Functional equivalence between same-different classifications and judged similarity of Markov patterns

BILL R. BROWN[2] AND DONALD F. DANSEREAU
TEXAS CHRISTIAN UNIVERSITY

This experiment represented an initial attempt at comparing same-different classifications and similarity judgments of Markov histoforms sampled from three different schema families. A measure of individual stimulus variability was more strongly related to classification responses than to judged similarity. Dichotomous responses, derived from the similarity data by imposing threshold criteria, were found to be highly similar to real classification responses when the threshold values were estimated directly from the latter; little correspondence, however, was obtained using optimal classification criteria. The results support the notion that somewhat different features are used in making classification responses and similarity judgments, and indicate the need for developing more sensitive tasks involving controlled scanning of pattern information.

Since the early work of Attneave and Arnoult (1956) and Fitts, Weinstein, Rappaport, Anderson, and Leonard (1956), there has been a continuing effort to specify stimulus variables relevant to human pattern perception. Prominent among such attempts has been the development of the VARGUS 7 computer system (Evans, 1967a) for producing histoform stimuli that are distortions of a standard pattern or prototype (schema). A population of such instances, all of which are deviations from the same prototype, constitutes a schema family. Different schema families can be produced by using different prototypes, and these populations of stimuli may be conceptualized as clusters of points located in a multidimensional attribute space. Constraint redundancy (Rc), a measure discussed by Evans (1967b), reflects the tightness or looseness of each schema cluster in that it measures the adherence of the population of instances to the prototype. The deviation of each specific instance from its prototype is reflected by the statistic, Proportion of Schematic Steps (POSS), developed by Bersted, Brown, and Evans (1969).

The fact that Ss can successfully discriminate among VARGUS 7 instances belonging to different schema clusters has been demonstrated in a number of studies (Bersted, Brown, & Evans, 1969; Brown & Evans, 1969; Brown, Walker, & Evans, 1968; Dansereau & Brown, 1969; Evans & Arnoult, 1967). All of these experiments, however, used such dichotomous response measures as categorization responses, same-different classifications, and oddity choices. There are three points relevant to the use of dichotomous measures that require consideration.

First, since dichotomous responses are not as sensitive as graded response measures, they likely do not reflect adequately the detailed pattern information upon which Ss base their discriminations. The use of such measures as judged similarity may thus facilitate the specification of the relations existing between performance and stimulus variables frequently manipulated in mixed schemata discrimination tasks (e.g., Rc and POSS). Correlational analyses and multidimensional scaling procedures certainly should be more meaningful with graded response data.

Second, Dansereau and Brown (1969) suggested that Ss based their same-different classifications upon the similarity or dissimilarity of any two histoforms being judged (i.e., the physical distance between the two instances). Classification of each pair of instances as belonging to the same or different schema clusters would thus necessitate the application of some threshold criterion by the Ss, and the accuracy of classification responses would depend on the appropriateness of the threshold value used. This conception will be referred to here as the "similarity threshold" model. At least two studies (Brown & Evans, 1969; Dansereau & Brown, 1969) have shown that Ss exhibit strong biases toward classifying stimuli as belonging to different schema families (thus implying the selection and implementation of an inappropriate threshold value). Graded similarity judgments may therefore be more useful than are dichotomous responses as indicators of the Ss' ability to discriminate among VARGUS 7 stimuli. In the same context, to the extent that Ss actually base their classification responses upon judged similarity of instances, imposing a threshold value (estimated from actual classification responses) upon similarity judgments may produce dichotomous responses highly similar to the real classification data.

Third, several authors (Attneave, 1957; Evans, 1967c) have theorized that Ss classify instances belonging to different schema clusters in a manner analogous to schema plus correction (Woodworth, 1938). This supposition clearly requires that Ss be sensitive to stimulus variability and respond to the distance of each specific instance from the prototype. In such a case, Ss would likely base their classifications upon the degree to which the two instances are of similar distances from a prototype(s). This conception will be referred to here as the "prototype encoding" model. Since this distance measure is at least partially correlated with that treated by the "similarity threshold" model (i.e., increasing the deviations of the instances from the prototype likely increases the physical distances among the stimuli within a schema cluster), imposing a classification threshold upon similarity judgments would be expected to produce dichotomous responses similar to classifications based on the "prototype encoding" notion.

This experiment represented an initial attempt at comparing same-different classification responses and similarity judgments of VARGUS 7 histoforms sampled from three different schema families. Specifically, the study was intended to accomplish the following objectives: (1) to examine the effect of different magnitudes of Rc on same-different classifications and judged similarity; (2) to specify how classification responses, similarity judgments, and dichotomous responses, derived from graded similarity data by imposing a classification threshold, vary as a function of the POSS values; and (3) to determine the extent to which classification responses derived from graded similarity data differ from actual same-different classifications.

METHOD

Subjects

The Ss were 32 undergraduate students enrolled in the introductory psychology courses at Texas Christian University.

Stimuli

The stimuli were 24 column histoforms

generated by the VARGUS 7 computer program at 50% and 70% Rc. The program uses a seven-element Markov process; three distinct families of stimuli were produced by creating distortions of three prototypes, which, in this case, were three different most-probable sequences of column heights. The stimulus populations sampled in this experiment can be found in Bersted et al (1968) and are there identified as Schemata 2, 3, and 4.

Procedure

The Ss were given 60 discrimination trials with either the 50% or 70% Rc VARGUS 7 stimuli. On a typical trial, 16 of the Ss (eight Ss per Rc condition) viewed two stimuli and then judged them as belonging to the same schema family or different schema families; the remaining 16 Ss (8 Ss per Rc condition) judged the same pairs of patterns as to their degree of similarity on a 20-point scale Fifty per cent of the 60 trials involved the pairing of instances from the three different schema clusters (different trials), whereas the remaining trials involved the pairing of instances from the same schema family (same trials). A Kodak Carousal projector, impulsed by tapes that controlled the timing sequence, projected each of the 60 stimulus pairs onto a 2 x 2 ft rear-view screen. The Ss were permitted to view each pair for 8 sec; the stimuli were then removed from the screen, and Ss were given 7 sec in which to mark their responses. No knowledge of results or external reinforcement was administered.

No one instance from any of the three schema families was presented more than once during the 60 trials. The two stimuli presented on each trial, whether from the same or from different schema families, were paired on the basis of their having equal POSS values. Two stimulus presentation orders were used for each Rc condition. These orders were obtained in a random fashion, except for the following constraints: (1) Each block of 10 trials had an equal number of same and different trials; (2) blocks of 10 trials had approximately the same distribution of POSS values for each of the two Rc levels; and (3) the three schema families were approximately equally represented, in terms of both same and different pairings, within each block of 10 trials. The Ss participated in groups of two, and were assigned randomly to the treatment conditions as they appeared for the experimental session.

RESULTS

A two-way analysis of variance, with one between-group factor (magnitude of Rc) and one within-group factor (same vs different trials), was performed, with total correct classification responses as the dependent variable. There was a significant Rc main effect [$F(1,14) = 10.38, p < .01$]. Increasing the magnitude of Rc facilitated classification performance on both same and different trials. No other significant effects were obtained in this analysis.

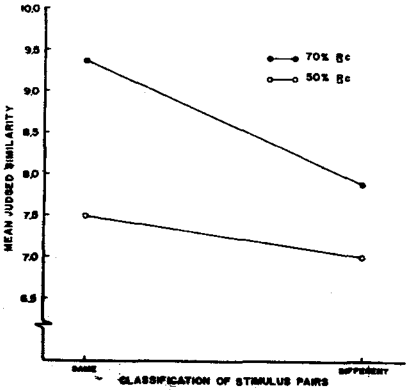

Fig. 1. Mean judged similarity as a function of classification of stimulus pairs and Rc conditions.

An analysis identical to that discussed above was performed using total similarity judgments as the dependent variable. There was a significant same-different trials main effect [$F(1,14) = 33.57, p < .01$], a significant Rc by Same-Different Trials interaction [$F(1,14) = 10.26, p < .01$], and a significant simple main effect of Rc for same trials [$F(1,14) = 5.13, p < .05$]. Figure 1 shows that mean similarity judgments for all Rc groups were higher on same trials than on different trials. Increasing the magnitude of Rc produced higher similarity ratings with stimuli sampled from the same schema cluster, but did not differentially affect similarity judgments of instances belonging to different schema families.

In order to obtain dichotomous classification responses from the graded similarity judgments, two different classification thresholds were applied to the similarity data within each Rc condition. The first data transformation (T1) involved imposing an optimal criterion that, in effect, produced an equal number of same and different classifications. The second data transformation (T2) involved the application of a classification criterion that produced a response bias (i.e., a bias toward classifying stimuli as belonging to different schema families) identical to that exhibited by Ss in making same-different classifications. The latter criterion for each Rc level thus represented an inappropriate or erroneous threshold rather than an optimal one. With respect to both T1 and T2 transformations, the threshold values were applied to the similarity data for each individual S across the 60 trials.

A three-way analysis of variance, with two between-group factors (Rc magnitude and real classifications vs T1 responses) and one within-group factor (same vs different trials), was performed, with total correct classification responses as the dependent variable. Only the Rc main effect was significant [$F(1,28) = 18.40, p < .01$]. Inspection of Fig. 2 shows that, on same and different trials, increasing the magnitude of Rc facilitated performance with respect to both the actual same-different classifications (the effect observed previously) and the T1 classification data. The T1 level of performance, however, did not differ significantly from that exhibited by Ss actually making same-different classifications, regardless of the magnitude of Rc.

An analysis identical to the one above, but using the T2 performance data, was performed with total correct classifications as the dependent variable. The Rc main effect was significant [$F(1,28) = 25.92, p < .01$], and the same-different trials effect was also significant [$F(1,28) = 7.17, p < .05$]. Inspection of Fig. 3 shows that, on same and different trials, increasing the magnitude of Rc facilitated performance for both actual same-different classifications and the T2 data. As with the T1 transformation, the T2 level of performance was not significantly different from the real same-different classifications. The T2 performance data did, however, approximate more closely the accuracy of the classification responses than did the T1 transformation data.

Pearson product-moment correlations,

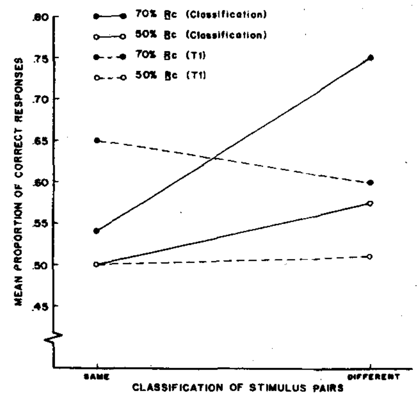

Fig. 2. Mean proportion of correct responses for actual same-different classifications and T1 transformation as a function of classification of stimulus pairs and Rc conditions.

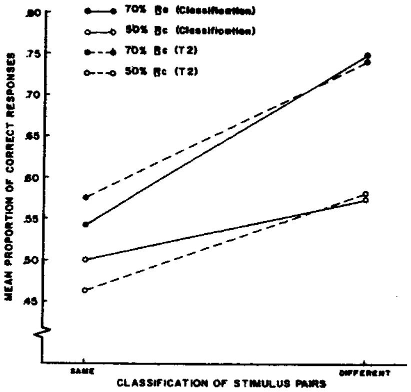

Fig. 3. Mean proportion of correct responses for actual same-different classifications and T2 transformation as a function of classification of stimulus pairs and Rc conditions.

computed over the two stimulus presentation orders, between-performance means, and POSS, were obtained for the same and different trials (N = 60, df = 58). Table 1 shows these correlations for the two Rc conditions and the level of significance attributed to each coefficient. In the majority of cases, the correlations were much higher for the 70% Rc condition than for the 50% Rc groups. With respect to the three types of classification responses, accuracy of judgments on same trials was facilitated as POSS increased in magnitude but decreased as POSS became higher on different trials. Similarly, on both same and different trials, the stimuli were rated more similar as POSS increased in magnitude.

One unexpected result was that for the 70% Rc condition, the relationships between performance and POSS were higher for the real same-different classifications than for those involving mean judged similarity. In the same context, the correlations for the T1 and T2 performance data were somewhat lower than were those for either Ss' actual classification responses or similarity judgments. The relationships between T1 performance means and POSS were also higher on same and different trials under the 70% Rc condition than were the corresponding correlations for the T2 data. The application of different classification criteria thus tended to produce weaker relationships between accuracy of judgments and the POSS statistic.

Table 1
Linear Correlations Between Performance Means and POSS Computed Over Two Stimulus Presentation Orders

	70% Rc		50% Rc	
Response Measure	Same Trials	Different Trials	Same Trials	Different Trials
Same-different Classifications	.63**	−.42**	.28*	−.33*
Similarity Judgments	.56**	.35**	.28*	.24
T1 Data	.41**	−.31*	.07	−.06
T2 Data	.28*	−.17	.11	−.04

* $p < .05$ ** $p < .01$

DISCUSSION

The results reported here provide support for the notion that Ss can discriminate, without external reinforcement or knowledge of results, among stimuli belonging to different schema clusters. Increasing the magnitude of Rc facilitated classification performance on both same and different trials, and also produced an increment in judged similarity of stimuli sampled from the same schema family. The Rc variable, however, had no effect on similarity ratings of stimuli on different trials. The latter result may be attributed to the fact that as the magnitude of stimulus distortion increased, Ss tended to use a smaller proportion of the 20-point rating scale. The absolute scale, in fact, shifted toward the "Very Dissimilar" point of the scale as Rc was decreased.

Increasing the magnitude of POSS on both same and different trials was accompanied by an increasing tendency on the part of Ss to classify instances as belonging to the same schema family. Consequently, classification responses on same trials became more accurate as the distance of the stimuli from the prototype decreased; classification errors on different trials, however, increased as the instances became closer to their respective prototypes. Similarly, as the magnitude of POSS increased, instances belonging to either same or different schema clusters were judged to be more similar. Together, the results for the two response modes thus indicate that Ss based their discriminations at least partially upon stimulus information other than the specific prototype characteristics.

One explanation for the above findings has been proposed by Bersted and Evans (1969). These authors, using a task having different response requirements from those of the present study, suggested that Ss are sensitive to repeated combinations of column heights in the instances, regardless of the absolute column heights represented in the combinations. The number of such repetitions becomes greater as POSS increases, and since the two stimuli on any one trial were equated with respect to POSS, the instances being judged exhibited exactly the same number of repeated combinations (although the repeated components were not necessarily the same for any two instances). To the extent that Ss are more sensitive to this "regularity dimension" than to specific prototype characteristics, they would be more likely to classify instances from different schema families as the same, or as being similar, as the magnitude of POSS increased.

Contrary to expectation, the relationship between graded similarity judgments and the POSS variable was not as strong as that involving same-different classifications. These results, coupled with the fact that dichotomous responses derived from the similarity data correlated even less with POSS, indicate that Ss do not use the same, detailed pattern information in making classification responses and similarity judgments. In order to obtain a close correspondence between real classification performance and the derived dichotomous data, it was necessary to apply to the similarity data a threshold criterion estimated from the actual same-different classifications. Derivation of dichotomous responses from similarity judgments in lieu of, or as a substitute for, real classification responses thus does not appear feasible with respect to accuracy of performance under the different Rc conditions.

The T1 and T2 data transformations were performed to determine if the "similarity threshold" model adequately represented the process by which Ss classify stimuli as belonging to the same or different schema clusters. There are several points, however, that require consideration in evaluating the relevance of this model to classification performance.

First, Ss making same-different classifications used inappropriate or erroneous classification criteria, as indicated by the bias toward judging stimuli as belonging to different schema families. If the "similarity threshold" conception is appropriate, the application of an optimal threshold value to the similarity data should thus have produced performance superior to classifications based on the erroneous criteria. Comparisons of the T1 data and several other imposed thresholds to actual classification performance failed to support this hypothesis. These results (coupled with the fact that same-different classifications were more highly correlated with POSS than were the similarity judgments and derived dichotomous responses) suggests that, to the extent that the "similarity threshold" model is correct, Ss making same-different classifications attend to stimulus attributes that are different from those used in similarity judgments. The Ss thus may base both classification responses and similarity judgments upon a measure of the distance between stimuli, but the measurement spaces within which the two types of responses are made would not be the same.

Second, since the distance measures treated by the "similarity threshold" and "prototype encoding" models are at least partially correlated, there is also the possibility that Ss were classifying on the basis of the distances between the instances and the prototype(s). Increasing the deviations of the instances from the prototype likely increases the physical distances among the stimuli within a schema cluster. The close correspondence obtained between classification performance and the T2 data is thus consistent with both the "similarity threshold" and "prototype encoding" conceptions, and indicates only that Ss were using a classification threshold.

Finally, the same-different task used in this experiment was not conducive to a detailed assessment of the stimulus information used by individual Ss across discrimination trials. In view of the fact that the "similarity threshold" and "prototype encoding" conceptions may both be relevant to the discrimination of stationary patterns, the development of more sensitive tasks involving controlled scanning of stimulus information is clearly required. The use of such tasks will permit more detailed treatment of individual performance data (including verbal protocols), and thus should facilitate the specification of the pattern components used in classifying stimuli clustered in a multidimensional attribute space.

REFERENCES

ATTNEAVE, F. Transfer of experience with a class-schema to identification learning of patterns and shapes. Journal of Experimental Psychology, 1957, 54, 81-88.

ATTNEAVE, F., & ARNOULT, M. D. The quantitative study of shape and pattern perception. Psychological Bulletin, 1956, 53, 452-471.

BERSTED, C. T., BROWN, B. R., & EVANS, S. H. A standard set of VARGUS 7 patterns at three levels of schematic redundancy. Psychonomic Monograph Supplements, 1968, 2(13, Whole No. 29), 251-282.

BERSTED, C. T., BROWN, B. R., & EVANS, S. H. Free sorting of stimuli clustered in a multidimensional attribute space. Perception & Psychophysics, 1969, 6, 409-413.

BERSTED, C. T., & EVANS, S. H. Distance from the schema as a predictor of classification and similarity judgments. Paper to be presented at annual meeting of the Psychonomic Society, November 1969.

BROWN, B. R., & EVANS, S. H. Perceptual learning in pattern discrimination tasks with two and three schema categories. Psychonomic Science, 1969, 15, 101-103.

BROWN, B. R., WALKER, D. W., & EVANS, S. H. Schematic concept formation as a function of constraint redundancy and knowledge of results. Psychonomic Science, 1968, 11, 75-76.

DANSEREAU, D. F., & BROWN, B. R. Acquisition and utilization of schematic concepts as a function of constraint redundancy and stimulus exposure time. Paper presented at annual meeting of the Southwestern Psychological Association, April 1969.

EVANS, S. H. VARGUS 7: Computer patterns from Markov processes. Behavioral Science, 1967a, 12, 323-328.

EVANS, S. H. Redundancy as a variable in pattern perception. Psychological Bulletin, 1967b, 67, 104-113.

EVANS, S. H. A brief statement of schema theory. Psychonomic Science, 1967c, 8, 87-88.

EVANS, S. H., & ARNOULT, M. D. Schematic concept formation: Demonstration in a free sorting task. Psychonomic Science, 1967, 9, 221-222.

FITTS, P. M., WEINSTEIN, M., RAPPAPORT, M., ANDERSON, N., & LEONARD, J. A. Stimulus correlates of visual pattern recognition: A probability approach. Journal of Experimental Psychology, 1956, 51, 1-11.

WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.

NOTES

1. This research was conducted by the Institute for the Study of Cognitive Systems and supported by a Department of Defense Project THEMIS contract, No. DADOO5-C-0176, under the Department of the Army.

2. Address: TCU Research Foundation, Fort Worth, Texas 76129.

(Accepted for publication August 11, 1969.)