										100	29	25	15	06	-31	02	-07	-34	04	15	-21	-15	15	-10
11.											100	00	-03	11	29	13	34	22	-31	-31	-22	10	-26	-51
12.												100	10	-04	-04	-47	-37	-52	-24	07	-81	-47	76	04
13.													100	69	-13	-42	-24	-35	13	05	09	-22	40	-02
14.														100	29	-29	-14	-20	-07	12	24	-37	26	10
15.															100	12	13	05	-14	-49	10	-20	08	-02
16.																100	76	57	27	-06	18	01	-67	-25
17.																	100	84	21	-01	07	01	-60	-17
18.																		100	03	02	23	24	-65	-11
19.																			100	25	35	23	-17	06
20.																				100	-11	01	00	15
21.																					100	42	-32	09
22.																						100	-34	-31
23.																							100	05
24.																								100

^a The names of the variables corresponding to the numerals are given in the text and also in Table 2.

Table 2. Factor Loadings of 24 Physical Variables on 10 Rotated, Orthogonal Factors (In Units of 0.01)

Physical Variable	Factor									
	A	B	C	D	E	F	G	H	I	J
1. Turns	.28	-.28	-.03	.83	.03	.18	.09	-.10	-.11	-.14
2. % Arcs	.90	.13	.12	-.10	-.23	-.14	-.14	.18	.09	.01
3. % Convex Arcs	.05	.09	.10	.00	-.01	-.03	-.08	.12	-.92	.24
4. % Convex Angles	-.31	-.09	.07	.18	-.16	.02	-.10	-.86	.13	.11
5. Perimeter (P)	-.49	-.13	.06	.85	-.09	-.03	-.02	.01	-.01	-.01
6. Area (A)	-.46	-.45	.59	.14	-.09	.15	-.16	.09	.05	-.18
7. P ² /A	-.42	.11	-.32	.80	-.07	-.11	.04	-.03	.03	.04
8. % Straight P	-.82	-.22	-.02	.30	-.02	.04	.07	-.28	-.02	.10
9. Orientation	.30	.32	.81	.04	.10	-.15	.04	-.08	-.26	.05
10. Elongation Index	-.10	-.05	-.88	.22	.10	.17	.05	.01	-.05	.04
11. Areal Distrib. Index	.12	.27	-.43	-.18	-.48	.09	.57	.11	.25	-.11
12. Variance (Var.) Index	-.02	-.38	-.10	.07	.00	.88	-.04	-.01	.14	.06
13. Mean Straight Length	-.93	-.18	-.10	.06	-.15	-.05	-.05	.04	.19	.07
14. Var. Straight Length	-.62	-.14	-.11	-.07	-.69	-.14	-.14	-.12	-.14	.09
15. Skew. Straight Length	.18	.06	.29	.12	-.72	.05	.03	-.17	.05	-.52
16. Mean Arc Length	.19	.83	-.04	.04	.17	-.14	.11	-.23	-.20	-.22
17. Var. Arc Length	.10	.96	.05	.03	-.11	-.01	.05	.13	.04	.04
18. Skew. Arc Length	.19	.78	.27	-.13	-.09	-.21	.06	.32	.04	.13
19. Var. Internal Angles	-.09	.26	.05	.70	.22	-.28	-.12	-.37	.16	.26
20. Skew. Internal Angles	-.08	.01	-.01	.03	.10	.09	-.12	-.13	-.24	.88
21. Var. on X	-.07	.05	.12	.14	-.09	-.91	-.16	-.01	.03	-.10
22. Skew. on X	.31	-.09	.12	.07	.23	-.61	.46	.15	.31	.25
23. Var. on Y	-.30	-.65	.08	.25	-.11	.51	-.11	.14	.11	-.06
24. Skew. on Y	.13	-.14	.02	-.06	-.05	-.04	-.93	-.05	-.02	.11
% of Total Variance	.17	.15	.10	.12	.06	.11	.07	.05	.05	.06

The factor analysis was based on only the linear components of the relationships between variables. To check for any strong non-linear relationships, scatterplots of most of the possible pairs of physical variables were examined. No strong non-linear relationships were found.

Table 3 shows the factor scores of each form on each of the rotated factor axes. The product-moment correlations of the factor scores on each factor with the complexity ratings for each of the 20 forms were computed for each S. Table 4 shows the correlations between rated complexity and each factor for each S, and between the mean rated complexity (averaging over Ss) and each factor. The .05 significance level for correlations based on 20 cases is $\pm .43$. If one considers each correlation individually, then only the correlations with factor D are significant, with a single exception; the complexity ratings made by S number 6 show a correlation of $-.51$ with Factor H. However, a more appropriate statistical evaluation of these data must take into account the fact that the correlations of complexity with Factor H were negative for all 11 Ss. Assuming the hypothesis that the true correlations of complexity with all factors but Factor D is zero, then the probability of obtaining 11 coefficients all of the same sign for at least one of the remaining nine factors is .009. Thus, Factor H may be considered a statistically significant predictor of complexity. But

Factor D is clearly the important determinant of complexity, with correlations ranging from .75 to .89, while those between complexity and Factor H ranged from $-.09$ to $-.51$, as shown in Table 4. The correlation of the group average ratings with Factors D and H are .90 and $-.25$, respectively, indicating that Factor D accounts for 81 percent of the variance in complexity ratings while Factor H accounts for only an additional 6 percent.

The correlations of each S's complexity ratings with the group average complexity ratings ranged from .80 to .97 indicating that the group data is a good representation of each individual's data.

DISCUSSION

Physical Analysis

The factor loadings of physical variables on rotated factors shown in Table 2 exhibit certain features that lead to obvious interpretations for most factors, while for a few factors the meaning is more elusive. The nature of the factors is discussed below, after which some methodological points concerning this analysis are discussed.

Since Factors D and H are the important variables for the psychophysical analysis to be discussed later, they will be examined in greater detail than the remaining factors.

Factor D might be called "physical complexity" because of its strong relationship to judged complexity and also because of the nature of the physical measures that define it. Turns, Perimeter, P^2/A , and Variance of Internal Angles, are the defining variables, and all four have large positive loadings on Factor D, as shown in Table 2.

Table 3.
Factor Scores of 20 Forms on 10 Rotated, Orthogonal Factors

Form Number ^a	Factor									
	A	B	C	D	E	F	G	H	I	J
1.	-.81	.62	-2.46	-.87	-1.65	-1.35	.90	-.20	-1.43	-.14
2.	-.48	2.03	-1.35	1.32	.51	1.27	-1.36	.32	1.07	-.52
3.	-.94	.01	-.50	.21	.83	.37	.15	.09	1.64	-.45
4.	-.41	2.42	1.12	.16	-.28	.63	1.66	-.23	-.26	.29
5.	.71	1.29	.86	-.79	-.37	-1.40	.19	2.56	.09	.32
6.	-.81	-.50	-.85	.74	-.24	.62	-1.87	.17	-1.80	-1.02
7.	.48	-.64	1.34	1.48	-2.66	.14	-.11	.17	-.19	-.35
8.	2.08	-.87	-1.22	-1.17	.49	.75	.25	1.88	-.06	-.99
9.	.77	-.06	.46	.12	.22	1.10	-.87	-.23	.12	.46
10.	1.63	-.85	-.83	.16	-.07	.46	.73	-1.09	.31	.15
11.	.97	1.19	.57	-1.09	.50	.89	-.39	-.78	-.18	.67
12.	-1.33	-.71	1.02	1.17	.34	.22	.81	.94	-.95	-.46
13.	.57	-.51	.24	-.46	-1.51	.84	.28	-1.51	.56	.33
14.	-.96	-.86	-.67	-.01	.46	.24	-.29	.47	-.45	3.78
15.	-.21	-1.04	1.29	-.08	1.22	.36	-.03	.54	-.95	-.44
16.	.85	-.45	.27	.94	.66	-1.86	-.72	-.09	.75	.58
17.	.29	.30	.25	.37	.17	-2.47	-1.45	-1.04	.60	-.32
18.	-.34	-.76	-.75	1.29	.83	-.43	2.39	-.26	1.10	-.46
19.	-.01	.29	.59	-1.01	1.65	-.49	.27	-1.73	-1.74	-.76
20.	-2.05	-.92	.65	-2.48	-.67	.10	-.55	.03	1.77	-.69

^a The forms corresponding to these numerals are shown in Fig. 1.

Table 4. Intercorrelations for 11 Ss between Complexity Ratings and Factor Scores of 20 Forms on 10 Rotated, Orthogonal Factors

Subject Number	Factor									
	A	B	C	D	E	F	G	H	I	J
1.	-.13	-.22	.07	.84	.10	-.11	.03	-.22	.00	.10
2.	-.10	-.30	-.04	.89	-.07	-.05	-.06	-.11	.06	.15
3.	-.23	-.02	-.15	.81	.10	-.23	-.29	-.14	.17	.10
4.	.21	-.14	.11	.81	-.09	-.05	-.18	-.20	.03	.09
5.	.02	-.16	.19	.84	.04	-.10	.15	-.26	-.17	.21
6.	.16	.05	.17	.61	.02	.03	-.03	-.51	-.01	-.08
7.	-.26	-.28	.13	.78	.11	-.05	.16	-.20	-.07	.21
8.	-.25	.03	.15	.87	-.08	.02	.04	-.09	-.01	.19
9.	.17	-.23	-.03	.83	.27	-.08	.08	-.16	-.02	.24
10.	-.09	-.11	.04	.89	-.04	.05	.02	-.21	.08	.10
11.	.03	.05	.02	.74	.17	.04	-.10	-.41	.04	.07
Group Mean	-.08	-.13	.06	.90	.05	-.05	-.02	-.25	.01	.14

Certain necessary relationships exist among these variables. If a set of forms is constructed on a standard size matrix, then the more turns a form has in its perimeter, the longer that perimeter must be. Thus, Turns and Perimeter are necessarily related for the set of forms used in this study. It is also necessary that a long perimeter and a large number of turns result in a smaller than average enclosed area, unless a very special set of random points is chosen in the construction of the form. Thus, P^2/A is related to Turns and Perimeter. The relationship of Variance of Internal Angles to these other three variables is not necessarily inherent in the type of forms used here, and, indeed, its factor loading on Factor D is not quite as large as those for the other defining variables. The empirical fact that it is related to Factor D supports the interpretation of this factor as "physical complexity," since this variable measures the lack of homogeneity of the internal angles in the forms.

Factor H is defined by a single physical variable, Proportion Convex Angles. Its loading on Factor H is negative, indicating that this factor is a measure of the proportion of *concave* angles in the forms, i.e., the angles that make "indentations" in the perimeter of the forms. It should be noted that this proportion has nothing to do with the magnitude of the indentations. The latter help to determine Factor D. For example, Form 18 in Fig. 1, though nearly cut in half by a deep concave angle, has a small factor score of $-.26$ on Factor H, as shown in Table 3, since only 7 of its 19 angles are concave. Its high factor score of 1.29 on Factor D was primarily due to the greater than average values of P^2/A and Variance of Internal Angles that are associated with the large angular indentations in this form.

The purpose of performing this factor analysis on the physical measures taken on the forms was to obtain an orthogonal set of physical variables that could be related to psychological judgments. Therefore, the "meanings" of the factors that are not related to the

judgment of complexity are not of great interest in the present study. The relationship of these factors to psychological variables in future studies may provide names for them. There are, however, some obvious features that should be pointed out in some of the remaining factors.

Factors A and B appear to be related to independent aspects of curvature in the perimeter of the forms. Factor A is specified by the large positive loading of Proportion (%) Arcs and by the large negative loadings of Proportion (%) Straight Perimeter and Mean Straight Length. Factor B is specified by high positive loadings of Mean Arc Length, Variance Arc Lengths, and Skewness of Arc Lengths. Thus, Factor A seems to be a measure of the relative amount of curvature in a form as a whole, while Factor B describes the details of the curvature when it appears in a form.

Factors G and F combine to describe the nature of a form's orientation in space. Factor G is specified by the large loadings of Elongation Index and Orientation. It, thus, describes and specifies the direction of the general elongation in a form. The large positive loadings of the Variance Index and Variance on Y, and the negative relationship of Variance on X to Factor F indicate that this factor describes the amount of elongation in a strictly vertical or strictly horizontal direction.

The Areal Distribution Index, Skewness on X, and Skewness on Y, all have high loadings on Factor G, suggesting that this factor describes the degree to which the area within the form is unequally distributed. It will be recalled that the quadrants of a form were defined by vertical and horizontal axes passing through the center of gravity of the form. The quadrants were numbered in a counterclockwise direction beginning in the "Northeast" quadrant. The Areal Distribution Index compares the area in quadrants 1 and 3 with the area in quadrants 2 and 4. Skewness on X in effect compares the areas on either side of the X mean with each other, and Skewness on Y in effect compares the areas on either side of the Y mean with each other.

Factor I, like Factor H, is uniquely described by a single physical measure: Proportion Convex Arcs has a loading of $-.92$ on Factor I.

The methodological issues to be discussed with regard to this factor analysis relate to the decision to factor the entire variance of the physical measures rather than the common variance.

The correlation matrix that was factored (Table 1) had unity in the principal diagonal. Depending on one's view this implies either that all the variance was common variance, or that it was appropriate to include both common and unique variance components in the analysis. The assumption was made that for judgment tasks of the type used in this experiment, the Ss could not be expected to restrict themselves to observations of only those stimulus attributes that were physically correlated with other stimulus attributes. That is, Ss could not be

expected to make judgments based on common variance alone. Therefore, the entire variance was analyzed. It is interesting to observe that Factors H, I, and possibly J would not have appeared in an analysis that used communalities in the principal diagonal, because each of these factors is uniquely specified by a single physical variable related to no other factors.

The importance of specifying these unique factors in the physical measures is illustrated by the fact that in the present experiment the unique factor H was involved in the S's complexity ratings. Thus, to appreciate fully the physical aspect of the psychophysics of judged complexity, it was necessary to allow Factor H to appear, by using unity rather than communalities in the principal diagonal of the correlation matrix.

Psychophysical Analysis

The results presented in Table 4 show that Factor D is clearly the most important, and perhaps the only meaningful physical correlate of complexity. Factor H, while a statistically significant correlate of complexity, is unimportant in terms of the variance it accounts for in the complexity ratings, and will be discussed later. The variables that have high loadings on Factor D, with the exception of Perimeter, are the same variables found to be related to complexity in the studies cited earlier by Attneave (1957) and Arnoult (1960). Both of these investigators used multiple regression techniques to relate complexity to the physical variables. The following section is a discussion of multiple regression in terms of factor analytic concepts. The relationship between the results of the present study and those that would have been obtained using a standard multiple regression analysis is discussed and related to the results obtained by Attneave and Arnoult. Following this some conclusions regarding the psychological implications of the methods and results are presented.

Multiple Regression and Factor Structure. In multiple regression analysis, a lack of independence between the predictor variables produces effects that make interpretation of the prediction system difficult. The multiple R that is obtained in such a case is composed of components due to the relationship of the criterion variable to the common, or "shared," variance in the predictors, and components that are due to the relationship of the criterion variable to unique, or independent, variance in the predictors. These components can not be separately identified without further analysis. If, however, the predictor variables are independent, then the multiple R is composed only of the independent relationships of each predictor to the criterion. Thus, independence between predictors is a desirable condition for multiple regression analysis.

If one obtains an independent set of predictors from an orthogonal factor analysis of a set of dependent predictors first, and then relates the criterion to these independent predictors, many problems of interpreta-

tion are circumvented. The squares of the correlations between the criterion variable and each factor then show the proportion of the criterion variance accounted for by each factor, and multiple R^2 is simply the sum of these independent contributions. The correlation of each factor with the criterion will, of course, change if the factor axes are rotated to a different position. This is as it should be since the rotated factors will have a different relationship than the original factors to the variables that were factor analyzed, and therefore will have a different "meaning." The multiple R, however, will remain constant when the factor axes are rotated.

Another aspect of the problem of dependencies between predictors in multiple regression is the ambiguity of the regression weights when these dependencies exist. An example using the present results will demonstrate this point. Table 1 gives the intercorrelations between Turns (T), Variance of Internal Angles (V), and P^2/A (X). These are the same variables found to be related to complexity judgments in the Attneave and Arnoult studies cited. In the present study the correlation of group average complexity ratings (C) with T was .80, with V was .77, and with X was .71. These data may be used to compute the following multiple regression equation for predicting C from T, V and X, where all variables are standardized:

$$C' = .55T + .15V + .46X$$

The use of this predicting equation accounts for 91 percent of the variance in C (multiple $R^2 = .91$). Notice that T receives about four times as much weight as V in this equation. Now consider the multiple regression equation resulting from the present data for predicting C from T and V alone:

$$C' = .59T + .54V$$

The use of this equation accounts for 89 percent of the variance in C. But notice now that T and V receive about equal weight, and that the omission of X as a predictor has not substantially reduced the amount of variance accounted for in C. The discrepancy in the relative weights of T and V between these equations is due to the effects of the linear dependencies between the predictors; particularly the strong dependency between V and X. The factor analysis done on the present data shows (see Table 2) that the interdependencies between these three variables were of sufficient magnitude to define a single, simple-structure factor. The finding that this factor was the only predictor that was related to complexity may be taken to mean that it is the "shared," or common, variance of T, V and X with respect to Factor D that is the physical correlate of judged complexity, and that the unique aspects of each of these variables with respect to Factor D were not

related to judged complexity.

It was pointed out above that the multiple R between a set of independent predictors and a criterion variable is invariant under orthogonal rotation of the predictors, but that the relationship of each predictor to the criterion will change as a result of such rotations. One such orthogonal rotation is that from the principal axes to the simple structure axes as used in this experiment. If the correlations shown in Table 4 between each simple-structure factor and the group complexity are squared and summed over factors the result is a good approximation to the multiple R^2 that would have been obtained if all the 24 physical variables measured were used as predictors of group complexity ratings in a standard multiple regression analysis of this data. The multiple R^2 so obtained is .92. The multiple regression equation given earlier in which only T , V and X were used as predictors results in a multiple R^2 of .91. Thus, only an additional one percent of the variance in complexity is accounted for by including all of the physical variables as predictors. But the factor analysis also shows that the variance in T , V and X that is not common to those three variables is not significantly related to complexity. Thus, the amount of variance that is both significantly and meaningfully accounted for in the complexity data is 81 percent; that amount accounted for by Factor D. The relative weights that T , V and X receive in determining this factor are about equal as is shown in Table 2, whereas the regression equation using T , V and X as predictors gives V far less weight than the other two variables.

The point to be made from this comparison of standard multiple regression techniques with multiple regression using orthogonal factors as predictors is that, while both techniques will reveal the maximum amount of prediction possible, the use of factor analysis enables one to separate the total relationship of criterion to predictors into components. This separation contributes to the understanding of the prediction system, and enables one to eliminate components from the system when they are probably chance relationships.

The Attneave and Arnoult data on complexity as a function of T , V and X show about the same intercorrelations between T , V and X as those shown in Table 1. Their multiple R s describing the gross relationship were of the same general magnitude as the correlation of complexity with Factor D in the present study. The additional information contributed by the methodology of the present study consists of the findings that T , V , X and Perimeter are the only variables that are related to complexity when the correlations of complexity with a large set of variables are considered; that the basis of these relationships is a single factor that is described by T , V , X and Perimeter; and that T , V , X and Perimeter all contribute about equally to this factor.

Conclusions. The results presented in Table 4 in combination with the high correlations between individu-

al complexity ratings and the group mean ratings indicate a high degree of reliability, both between and within Ss, in the use of the rating scale; in the perception of the physical characteristics of the stimuli; and in the use of these characteristics to produce ratings.

There is, however, one indication of unreliability between Ss. For Ss 6 and 11 the correlation of complexity with Factor H was higher than for the remaining Ss, while the correlation of complexity with Factor D was lower for these two Ss than for the remaining Ss. Factor H, while generally unimportant, was a statistically significant variable. Thus, there may be some minor individual differences in the relative importance assigned by Ss to the physical variables when making complexity judgments.

Finally, we may consider whether or not the use of factor analysis has contributed to an understanding of the perception of complexity beyond just the mathematical relationship of Factor D to rated complexity. It is noteworthy that Factor D appeared as the result of a purely analytical rotation to simple structure, and precisely those variables that Attneave (1957) and Arnoult (1960) found to be related to judged complexity, using different forms and different Ss, were the variables defining this simple-structure factor. This result suggests that Factor D measures a statistically independent physical characteristic of closed figures that is used by Ss in making judgments of the complexity of the figures. It can be argued, of course, that any rotation of the factor axes would have accounted for just as much variance in the complexity judgments, and that therefore the conclusion is quite arbitrary that Ss are indeed perceiving Factor D when they are making their complexity judgments. However, if one assumes that perceptual judgments are made on a kind of "principle of least effort" then one would expect that, when making judgments about perceptual input, an observer would use the fewest dimensions possible, compatible with the task, and that these dimensions would be as uncomplicated as possible. That is to say, he would try to use a single variable of the simple-structure type. Indeed, the reason for the very existence of the simple-structure concept in factor analysis is to make it possible for the analyst to "perceive" the meaning of the factors.

Support for the implied hypothesis made here that the human is a "simple-structure" observer must come from further studies in which other psychological judgments about forms are related to the physical factor structure of the forms.

References

- Arnoult, M. D. Prediction of perceptual responses from structural characteristics of the stimulus. *Percept. mot. Skills*, 1960, 11, 261-268.
- Attneave, F. Physical determinants of the judged complexity of shapes. *J. exp. Psychol.*, 1957, 53, 221-227.
- Attneave, F., & Arnoult, M. D. The quantitative study of shape and pattern perception. *Psychol. Bull.*, 1956, 53, 452-471.

- Hake, H. W.** Contributions of psychology to the study of pattern vision. U. S. Air Force, Wright Air Development Technical Report, 1957, No. 57-621.
- Michels, K. M., & Zusne, L.** Metrics of visual form. *Psychol. Bull.*, 1965, 63, 74-86.
- Vanderplas, J. M., & Garvin, E. A.** The association value of random shapes. *J. exp. Psychol.*, 1959, 57, 147-154.
- Vanderplas, J. M., Sanderson, W. A., & Vanderplas, Janet N.** Statistical and associational characteristics of 1100 random shapes. *Percept. mot. Skills*, 1965, 21, 414.

Note

1. The research reported in this paper was sponsored by the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright Patterson Air Force Base, Ohio, under contract number AF 33 (615)-1086 with Antioch College, Yellow Springs, Ohio. This paper has been identified by the Aerospace Medical Research Laboratory as AMRL-TR-66-67. Further reproduction is authorized to satisfy needs of the United States Government.

(Accepted for publication July 26, 1966.)