

A feature utilization analysis of the perception of pattern class structure*

LEONA S. AIKEN, Temple University, Philadelphia, Pennsylvania 19122
and
DONALD R. BROWN, Purdue University, West Lafayette, Indiana 47907

Patterns were computer-generated about two prototypes to form schematic clusters of three diameters about the cluster centroids. In a schematic concept-formation task, recovery of a priori schematic class membership by Os varied inversely as physical cluster diameter, with concept acquisition across trials evidenced in only the low-cluster-diameter condition. For each O who failed to classify according to the schema rule, linear discriminant function analysis was applied to his classes. In all cases, O-generated classes were successfully recovered by the physical pattern features used as predictors, the mapping of these classes by the LDF exceeding that by the schema rule in accuracy at all cluster-diameter levels.

The concern of the present study, observer sensitivity to relational invariants from specific instances of stimulation, has been discussed in a variety of contexts and has been variously labeled from "pattern conception" by Miller and Chomsky (1958) to "schematic concept formation" by Evans and his associates (e.g., Evans, 1967).

Within our specific context, visual pattern perception, it has been previously noted that Os are sensitive to properties of sets of patterns as well as to the properties of the instances which define the sets (e.g., Brown & Owen, 1967; Mavrides & Brown, 1969). As Os are sensitive to the rules which define finite-state languages (Kolers, 1967), so are they sensitive to the statistical structure which serves as the rules by which sets of visual patterns are generated (e.g., Bersted, Brown, & Evans, 1969; Mavrides & Brown, 1969). Given, in addition, the physiological (e.g., Hubel & Wiesel, 1968) and psychophysical (Aiken & Brown, 1969a, b, c; Behrman & Brown, 1968; Brown & Andrews, 1968; Brown & Brumaghin, 1968; Brumaghin & Brown, 1968; Stenson, 1968; Sternberg, 1967) data, it seems reasonable to think that Os are processing individual patterns in terms of abstracted features and sets of patterns in terms of the distributions of such features in a multidimensional cue-dimension space (Beach, 1964a, b). Data in support of this general conceptualization have been provided for discrimination and reproduction tasks (Mavrides & Brown, 1969) and for a sorting task (Mavrides & Brown, 1970). The present study was designed to provide further data in support of these general notions. In particular, we were

interested in evaluating the accuracy with which classification behavior can be mapped into an a priori classification structure and in providing quantitative data to assess the extent to which classification can be predicted from pattern features.

METHOD

Patterns

Patterns were generated to form clusters in a coordinate space about two different *prototypes* or class centroids (see Fig. 1). Construction of patterns about these two prototypes began with the generation of three sets of bivariate normal data. Both variates of each distribution had an expected value of zero; the intercorrelation between the two variables was zero in all sets. The three sets corresponded to three variability levels, that is, standard deviations of 4, 7, and 10.

Describing the prototype in polar coordinates, patterns were generated by perturbing the lengths of the radii of the prototype patterns, and the angles between these radii, by the addition of the randomly generated values. This procedure was followed for both prototypes at each of the three variability levels. At each variability level and for each prototype, 100 patterns were generated. Seventy-five of each set were selected for use in the experiment, any patterns containing three-collinear vertices being discarded so that all patterns would be eight-sided. The 75 patterns of each set constituted a *schema class*.

Patterns were normalized to a constant area of 1,250 units, centered in a 100 by 100 unit field, and were plotted using a Hybrid computer (P. M. Aiken, 1969). Hybrid plotting was used to avoid the discontinuities associated with conventional digital incremental plotting. The patterns were then cut from black construction paper, mounted on white

cardboard squares, and were reproduced as 35-mm slides. A sample of patterns from each prototype at each variability level is given in Fig. 2.

Since the expected perturbation of any radial or angle of a prototype was zero, the mean pattern in a cluster, computed by finding the mean X and Y coordinates of corresponding vertices across the patterns, should have been equal to the prototype. The mean pattern for each cluster of 75, compared with the prototype from which it arose in Fig. 3, in all cases closely approximated its prototype.

Observers

Thirty-six undergraduates served as Os to satisfy a course requirement.

Experimental Design and Procedure

The 150 patterns at each variability level were divided into three blocks of 50 patterns each, which contained 25 patterns from each prototype. These blocks formed the basis of a 3 by 3 balanced Latin square design to control for order effects. Twelve Os were assigned to each variability level,

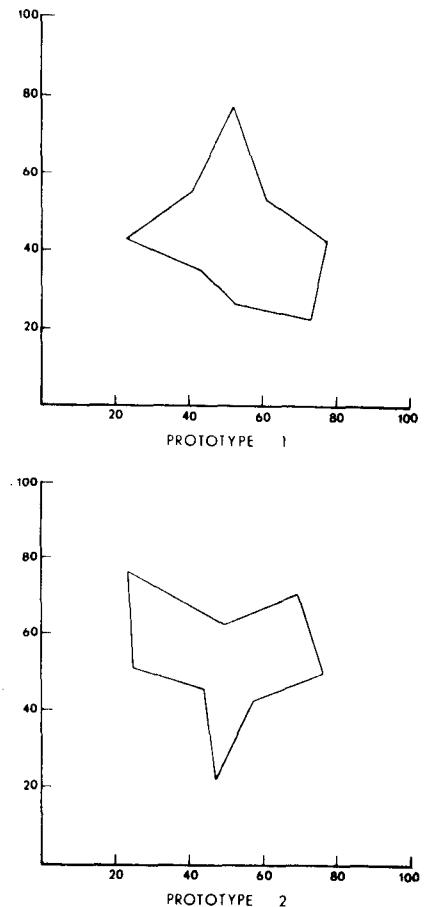


Fig. 1. Two prototypes used in the experiment.

*This research was supported by Research Grant HD-00909 from the National Institute of Child Health and Human Development.

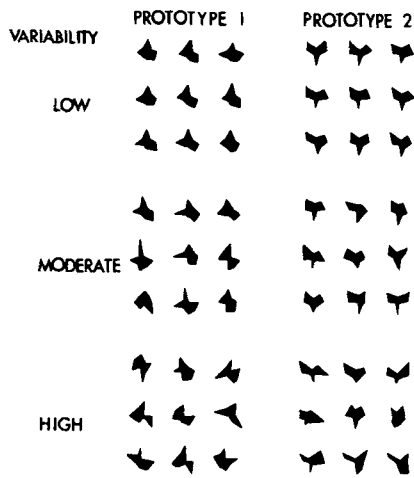


Fig. 2. Sample of patterns from each prototype at each variability level.

the 12 Os at each level being subdivided randomly into three groups ($N = 4$) corresponding to the rows of the Latin square.

Os were run individually in 1-h sessions. Each O viewed all 150 patterns within a variability level in a different random sequence. The patterns from both prototypes were intermixed randomly throughout the administration. Os were informed that the patterns they would view contained two classes, the nature of which they were to discover. Each O was asked to arbitrarily call the first pattern A or B, thereafter responding A or B to each stimulus in a fashion consistent with his

own idea of the nature of the classes. Exposure time for each pattern was 4 sec, and Os responded by depressing one of two response keys. No feedback was given.

RESULTS

Control Design

An analysis of variance of the Latin square control design at each variability level in terms of accuracy of classification according to the schema rule showed no order effects to be present ($F < 1$ in all cases). At the moderate and high variability levels, no difference in difficulty was found among the three blocks of 50 patterns. At the low variability level, one block of patterns proved slightly more difficult than the other two ($p \leq .05$).

Schematic Concept Formation

The manipulation of variation about the prototype had a marked effect upon classification accuracy according to the schema rule ($p \leq .001$). As seen in Fig. 4, at the low variability level, average accuracy of classification exceeded 90% in the third set of 50 patterns, while being markedly lower for both the moderate and high variability patterns. While all Os received 150 trials, in order to allow an opportunity for acquisition of the classification rule, only in the low variability condition did performance improve over time ($p \leq .05$).

Pattern Feature Analyses

For the purpose of feature analyses of

O-generated classes, all patterns were quantified by the 80 physical measures described by Brown and Owen (1967). This large set of measures is primarily comprised of indicators of compactness, jaggedness, elongation, and vertical and horizontal skew of the patterns.

To test the ability of physical pattern features to distinguish between E-generated classes, stepwise discriminant analyses were constructed based upon the a priori schematic classes. Predictors for the analyses were nine physical measures chosen to represent the factor structure of the 80 measures from which they arose. These measures, as well as the point-biserial correlation between each measure and class membership at all variability levels, are given in Table 1.

In the stepwise discriminant analyses of the a priori classes, recovery of these classes on the basis of two physical measures, CG_y and CG_x , was almost complete, there being 100%, 97%, and 86% recovery for the low, moderate, and high variability patterns, respectively. While other predictors were significant in terms of accounting for between-class variation with the inclusion of only CG_y and CG_x , class recovery was maximal at all variability levels (see Table 2). The location of a patterns in two-spaces defined by these two physical measures is shown in Fig. 5a-c.

Observer Discriminant Analyses

Os were considered as being schematic

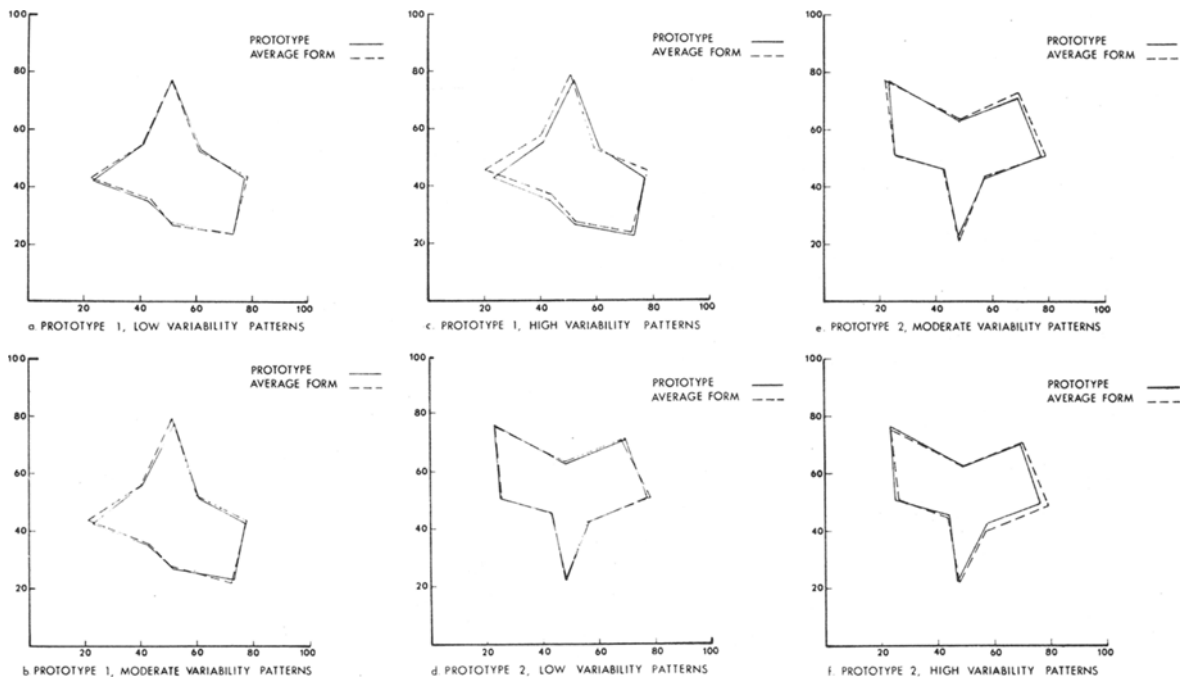


Fig. 3. Mean pattern compared with prototype at each variability level and for both prototypes.

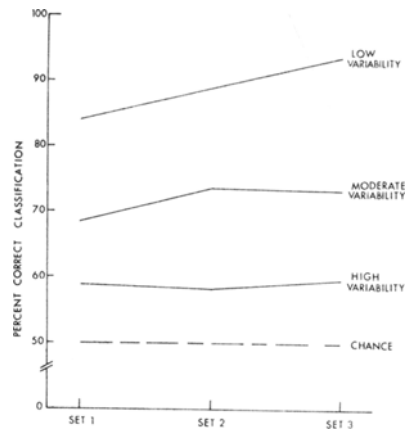


Fig. 4. Mean percent of stimuli correctly classified according to schema rule at each level of variability.

Table 1
Physical Measures* Used in Pattern Analyses

Measure	Point-Biserial Correlation with Class Membership (N = 150)		
	Low Variability	Moderate Variability	High Variability
P	-.10	-.13	-.03
PM ₂	.05	.01	.05
AE	.57	.38	.25
AJ	.38	.07	-.13
A ₃₁	-.70	-.43	-.07
V/H	-.02	.12	.10
S/P	.18	.20	.09
CG _x	.82	.55	.32
CG _y	-.96	-.85	-.66

*See Brown & Owen, 1967, for a more complete definition of pattern measures.

P—Number of units in the perimeter

PM₂—Maximum second perimeter moment of X

AE—Size in degrees of the smallest exterior angle

AJ—Size in degrees of the smallest interior angle

A₃₁—Third moment of interior angles

V/H—Ratio of vertical to horizontal extent

S/P—Ratio of largest side to perimeter

CG_x—X coordinate of areal center of gravity

CG_y—Y coordinate of areal center of gravity

nonschematic, based upon the nature of the classes each generated in his final 50 trials. An O was defined as schematic if he had achieved 90% accuracy of classification according to the a priori schema rule. By this criterion, 12 Os, 9 in the low variability and 3 in the moderate variability condition, qualified as schematic.

For the 12 schematic Os, as expected, individual LDFs constructed from their classes in the final 50 trials closely resembled the model LDF equations generated from the schema classes, i.e., equations contained primarily the measures CG_y and CG_x as significant predictors. For all schematic Os, one of these measures was the strongest predictor in the equation, and for 8 of the 12 schematic Os, these two measures were the best predictors, as had been the case in the model equation. As seen in Fig. 6, no distinction could be made in accuracy of mapping of O-generated classes by the LDF vs the a priori schematic rule.

For the 24 nonschematic Os, their groupings were successfully recovered by the LDF based on features. The mapping of O-generated classes was quite accurate in terms of the feature analysis, being at worst about 80% accurate in the moderate-variability condition (see Fig. 7). For all Os, significant predictors among the nine measures could be found which differentiated between their specific classes. In all cases significant differentiation between classes being achieved (see Table 3).

In contrast to the schematic Os, for nonschematic Os, individual LDFs based upon their classes were necessarily at wide variance with the model equations. While in 37% of these LDFs an areal skew measure did occur as the best or second best predictor, never did both the horizontal and vertical areal skew measures occur together, as had been the case in the model equation. For 25% of the Os, their

Table 2
Final Model Discriminant System at Each Level of Variability

Low Variability Patterns				
Predictor*	CG _y	CG _x	S/P	A ₃₁
p**	.001	.001	.001	.005
N***	150	150	150	150
Moderate Variability Patterns				
Predictor	CG _y	CG _x	A ₃₁	
p	.001	.005	.05	
N	143	145	144	
High Variability Patterns				
Predictor	CG _y	CG _x	AE	
p	.001	.025	.025	
N	120	129	124	

*Significant predictors in order of occurrence in progressive steps of LDF analysis.

**Level of significance of variable as a discriminator of class centroids in a system containing other variables.

***Number correctly classified at point of entry of variable (of 150).

classes were best distinguished by feature measures of angularity (A_J, A_E, A₃₁) alone. For another 17% of Os, their classes were best described by measures of compactness and elongation (P, PM₂, V/H).

DISCUSSION

When the cluster diameter of classes in a physical feature space is allowed to increase, as with the pattern generation procedure described herein, we find this manipulation reflected in a decrease in the ability of the O to map the a priori class structure inherent in the patterns. With weakly constrained pattern clusters, Os fail to exhibit knowledge of class structure, as reflected in their near-chance accuracy of pattern classification according to the a priori rules.

Given this failure to recover the class structure, a consequent question concerning the nature of the O's perceptual behavior arises: If the O fails to extract the structure imposed in pattern generation,

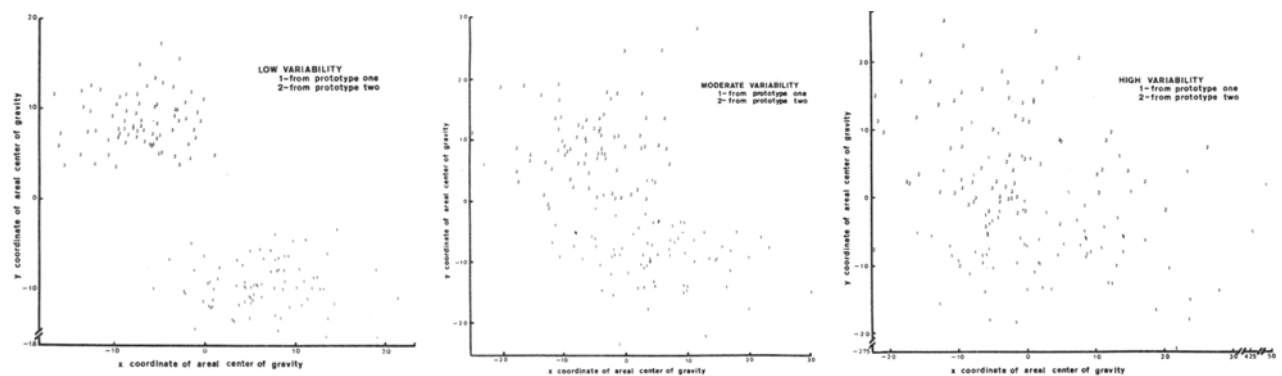


Fig. 5. Location of schematic patterns in a two-space defined by the X and Y coordinates of the areal center of gravity of each pattern for the low- (a), moderate- (b), and high- (c) variability patterns.

Table 3
Summary of Individual Linear Discriminant Systems for Nonschematic Ss

S	Number of Predictors ^a	Model Predictors ^b	NLDFc	NSCFd	p ^e
Low Variability Condition					
1	2	0	46	35	.001
2	3	1	45	44	.001
3	2	1	43	42	.001
			\bar{N} : 44.7	40.3	
Moderate Variability Condition					
1	4	0	44	30	.001
2	3	0	38	32	.01
3	1	1	41	40	.001
4	2	0	38	36	.001
5	2	1	36	35	.001
6	2	0	42	38	.001
7	2	0	37	25	.01
8	2	1	38	26	.005
9	2	1	37	39	.001
			\bar{N} : 39.0	33.0	
High Variability Condition					
1	1	0	33	26	.005
2	2	0	41	29	.001
3	2	1	39	31	.001
4	1	1	38	38	.001
5	6	0	44	33	.001
6	4	0	42	31	.001
7	3	0	41	26	.001
8	4	0	43	27	.005
9	3	0	38	27	.005
10	2	0	37	30	.005
11	3	0	43	31	.001
12	2	0	41	30	.001
			\bar{N} : 40.0	29.9	

^aNumber of significant predictors ($p \leq .05$) in the final LDF. ^bFrequency of occurrence of the best two predictors in the model equations, i.e., CG_y and CG_x , as the best two predictors in observer-generated equations. ^cNumber of patterns (of 50) correctly classified by LDF. ^dNumber of patterns (of 50) correctly classified according to schema rule. ^eLevel of significance of F approximation to Wilk's lambda, a test of overall equality of class means in the predictor space.

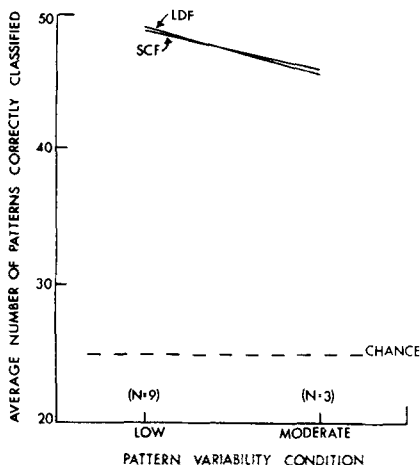


Fig. 6. Average number of patterns in O-generated classes correctly classified according to SCF and LDF classification systems for schematic Os.

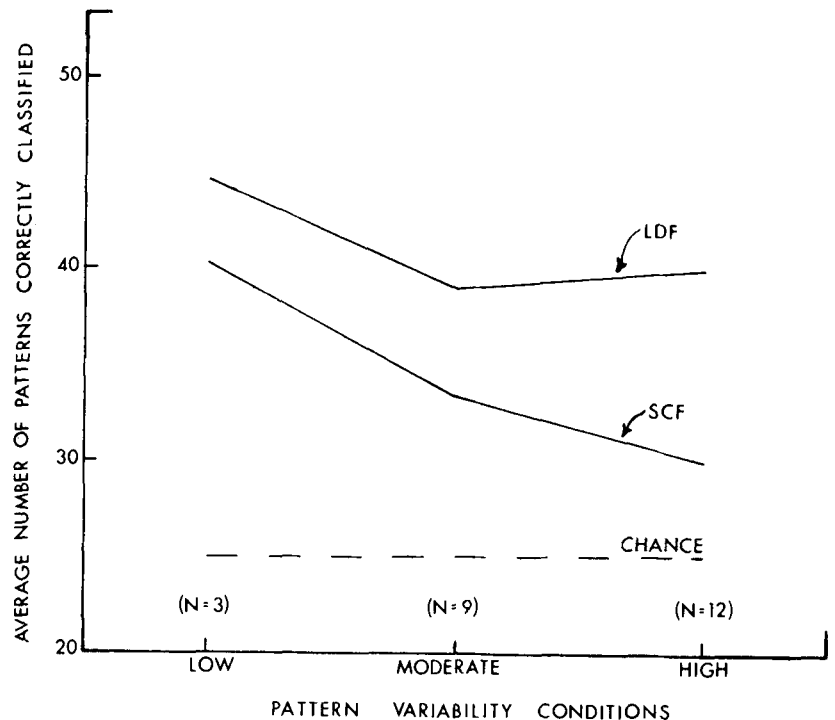


Fig. 7. Average number of patterns in O-generated classes correctly classified according to SCF and LDF classification systems for nonschematic Os.

does he simply behave randomly, or alternatively, or does he attempt to impose structure through the use of pattern features which appear salient to him? For all Os who failed to classify according to the a priori rules, physical features could be found which, to a large extent, reproduced their unique classifications, strongly suggesting that the first alternative, i.e., random behavior in the absence of clearly defined classes, may be eliminated.

If an O consistently used classification rules based on features which differ from the features defining a priori cluster, then this would be reflected operationally in the divergent weights assigned various physical pattern features in O-generated discriminant equations from those weights occurring in model equations. This was, in fact, what occurred for nonschematic Os: The significant physical measures in the individual LDFs differed, in some cases markedly, from those in the schematic model equation. Thus it appears that the O, in dealing with a relatively weakly structured stimulus configuration, will impose a consistent structure of his own generation, to organize the stimulus configuration.

Fenker¹, expanding upon the work of Rodwan and Hake (1964), has suggested a generalization of the Brunswik lens model to include discriminant analysis as a procedure for investigating feature utilization in complex pattern classification. The results of this study

indicate that such an approach is applicable to classification behavior and should be useful in future research in this area. The assessment of observer feature utilization in classification is easily extended to an examination of its relationship to perceived similarity, which variable is expected to be closely related to perceived class membership.

The problem of how an O reduces input from the environment by the imposition of class structure has been dealt with in a perceptual context by other researchers (e.g., Rodwan & Hake, 1964; Imai & Garner, 1968; Handel & Preusser, 1969; Wing & Bevan, 1969; and Aiken & Brown, 1971). However, this problem is best considered as a special case of a more general question concerning the development of models which map the organism's attempts to structure his environment.

REFERENCES

- AIKEN, L. S., & BROWN, D. R. A spatial analysis of the discriminability of forms in noise. *Perception & Psychophysics*, 1969a, 5, 171-175.
- AIKEN, L. S., & BROWN, D. R. Visual form perception: Congruence among spatial configurations. *Perception & Psychophysics*, 1969b, 5, 155-160.
- AIKEN, L. S., & BROWN, D. R. Visual form perception: Task demands and congruence among spatial solutions. *Psychonomic Science*, 1969c, 15, 219-220.
- AIKEN, L. S., & BROWN, D. R. Feature utilization of continuously varying attributes in visual pattern classification. *Perception & Psychophysics*, 1971, in press.

- AIKEN, P. M. FORTRAN programming of the Purdue Hybrid Computer. Purdue Laboratory for Applied Industrial Control, April 1969.
- BEACH, L. R. Cue probability and inference behavior. *Psychological Monographs*, 1964a, 78 (Whole No. 582).
- BEACH, L. R. Recognition, assimilation, and identification of objects. *Psychological Monographs*, 1964b, 78 (Whole No. 583).
- BEHRMAN, B. W., & BROWN, D. R. Multidimensional scaling of form: A psychophysical analysis. *Perception & Psychophysics*, 1968, 4, 19-25.
- BERSTED, C. T., BROWN, B. R., & EVANS, S. H. Free sorting with stimuli clustered in a multidimensional attribute space. *Perception & Psychophysics*, 1969, 6, 409-413.
- BROWN, D. R., & ANDREWS, M. H. Visual form discrimination: Multidimensional analyses. *Perception & Psychophysics*, 1968, 3, 401-406.
- BROWN, D. R., & BRUMAGHIM, S. H. Perceptual equivalence, pattern perception, and multidimensional methods. *Perception & Psychophysics*, 1968, 4, 253-256.
- BROWN, D. R., & OWEN, D. H. The metrics of visual form: Methodological dyspepsia. *Psychological Bulletin*, 1967, 68, 243-259.
- BRUMAGHIM, S. H., & BROWN, D. R. Perceptual equivalence between visual and tactual pattern perception: An anchoring study. *Perception & Psychophysics*, 1968, 4, 175-179.
- EVANS, S. H. A brief statement of schema theory. *Psychonomic Science*, 1967, 8, 87-88.
- HANDEL, S., & PREUSSER, D. The effects of sequential presentation and spatial arrangements of the free classification of multidimensional stimuli. *Perception & Psychophysics*, 1969, 6, 69-72.
- HUBEL, D. H., & WISEL, T. N. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception*. New York: Holt, Rinehart, & Winston, 1958. Pp. 150-166.
- IMAI, S., & GARNER, W. R. Structure in perceptual classification. *Psychonomic Monograph Supplements*, 1968, 2(9, Whole No. 25), 153-172.
- KOLERS, P. A. Some psychological aspects of pattern recognition. In P. A. Kolars and Murray Eden (Eds.), *Recognizing patterns: Studies in living and automatic systems*. Cambridge: M.I.T. Press, 1967. Pp. 4-61.
- MAVRIDES, C. M., & BROWN, D. R. Discrimination and reproduction of patterns: Feature measures and constraint redundancy as predictors. *Perception & Psychophysics*, 1969, 6, 276-280.
- MAVRIDES, C. M., & BROWN, D. R. Schematic concept formation: Feature measures and constraint redundancy as predictors. *Perception & Psychophysics*, 1970, 7, 239-243.
- MILLER, G. A., & CHOMSKY, N. Pattern conception. Proceedings of the University of Michigan Symposium on Pattern Recognition, 1958. (Mimeographed)
- RODWAN, A. S., & HAKE, H. W. The discriminant-function as a model for perception. *American Journal of Psychology*, 1964, 77, 380-392.
- STENSON, H. H. The psychophysical dimensions of similarity among random shapes. *Perception & Psychophysics*, 1968, 3, 201-214.
- STERNBERG, S. Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics*, 1967, 2, 45-53.
- WING, H., & BEVAN, W. Structure in the classification of stimuli differing on several continuous attributes. *Perception & Psychophysics*, 1969, 6, 137-141.

NOTE

1. Fenker, R. M. Multivariate extensions of Brunswik's lens model. Unpublished manuscript.

(Accepted for publication July 12, 1970.)