Free sorting with stimuli clustered in a multidimensional attribute space[1]

CHRIS T. BERSTED,[2] BILL R. BROWN AND SELBY H. EVANS
TEXAS CHRISTIAN UNIVERSITY

An experiment was conducted to determine whether Ss learn, without knowledge of results or prior familiarization with the prototypes, to sort histoform stimuli generated by three probabilistic rules into categories consistent with these rules. Seven Ss were asked to sort 30 stimuli (10 from each schema population) on each trial. No constraints were placed on the number of categories to be used. A statistic was developed to measure consistency between schema-defined categories and subject-defined categories. Four of the seven Ss met the learning criterion, and demonstrated increasing consistency with the schema-defined categories across trials. In general, more categories than necessary were used to classify the stimuli correctly. The results raise several questions about classification strategies used by Ss in perceptual learning tasks with stimuli generated by probabilistic rules, and indicate the need for further research concerning the variables that influence perceptual category formation.

Schema theory (Attneave, 1957; Oldfield, 1954; Woodworth, 1938) has suggested that people encode stimuli in the form of schema plus correction. Evans (1967a) has suggested schematic concept formation (SCF) as a process for the acquisition of schemata. Schematic concept formation results in the abstraction of probabilistic schema rules for use in the classification and encoding of stimuli, and seems to have much in common with what Gibson (1955) has referred to as the abstraction of higher-order variables. Both higher-order variables and probabilistic schema rules would lead to greater differentiation between stimuli of different families.

A schema family is analogous to a cluster of points in multidimensional space, with points representing the individual instances. Constraint redundancy (Rc), as defined by Evans (1967b), refers to the relative adherance of instances to the schema rule, or the distance of instances from the centroid in multidimensional space. Schematic concept formation occurs when an individual is exposed to overdetermined stimuli (i.e., those that have more attributes associated with them than are necessary for near perfect classification) without external reinforcement or knowledge of results (Evans, 1967a).

Schematic concept formation has been demonstrated in a modified oddity task (Brown, Walker, & Evans, 1968), same-different discrimination task (Brown & Evans, 1968), classification tasks (Evans & Arnoult, 1967; Rosser, 1967), and a modified reproduction task (Bersted, Brown, & Evans[3]). The first two tasks require that Ss be able to make discriminations between classes of stimuli, but they do not demonstrate that the same hypothesis or decision rule is entertained by a S on each trial. A S may thus use different characteristics of the stimuli as a basis for discriminations on different trials. For example, different small segments of the schema rules, or subschemata, might be used on different trials. Similarly, the first two tasks do not permit a determination of the number of categories used.

The number of attributes or subschemata to which Ss are responding and the number of categories they use are related, but to some extent independently variable, kinds of behavior. Several subschemata associated with a schema family might result in classification of all instances exhibiting these subschemata into one category. This category may be called a schema-defined superordinate category, with subschemata as defining attributes of the category. On the other hand, the use of several subschemata could result in the classification of instances from one schema family into several categories. Neither the discrimination nor the oddity task can answer the question of whether or not subschemata represent the defining attributes of superordinate categories (i.e., schemata) or the bases of separate categories.

The Bersted et al[3] investigation demonstrated that Ss at least partially learned the schema rules in a modified reproduction task. The question of whether or not Ss used the same number of classification rules, as defined by the schema categories, remains unanswered. It would be possible to obtain reproduction results like those reported if Ss used subschemata as a basis for separate categories.

The classification tasks employed by Evans and Arnoult (1967), and by Rosser (1967), were constrained in that the Ss were told how many categories they should use. This procedure does not permit the assessment of the number of categories that would have been formed without the above constraint. In fact, the reversals (i.e., Ss classifying stimuli together from the same schema, but reversing the labels attached to them), found by Evans and Arnoult suggest that only a small portion of the stimulus was used for classification. The reversals themselves support the hypothesis that subschemata were bases for separate categories, and were not defining attributes of schema-defined superordinate categories.

Previous work by Evans (1964) suggested that a computer simulation of a model for SCF used more categories than necessary to classify the stimuli unless restricted to the number defined by the stimulus generation rules. An investigation by Shipstone (1960) used an unconstrained free sorting task with verbal stimuli generated by deterministic rather than probabilistic rules. Her data also suggested that at least some Ss will use more categories than are necessary to classify the stimuli correctly.

The present research was designed to determine whether or not SCF occurs with histoform stimuli generated by probabilistic rules in an unconstrained free sorting task. It was hypothesized that without knowledge of results, Ss would learn to sort the stimuli in a manner consistent with the probabilistic stimulus generation rules. The procedure also permitted the assessment of whether Ss do in fact use more categories than are necessary to classify the stimuli successfully. Successful use of the same number of categories as defined by the stimulus generation rules indicates either the use of a large portion of each schema rule for classification, or the use of subschemata as defining attributes of a superordinate category. The latter alternative implies just as much use of learned schema rules for classification as the first alternative. Successful use of more categories than are defined by the stimulus generation rules, however, would suggest that subschemata are used as the basis of separate categories, or possibly that some other variable is used for the subdivision of stimuli within a schema-defined category.

METHOD

Stimuli

The VARGUS 7 pattern-generating system (Evans, 1967c) was used to produce histoform stimuli 24 columns in length. A sample of 400 stimuli was generated for each of three different probabilistic generation rules (Schemata 3, 4, and 5 as designated by Bersted, Brown, and Evans, 1968) at 70% constraint redundancy. Ten stimuli were selected from each sample and shuffled together to form the stimulus deck for one trial. A different deck, with new stimuli, was used on each trial.

Subjects

The Ss were eight female undergraduate students paid at an hourly wage. One of the Ss had to be dismissed after 3 days because of a death in the family. The Ss were run individually.

Procedure

Instructions were read to the Ss on the first day of the investigation. The task was illustrated by pointing out that people can recognize handwriting patterns of individuals even though the handwriting may show variations from time to time. In addition, the Ss were asked to judge whether each of five pairs of words were written by the same person or by two different persons. The rationale for using these instructions was to communicate the idea that superficial differences can be associated with stimuli from the same category.

The Ss were then instructed to sort the stimuli (printed on 3 x 5 in. cards) into stacks in a manner analogous to sorting the handwriting of different individuals. The Ss completed two trials (30 stimuli per trial) a day for 7 days. The Ss were never exposed to a stimulus they had seen previously. Prior to the 15th trial (eighth day), all Ss, with the exception of one who was already using three categories, were instructed to reduce the number of categories if possible. Prior to the 19th trial (10th day), the Ss were instructed to use only three categories. On the 13th day, those Ss who had met the learning criterion (to be described later) were instructed to sort the stimuli into six stacks.

Scoring

The evaluation of performance in the present task posed two requirements. First, an objective method was needed for determining whether each S's sort was consistent with the stimulus-defined categories. Second, a measure was needed for evaluating any changes in sorting consistency across trials.

A statistic was developed for assessing performance consistent with the stimulus-generation rules. This statistic, termed the Index of Schematic Responding (ISR), was based on the assumption that the greater the number of stimuli from the same schema family grouped together, the better or more consistent the sort with the schema rules; conversely, it was assumed that the greater the number of stimuli from different schema families placed together in a category, the poorer the sorting performance. The formula for the computation of ISR for one S on any trial is given by Eq. 1:

$$ISR = 150 + \sum_{i=1}^{N} (F_i - D_i) \qquad (1)$$

where N is the number of categories used by the S, and the quantities in parentheses are based on a partition of the set of all distinct unordered pairs of stimuli in the i^{th} category as follows: F_i is the number of such pairs in which both stimuli are from the same schema family; D_i is the number of such pairs in which the stimuli are from different schema families. The additive constant is chosen to avoid negative values.

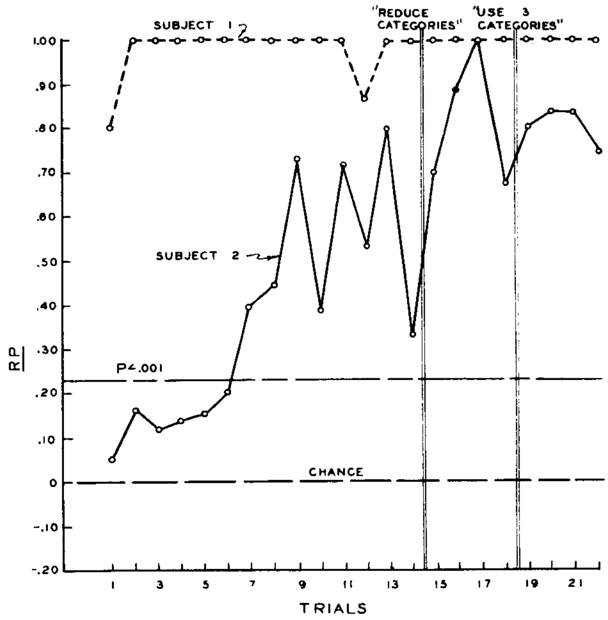

Fig. 1. Relative performance of Ss 1 and 2 across trials.

In order to determine whether the ISR value obtained on any trial could reasonably have been expected by chance, sampling distributions for the statistic were estimated by a Monte Carlo procedure for each of the different numbers of categories used by the Ss. A computer program, based on the null hypothesis that instances were randomly placed in categories, was used to determine these distributions. The distributions provided an estimate of the probability of any value of the statistic for each possible number of categories. The estimated probability levels afforded a statistical test of the statistical significance of any ISR value for any number of subject-defined categories.

In order to evaluate changes in sorting consistency across trials, a measure was needed to reflect consistency between subject- and schema-defined categories irrespective of the number of categories. Since the range of possible values of the ISR statistic decreases as the number of subject-defined categories increases, plotting absolute ISR values across trials would not necessarily reflect increases in consistency of sorting if a S used a different number of categories on successive trials. The measure selected was a proportion that reflected the difference between the obtained ISR score and the expected value of the ISR statistic, relative to the maximum possible difference in ISR for each number of possible categories. This measure, called Relative Performance (RP), was calculated by Eq. 2:

$$RP = \frac{ISR - E\,V(ISR)}{MAX(ISR) - E\,V(ISR)} \qquad (2)$$

where ISR is the obtained score for one S on a given trial, E V(ISR) is the expected value for ISR, given the number of categories used on that trial, and MAX(ISR) is the maximum value for ISR, given the number of categories on that trial. With increases in the number of categories, the expected value of ISR increases and the maximum ISR value decreases.

RESULTS

A stringent learning criterion was selected ($p < .001$ of an individual sort occurring by chance) in order to assure that sorting consistent with the stimulus-generation rules had

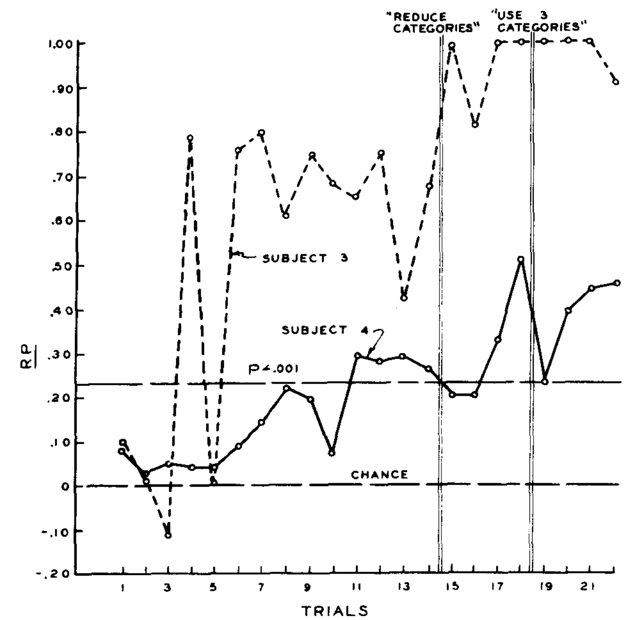

Fig. 2. Relative performance of Ss 3 and 4 across trials.

occurred. If one were to object to repeated testing of the null hypothesis over a series of trials, then performance exceeding this criterion across several trials would surely demonstrate sorting consistent with the stimulus-generation rules. Four of the seven Ss met this criterion during the first 7 days.

Subject 1 (Fig. 1) sorted consistently with the schema-defined categories thoughout the experiment. The only real change in performance was a reduction from five categories on the first trial to three categories on the second and subsequent trials. Subject 2 (Fig. 1) exhibited gradually increasing use of the schema-defined categories across trials. Subject 3 (Fig. 2) showed rapid improvement and then asymptotic or slowly deteriorating performance until told to reduce the number of categories used. Subject 4 (Fig. 2) exhibited slowly increasing performance and did not meet the criterion until the sixth day of the experiment (11th trial). It should be noted, however, that all of the RP values given for S 4 were above chance, not distributed about chance as would be expected if her performance was completely independent of the schema-defined categories.

The three remaining Ss did not consistently meet the learning criterion, although one did meet the criterion on one trial. These Ss exhibited vacillating performance across trials, with ISR values falling above and below the expected value of this statistic. The large variability suggests that the Ss changed their sorting strategies on different trials.

Figure 3 presents the mean performance across trials of the four Ss who met criterion, as measured by the RP statistic. The curve in this figure shows sorting behavior that is increasingly consistent with the schema-defined categories. Figure 4 presents the mean number of categories used by these four Ss across trials. Included for reference is the mean number of categories used by the remaining three Ss (i.e., Ss not meeting criterion). Together, Figs. 3 and 4 show sorting consistency increasing, but little if any reduction in the number of categories used. For the four trials immediately preceding the instructions to reduce the number of categories, a mean of 7.62 categories was used by the Ss meeting criterion. For all trials on which the criterion was met, the maximum number of categories used was 16, and the minimum number used was 3. Table 1 summarizes the mean number of

Table 1
Mean Number of Categories Used on the Four Trials Preceding Instructions to Reduce Categories for Each Subject

	S No.	Mean	S. No.	Mean	S No.	Mean	S No.	Mean
Subjects Meeting Criterion	1	3.0	2	5.75	3	7.0	4	14.75
Subjects Not Meeting Criterion	5	5.25	6	7.50	7	12.5		

categories used by the four Ss who met the learning criterion for the four trials prior to instructions to reduce the number of categories, and for the three Ss who did not meet the learning criterion.

On the last day of the experiment, the four Ss who had met the learning criterion were instructed to use six categories. This procedure was introduced to obtain information on what characteristics Ss might use when they form more categories than necessary. Previous investigations (Bersted et al[3]; Rankin, Markley, & Evans[4]) have demonstrated that Ss can be very sensitive to differences in stimulus variability (i.e., the degree of adherence to the schema rule). It was therefore expected that Ss might differentiate, within each schema-defined category, between stimuli close to the schema and those more distant.

In order to test the hypothesis that Ss were forming categories within each schema family on the basis of adherence of stimuli to the prototype, a measure of this adherence was needed. The measure, Proportion of Schematic Steps, or POSS (Bersted et al, 1968), represents the extent to which individual stimuli adhere to the schema rule (based on the proportion of transitions between column heights that follow the schema rule). For any particular sample of stimuli generated at a given Rc level, the mean POSS value approximates the transitional probabilities between the most probable column heights of the sequence (i.e., the schema rule). The POSS values for individual stimuli are distributed about this mean.

On the final day, when the four Ss were asked to use six

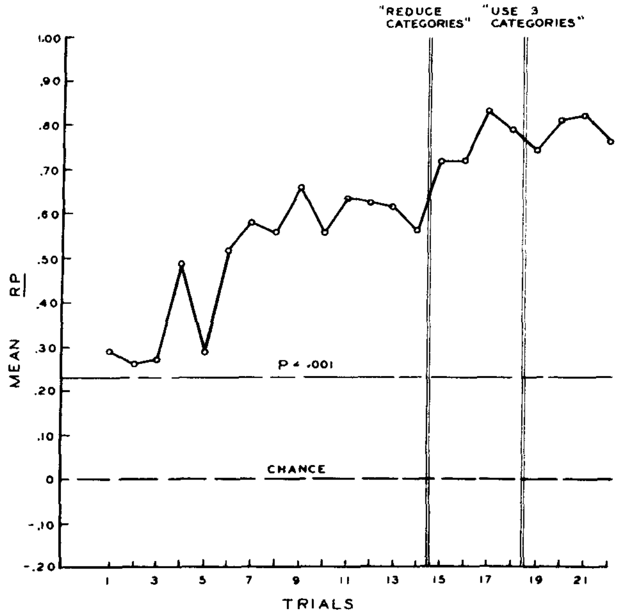

Fig. 3. Mean relative performance across trials for the four Ss who achieved the learning criterion.

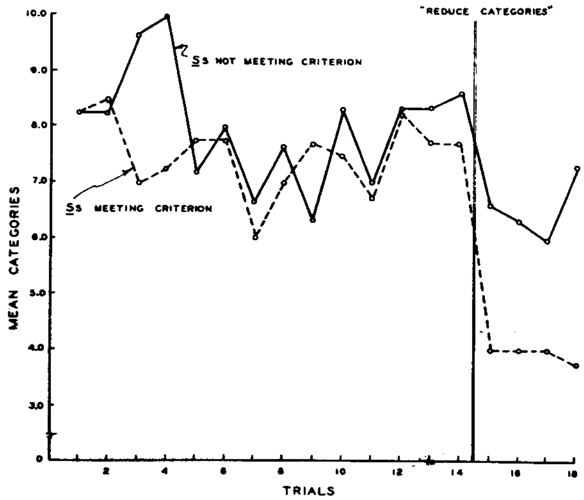

Fig. 4. Mean number of categories used across trials for the four Ss who achieved the learning criterion, and for the three Ss who did not meet the learning criterion.

categories, each of these Ss placed 5 stimuli into each of six categories. Two of the Ss grouped the 10 stimuli from each schema family into two separate stacks of five each. The other two Ss each made two reversals (i.e., four from one schema family placed with one stimulus from another schema). For each of these four Ss, an independent t test (POSS as the dependent variable) was performed on each pair of stacks that consisted primarily of stimuli from the same schema family. The predominantly significant results presented in Table 2 suggest that at least three of the four Ss adopted a strategy of sorting stimuli from the same schema family into one of two stacks: One stack contained stimuli that showed few deviations from the schema rule (i.e., high POSS values), and the other stack contained stimuli that showed more deviations from the schema rule (i.e., low POSS values).

DISCUSSION

The hypothesis that SCF occurs in an unconstrained free-sorting task was supported in that four of the seven Ss met the learning criterion. These results represent the first demonstration of SCF in which Ss were not given information concerning the number of categories (i.e., number of types of stimuli) or responses available in the task. When Ss have been told how many different types of stimuli there are within a task, generalization to the unconstrained real world environment is perhaps risky. In the natural environment, SCF is presumed to occur on the basis of information from perceiving objects, with no knowledge of the number of categories of which the objects are members. On the other hand, when Ss are limited to only certain specified responses (e.g., same-different discriminations), there is no way to determine whether or not the Ss are using the number of categories defined by the schema rules.

Although SCF was explicitly demonstrated in this study, the question remains as to why some Ss sort consistently with the schema-defined categories, but use more categories than are defined by the schema rules. It is obvious (Figs. 1 and 2) that Ss can, when instructed, combine this larger number of categories into the appropriate number of categories with only slight changes in consistency. Evans's (1964) computer simulation suggested the use of subschemata as a possible explanation for this multiplication of categories by Ss. The computer in many cases classified stimuli on the basis of only several attributes

Table 2
Results of t-tests, with POSS as the Dependent Variable, for the Two Categories Formed within Each Schema Family

	Schema 3	Schema 4	Schema 5
Subject 1	*		
Subject 2	***	**	*
Subject 3	***	****	*
Subject 4	****		****

* $p < .10$ *** $p < .025$
** $p < .05$ **** $p < .01$

within each stimulus, with the result that patterns from the same schema family were classified into several categories.

A second alternative, however, exists. Previous research (Bersted et al[3]; Rankin et al[4]) has suggested that Ss can be very sensitive to the degree of adherence to the schema rules of the stimuli. Specifically, it was suggested (Bersted et al[3]) that Ss are able to detect small changes in the variations from the schema of the stimuli that results in classification not only in terms of the schema family, but also in terms of the variability of the individual stimuli. The findings reported in Table 2 support this notion. A better test of this hypothesis, however, would be to determine whether or not the above type of behavior was exhibited in the unconstrained part of the experiment. Unfortunately, these data were not collected. In any case, the present data do show that the division of stimuli from a schema family into groups that adhere to the schema rule to differing degrees is a likely result in a constrained situation with highly practiced Ss.

It is of interest that Shipstone (1960) found a mean of eight categories used by her Ss in a free sorting task. Shipstone suggested that Ss may naturally respond with a relatively fixed number of categories in accordance with Miller's (1954) magic number seven, plus or minus two. Figure 4 shows that for the Ss in the present investigation who met the learning criterion, the mean number of categories used during the first 14 trials was between 6.0 and 8.5. The Ss not meeting the learning criterion used a comparable mean number of categories (Fig. 4). These means, however, are misleading because of the high variability between the individuals (Table 1).

It may be suggested, nevertheless, that some Ss have a preferred number of categories; this suggestion has two important implications. First, if the number of schemata in an SCF task is increased, performance consistent with the schema rules might also increase. This follows because, as the number of schemata in a task approaches the number of preferred categories of a S, one would expect fewer schema classes to be subdivided into classes associated with subschemata or with variability levels. Brown and Evans (1969) have demonstrated that fewer "different" responses are made when three schemata are present in a same-different task than when only two schemata are present, a finding consistent with the above suggestion. A second implication of this suggestion is that variations in instructions (e.g., information as to how many schemata are present) may lead to better performance in comparison to the absence of this information because of fewer variations between the number of subject-defined categories and schema-defined categories. It should be emphasized that these latter suggestions are not tied to any one preferred number of categories, but merely to the assumption that individual Ss tend to have some preferred number of categories.

The results from the present experiment are important for several reasons. First, they demonstrate that SCF can occur even when Ss are not given limited response alternatives or told the number of categories present within the task. Second, the results suggest that Ss tend to use more categories than are necessary to classify the stimuli correctly. The use of more categories than necessary suggests either that subschemata are used to define different categories, or variability levels of the stimuli result in the multiplication of categories. The relevance of each of the latter two alternatives to the SCF process remains to be determined.

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.

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.

BROWN, B. R., & EVANS, S. H. A model for schematic concept formation: Further development and assessment. Paper presented at the Psychonomic Society, St. Louis, November 1968.

BROWN, B. R., & EVANS, S. H. Perceptual learning in pattern discrimination task 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.

EVANS, S. H. A model for perceptual category formation. Unpublished doctoral dissertation, Texas Christian University, 1964.

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

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

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

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

MILLER, G. A. The magical number, seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

OLDFIELD, R. C. Memory mechanisms and the theory of schemata. British Journal of Psychology, 1954, 45, 14-23.

ROSSER, E. M. Categorization and discrimination of tone sequences. Unpublished doctoral dissertation, Harvard University, 1967.

SHIPSTONE, E. I. Some variables affecting pattern conception. Psychological Monographs, 1960, 74(17, Whole No. 504).

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. DAD005-C-0176, under the Department of the Army.

2. Address: Institute for the Study of Cognitive Systems, Texas Christian University, Fort Worth, Texas 76129.

3. Bersted, C. T., Brown, B. R., & Evans, S. H. Acquisition and utilization of schematic concepts as a function of duration of rest interval, constraint redundancy shifts, and length of intertask interval. In preparation.

4. Rankin, W. C., Markley, R. P., & Evans, S. H. Pythagorean distance and the judged similarity of schematic stimuli. In preparation.

(Accepted for publication April 16, 1969.)