Journal of Experimental Psychology: General
1987, Vol. 116, No. 3, 238–249

0096-3445/87/$00.75

Similarity, Confusability, and the Density Hypothesis

James E. Corter
Teachers College, Columbia University

Results from several letter- and digit-identification studies have been interpreted (Appelman & Mayzner, 1982; Krumhansl, 1982) as providing support for the hypothesis that psychological similarity is influenced by the local density of items in the stimulus space. This conclusion is questioned on the grounds that density was not directly manipulated in the studies, thus alternative explanations based on other stimulus characteristics cannot be excluded. In the present article six experiments are reported in which stimulus density was manipulated. In three experiments using similarity ratings and two using discrimination confusions, no effect of stimulus density was found. However, the identification counterpart of one of the discrimination studies did provide evidence of an effect of density on response probabilities. It is concluded that stimulus density can affect identification performance through its influence on the choice process implicit in any identification task, but is not an important determinant of psychological similarity.

The notion of similarity plays an important role in psychological theories of categorization, learning, memory, and choice. Data from many types of tasks may be analyzed as or converted into measures of similarity: direct ratings of the similarity or dissimilarity of stimulus pairs, discrimination confusions, identification errors, errors in classification or paired-associates learning, correlation or co-occurrence data, reaction times, and so forth.

Probably the most widely used models for similarity data have been geometric ones. The introduction of multidimensional scaling algorithms (Shepard, 1962a, 1962b; Torgerson, 1958) provided useful methods for analyzing and representing similarity or other proximity data. In recent years multidimensional scaling and related methods have been applied especially to analyze perceptual data (including confusion data of various sorts) and in market research. However, Tversky (1977) questioned the appropriateness of several of the assumptions underlying geometric models of similarity. These models assume that objects can be represented as points in a coordinate space such that the observed dissimilarities are directly related to the distances between the points in the space. However, such a representation of the proximity data as distances in the space requires that the proximity data satisfy both certain dimensional assumptions and the axioms of a metric space: minimality, symmetry, and the triangle inequality.

Minimality requires that distances in a metric space satisfy $\delta(x, y) \geq \delta(x, y) = 0$. That is, the distance of an object to itself must be 0, and the distance of an object to any other object must be greater than 0. If the proximity data are dissimilarities that are monotonically related to the underlying distances, then minimality requires that the dissimilarity of an object to itself is the same for all objects, and less than the dissimilarity of the object to any other. Many methods of gathering proximities do not provide any estimate of "self-similarity." One type of task that can provide such estimates is a discrimination task, in which subjects are required to respond "same" or "different" to stimulus pairs. The proportion of correct "same" responses can be used as a measure of self-similarity. Data from such experiments (e.g. Rothkopf, 1957) suggest that minimality is often not satisfied.

Symmetry requires that the distance (or dissimilarity) of x to y be equal to the distance (dissimilarity) of y to x: $\delta(x, y) = \delta(y, x)$. This assumption too is often violated, in similarity data (Tversky, 1977), identification confusions (e.g., Appelman & Mayzner, 1982; Gilmore, Hersh, Caramazza, & Griffin, 1979; Keren & Baggen, 1981; Townsend, 1971), and discrimination confusions (e.g., Rothkopf, 1957; Tversky, 1977).

The triangle inequality states that for any three objects, x, y, and z, $\delta(x, y) + \delta(y, z) \geq \delta(x, z)$. This property cannot be falsified by ordinal or interval dissimilarity data because adding a large enough positive constant to all the dissimilarities ensures that it will be satisfied. However, examples can be constructed so that the additive constant needed is quite large relative to the actual observed dissimilarities, casting some doubt on the psychological plausibility of the triangle inequality constraint.

The work described here was reported in the author's doctoral thesis submitted to the Department of Psychology, Stanford University, 1983, and was supported in part by National Institute of Mental Health Training Grant 5 T32 MH 15157-05.

I wish to thank my advisor, Amos Tversky, for his support and advice during this project. Working with him was both a great honor and a great pleasure. I also wish to thank the other members of my dissertation committee, Roger Shepard, Ewart Thomas, and Brian Wandell, for their suggestions. J. Douglas Carroll and Yoshio Takane provided interesting comments on later versions of the manuscript, and William Prinzmetal made a number of useful suggestions. Thanks are due to Leslie Jonath for her help in conducting Experiments 5 and 6. Finally, I gratefully acknowledge the support of my wife, Lisa Church, during all stages of the work.

Correspondence concerning this article should be addressed to James E. Corter, Box 41, Teachers College, Columbia University, New York, New York 10020.

The Contrast Model

Tversky (1977) proposed an alternative model of similarity, based not on a geometric or spatial representation, but rather on a set-theoretic approach. His contrast model (sometimes referred to as the feature-matching model) assumes that objects are characterized not by their values on a set of quantitative dimensions, but rather by sets of qualitative features. The contrast model describes the similarity of objects x and y as a linear combination (i.e., a contrast) of functions of the feature sets associated with x and with y. Specifically, the model can be written in terms of dissimilarities as

$$d(x, y) = -\theta F(X \cap Y) + \alpha G(X - Y) + \beta G(Y - X), \quad (1)$$

where X and Y refer respectively to the features of objects x and y; F and G are functions on these feature sets, with positive ranges; and α, β, and θ are positive constants that reflect the weighting of the corresponding terms. Equation 1 says that the dissimilarity of x and y is a negative function of some measure of the features common to x and y $(X \cap Y)$, and a positive function of the measure of the features that belong to x but not to y $(X - Y)$ and the features that belong to y but not to x $(Y - X)$. These latter two feature sets are referred to as the distinctive features of x and of y.

The contrast model can account for violations of minimality and asymmetry, as well as certain other phenomena associated with proximity data, such as systematic differences between ratings of similarity and of dissimilarity and changes in similarity with context (Tversky, 1977). Briefly, violations of minimality (e.g., as when x has a higher self-similarity than y) are assumed to be related to differences in the measures of the feature sets X and Y. The higher the measure of X, the more similar x will be to itself. Predicting asymmetries in proximity measures requires an additional assumption, the *focusing hypothesis*. The focusing hypothesis suggests that certain experimental tasks will cause the subject to focus more on one stimulus (x) than on the other (y), which will result in the features of x being weighted more heavily. That is, α in Equation 1 will be increased relative to β. If the objects have different total feature measures (i.e., if one object is more prominent or has more salient features than the other), then the two distinctive feature terms of Equation 1 must have different measures, and asymmetries will result. If α is greater than β, and x has greater total feature measure, then $d(x, y) > d(y, x)$.

Thus, the contrast model can account for violations of the axioms underlying traditional geometric models of similarity. In addition, the description of the attributes of objects as discrete qualitative properties rather than as values on a set of quantitative dimensions has obvious appeal in many psychological applications.

The Distance–Density Model

Krumhansl (1978) suggested that the violations of the metric axioms summarized by Tversky (1977) could be handled within a geometric-type model, in which the geometric distance between two stimuli x and y is modified by the density of other points in the space, especially in the regions of x and y. This results in a modified distance function, d', which reflects the psychological dissimilarity of x and y. One possible form of this dissimilarity function is given as

$$d'(x, y) = d(x, y) + \alpha D(x) + \beta D(y), \quad (2)$$

where $d(x, y)$ is the "true" distance between x and y in the stimulus configuration, $D(x)$ and $D(y)$ are some measures of local density in the regions of x and y, (e.g., the number of other stimuli within a certain radius), and α and β are positive constants. Psychologically, the model suggests that dense regions of the stimulus space are expanded in some way, perhaps because finer discriminations are made between stimuli in that region. Thus, increasing the density in a region would increase the judged dissimilarity of objects in that region.

Violations of minimality are predicted by the distance–density model whenever the densities of all objects are not equal because according to Equation 2 the dissimilarity of x to itself is equal to $(\alpha + \beta)D(x)$. Thus, differences in self-similarities are predicted to be related to differences in the local densities of objects, rather than to differences in the measures of their feature sets (i.e., their salience or prominence) as predicted by the contrast model. In conjunction with the focusing hypothesis, the distance–density model also predicts asymmetries. If the experimental task causes the subject to focus more on stimulus x, then α in Equation 2 will be increased, which corresponds to a tendency on the part of the subject to pay relatively more attention to the density in the region of x. If the stimulus density around x is greater than around y, then the model predicts that $d(x, y) > d(y, x)$.

Evidence for the Distance–Density Model

In support of the distance–density model, Krumhansl (1978, 1982) presented reanalyses of several proximity matrices derived from previous studies. For example, she analyzed the data of Rothkopf (1957), who collected "same–different" confusions for Morse code signals. To test if density did seem to affect the confusability of stimulus pairs, she selected both *interior* stimulus pairs (i.e., those pairs for which at least one signal fell near the center of the configuration in a multidimensional scaling solution for the data), and *exterior* pairs. For interior and exterior pairs with approximately equal distances in the solution, the average confusability for interior (i.e., denser) pairs was significantly less than for exterior pairs, as predicted by the model.

Evidence that differences in self-similarity may be related to differences in densities was obtained from analyses of several sets of "same–different" confusions (Balzano, 1977; Rothkopf, 1957) and learning data. However, Krumhansl (1978) notes that there are certain constraints between the diagonal and off-diagonal entries of a matrix containing learning or identification data. That is, a stimulus that is very often confused with other stimuli will have large off-diagonal entries, which means that the corresponding diagonal entry (i.e., the frequency of correct responses or self-confusions) will necessarily be small. Thus, the quantities (density and probability of correct response) predicted to be related by the model are not independent for this type of data.

Evidence supporting the distance–density model's explanation of asymmetries was also obtained from the Rothkopf (1957) data. Krumhansl (1978) found that those signal pairs that showed large asymmetries in confusion probabilities also showed asymmetries in stimulus densities. Note that Tversky (1977) presented an analysis of these data that demonstrated a relation between confusion asymmetries and number of features (i.e., length of signal), as predicted by the contrast model.

Appelman and Mayzner (1982) analyzed several sets of previously published letter-identification data, concluding that self-confusions (i.e., the probability of correct identification of a letter) were better predicted by density than by number of features, and that asymmetries were more strongly related to differences in stimulus densities than to differences in number of features. However, the direction of asymmetries was opposite to that predicted by the distance–density model (assuming that the presented stimulus is focused on and that its density is more heavily weighted). In Appelman and Mayzner's phrase, they found that the denser letter "gives more confusions than it gets." Krumhansl (1982) analyzed some digit identification data reported by Keren and Baggen (1981), concluding that the contrast and distance–density models were roughly comparable in overall fit and ability to predict asymmetries, but that self-confusions were better predicted by density than by number of features.

Need for Experimental Testing of the Density Hypothesis

The evidence supporting the notion that psychological similarity is partly a function of local stimulus density depends on reanalyses of proximity data sets. One problem with these analyses is a certain circularity between the measures of density used and the dependent measure, the actual proximities. In many of the analyses reported by Krumhansl (1978), proximities were correlated with measures of density indirectly derived from those same proximities (such as the number of other stimuli falling within a certain radius in a multidimensional scaling solution). The analyses in Appelman and Mayzner (1982) and Krumhansl (1982) used more independent measures of distance and density, in that stimulus densities were obtained by a featural analysis of the stimuli, and not estimated using multidimensional scaling solutions derived from the proximity data themselves.

The major problem in interpreting the results of these analyses is that such nonexperimental methods can provide only suggestive evidence for or against the density hypothesis. The reason is that density is confounded with a number of different variables for these data sets, thus alternative explanations based on these other variables cannot be excluded. For example, in Keren and Baggen's (1981) digit data, density and number of common features are correlated .77 across digit pairs (Krumhansl, 1982). For the letter data of Gilmore et al. (1979) analyzed by Appelman and Mayzner (1982), the corresponding correlation is .62. In a number of the data sets (e.g., Rothkopf, 1957), density is highly correlated with typicality. Consequently, effects such as asymmetries in confusion probabilities might be caused by differences between stimuli in density, in salience or prominence (as predicted by the contrast model), in typicality, or other stimulus characteristics. Distinguishing these potential explanations requires experimental control of stimulus materials.

Furthermore, the types of data that have provided evidence suggestive of density effects consist mainly of discrimination and identification confusions. If stimulus density affects psychological similarity, density might be expected to have an influence on other types of proximity data as well, for example, on direct ratings of similarity and dissimilarity. Identification tasks in particular have characteristics that set them apart from other paradigms in which proximity data are gathered. In most of these paradigms, two stimuli are presented perceptually, and the subject makes some response or judgment based on the similarity of the pair. In an identification task only one stimulus is actually presented. A "pair" of stimuli exists only in the sense that an incorrect response may be made to a presented stimulus. Also, the subject must choose a response from the set of valid alternatives. Therefore, a choice process is implicit in the task.

In summary, meaningful testing of the density hypothesis requires experimental manipulation of stimulus materials. In addition, evaluating the generality of any such effects calls for testing of the hypothesis across different types of proximity measures, such as ratings and confusion probabilities. The first requirement, experimental manipulation of density, can be achieved in studies of proximity relations by using artificial stimuli or by selecting subsets of the stimuli from a large population. These methods were used in the studies reported by Tversky (1977) and Tversky and Gati (1978), which provided evidence for the influence of common and distinctive features on similarity relations.

In the studies reported in the present article (three using direct ratings of similarity, two using discrimination confusions, and one using identification errors), between-subjects designs were used to assess the effects of manipulating density on proximity measures. In each of the studies there were two conditions. One group of subjects rated stimuli that were designed to have one relatively dense region in the neighborhood of a particular target stimulus. The other group was given a stimulus set that had the same total number of stimuli, but in which the density was increased for a different target stimulus. The prediction of the density model is clear in such an experiment: The increased density in the region of the target stimulus should decrease the similarity (or confusability) of that stimulus to all other stimuli, relative to the other condition. This method of manipulating density was used in Experiments 1–4. In Experiments 5 and 6, density was manipulated (again in a between-subjects design) by selecting different subsets of a population of familiar stimuli (block letters). Each group saw a different subset of six letters, designed to have high density for a particular target letter.

Similarity-Rating Studies

Experiment 1

Experiment 1 tested for density effects in similarity ratings of pairs of ellipses varying in height and width. In each condition

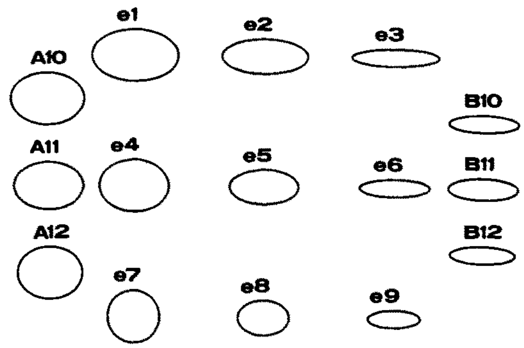

Figure 1. Ellipse stimuli of Experiment 1. (Frame ellipses are designated e1–e9, and the context ellipses for Conditions A and B are designated A10–A12 and B10–B12.)

of the between-subjects design, the local density for a certain target ellipse was increased by adding ellipses with nearly identical values of height and width.

Method

Subjects. A total of 51 Stanford undergraduates participated in the study as part of a course requirement.

Materials. Two sets of 12 ellipses each were constructed. In each set, 9 of the ellipses were frame *stimuli:* That is, these ellipses were common to both sets of stimuli. The remaining three were unique to one of the two sets, and constituted the density manipulation for that set. In Condition A these three *context* stimuli were constructed so as to be highly similar to Ellipse e4, and in Condition B so as to be highly similar to Ellipse e6. Ellipse e4 is therefore referred to as the target ellipse for Condition A, and Ellipse e6 is the target ellipse for Condition B. According to the density hypothesis, the rated similarity of the target ellipse to the other frame ellipses should be decreased (relative to the other condition) by the additional context stimuli. Both sets of ellipses are shown in Figure 1.

Procedure. Subjects were randomly assigned to one of the two conditions ($n = 24$ for Condition A, $n = 27$ for Condition B). In order to make the composition of the stimulus set (hence the density manipulation) more salient, they were initially asked to look through a training booklet and study the set of ellipses for their condition. The booklet contained one ellipse per page. When they indicated that they had familiarized themselves with the set, they were given a test booklet containing the 66 pairs of the 12 ellipses in a random order, 1 pair per page. Ratings of visual similarity were made by circling numbers on a 9-point scale on a separate answer sheet. The pairs were presented in a separate random order for each subject.

Results

The average ratings of similarity among the ellipses in the two conditions are shown in Table 1. According to the density hypothesis, the increased density in the region of Ellipse e4 in Condition A should cause Ellipse e4 to be rated less similar to the other frame stimuli, and analogously for Ellipse e6 in Condition B. The hypothesis was tested by defining a density statistic that compared the similarity ratings involving target–frame stimulus pairs for the two conditions.

The actual statistic used to test the hypothesis was

$$\psi = S_{4,}{}^{*} - S_{6,}{}^{*}$$

where

$$S_{4,}{}^{*} = S(e4, e1) + S(e4, e2) + S(e4, e3) + S(e4, e5) + S(e4, e7) + S(e4, e8) + S(e4, e9),$$

and

$$S_{6,}{}^{*} = S(e6, e1) + S(e6, e2) + S(e6, e3) + S(e6, e5) + S(e6, e7) + S(e6, e8) + S(e6, e9).$$

$S(e4, e1)$ refers to the rated similarity of Ellipse e4 and Ellipse e1. The prediction of the density hypothesis was tested by comparing the value of this statistic for the two conditions. Note that the density hypothesis predicts that the value of this statistic be smaller for Condition A (density increased for Ellipse e4) than for Condition B (density increased for Ellipse e6). That is, the similarities involving a dense target ellipse should be smaller than for the same ellipse in its nondense condition.

Table 1
Rated Visual Similarity of Ellipses: Experiment 1

Ellipse	e1	e2	e3	e4	e5	e6	e7	e8	e9	B10	B11	B12
e1	—	4.04	1.59	6.33	4.44	2.30	5.00	4.78	3.33	2.70	3.81	2.67
e2	5.42	—	4.44	3.22	5.85	5.74	2.48	3.59	4.67	5.30	5.96	4.96
e3	2.71	4.63	—	1.37	2.89	6.04	1.52	2.11	3.93	7.85	6.63	6.70
e4	6.08	4.17	2.25	—	4.41	2.22	7.44	5.37	3.26	2.33	3.41	2.56
e5	5.38	6.50	3.67	4.63	—	4.56	3.93	6.22	5.96	3.96	5.04	4.67
e6	3.29	5.08	6.83	2.42	4.29	—	2.44	3.63	6.44	8.41	7.30	8.30
e7	3.08	2.33	1.54	5.33	2.79	1.63	—	5.04	2.89	1.26	1.59	1.44
e8	5.17	4.00	2.46	5.79	5.21	3.08	4.00	—	5.07	3.04	3.81	3.37
e9	3.50	5.38	5.17	2.96	4.88	5.92	1.79	4.25	—	6.48	6.19	6.74
A10	5.50	3.38	1.96	7.63	4.13	2.71	5.71	5.42	2.42	—	6.78	8.15
A11	6.38	4.13	2.25	7.42	5.25	3.33	4.71	6.42	2.88	6.88	—	7.63
A12	6.58	4.21	2.04	7.79	5.13	3.04	4.38	5.33	2.67	6.92	7.42	—

Note. Data above the diagonal is from Condition B (density increased for Ellipse e6). Data below the diagonal is from Condition A (density increased for Ellipse e4).

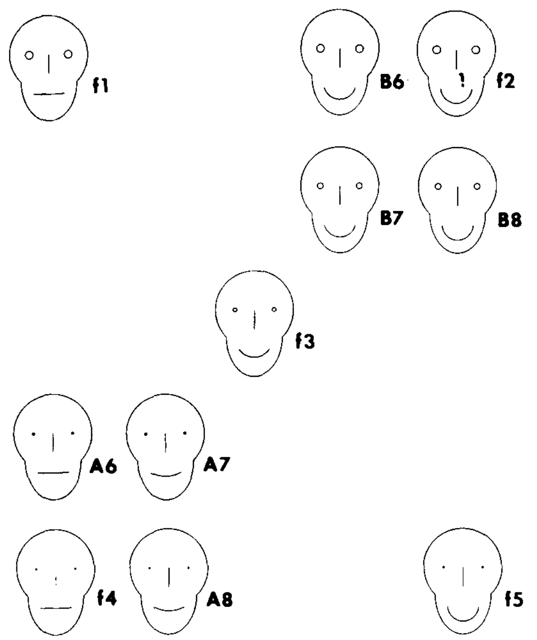

Figure 2. Face stimuli of Experiment 2. (Frame faces are designated f1–f5, context faces for Condition A by A6–A8, and context faces for Condition B by B6–B8.)

The difference in the value of the statistic for the two conditions ($\psi_A - \psi_B$) was calculated as 0.120, which is in the direction opposite to that predicted by the density hypothesis. The two-group t test for this comparison was not significant, $t(49) = 0.824, p > .05$.

One potential objection to this analysis is that the mean and variance of ratings may vary considerably from subject to subject. Although any differences between the groups in this regard should be small, given the sample size, it could be argued that such differences might be obscuring any density effect. Therefore, in a second analysis the 36 similarity ratings among the frame ellipses (i.e., for those stimulus pairs common to the two conditions) were separately normalized for each subject. These normalized ratings were then tested by the same statistic used in the previous analysis. The results did not change under this analysis. The difference $\psi_A - \psi_B$ was again positive, 0.368, $t(49) = 0.507, p > .05$.

Experiment 2

Experiment 2 was designed to test for density effects in ratings of the similarity of a set of schematic faces varying on eye size and curvature of mouth. The design of the experiment was like that of Experiment 1: Density was manipulated (in a between-subjects design), and the effect on the rated similarity of the faces was assessed.

Method

Subjects. A total of 49 Stanford undergraduates participated in the study as part of a course requirement.

Materials. Two sets of eight faces varying on mouth curvature and eye size were constructed. In each of the conditions five of the faces were frame stimuli, and three were added context faces. The two sets of faces used are shown in Figure 2. In Condition A the three context faces used were constructed to be highly similar to Face f4, while in Condition B they were highly similar to Face f2.

Procedure. The experimental procedure was the same as for Experiment 1. Subjects were first shown a booklet with the set of faces for their condition, and asked to familiarize themselves with the set. They then worked through a test booklet that contained one of the 36 pairs of stimuli on each page, assessing the visual similarity of each pair on a 9-point scale and recording their rating on a separate answer sheet.

Results

The average similarity ratings for the stimuli are given in Table 2. As in Experiment 1, the test of the density hypothesis was performed by comparing the value of a statistic involving the ratings for the two conditions. Because in Condition A Face f4 was the target stimulus (i.e., the face to which the added context faces were most similar), the density hypothesis predicts that the average similarity of this face to the other frame faces should be decreased, and analogously for Face f2 in Condition B. So the hypothesis may be tested by comparing, for the two conditions, the value of the test statistic

$$\psi = S_{2,}{}^{*} - S_{4,}{}^{*},$$

where

$$S_{2,}{}^{*} = S(f2, f1) + S(f2, f3) + S(f2, f5),$$

and

$$S_{4,}{}^{*} = S(f4, f1) + S(f4, f3) + S(f4, f5).$$

The difference in calculated values for the two conditions ($\psi_A - \psi_B$) is 0.342, which is not significant at the .05 level, $t(47) = 0.444$. For the z-score analysis, in which the ratings of frame–frame pairs were normalized for each subject before the statistic was calculated, the value of the comparison was 0.397, which is not significant at the .05 level, $t(47) = 0.819$. Again, the values of the comparisons were positive, rather than negative as predicted by the density hypothesis.

Experiment 3

Experiment 3 replicated the previous experiments using a set of letter-like figures constructed from line segments.

Table 2
Rated Visual Similarity of Faces: Experiment 2

Face	f1	f2	f3	f4	f5	B6	B7	B8
f1	—	5.20	3.68	5.36	2.28	5.04	3.92	3.80
f2	5.38	—	5.20	2.48	5.44	8.40	7.16	7.48
f3	4.04	5.33	—	3.32	6.04	6.00	6.48	5.92
f4	6.00	2.25	3.63	—	5.36	3.16	3.00	2.80
f5	2.29	5.63	5.29	5.25	—	5.40	5.72	5.92
A6	6.13	2.58	4.96	8.21	4.83	—	7.36	7.28
A7	3.88	3.50	6.08	6.08	5.04	7.25	—	8.56
A8	3.63	3.38	5.17	7.08	6.21	6.33	8.04	—

Note. Data above the diagonal is from Condition B (density increased for Face f2). Data below the diagonal is from Condition A (density increased for Face f4).

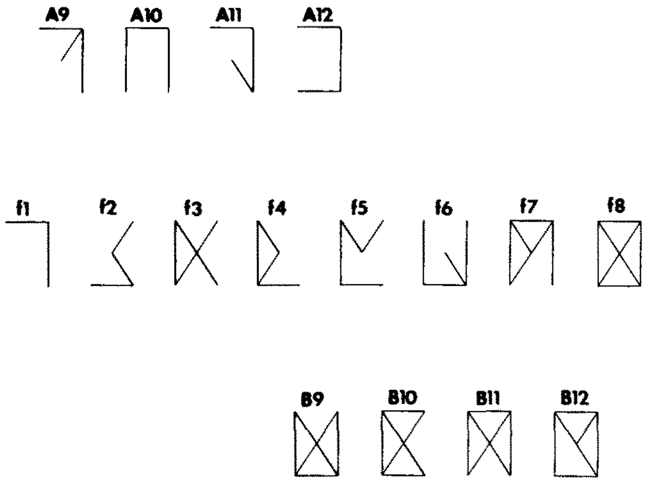

Figure 3. Figure stimuli used in Experiments 3 (similarity rating) and 4 (discrimination confusions). (Frame figures are f1–f8, context figures for Condition A are A9–A12, and context figures for Condition B are B9–B12.)

Method

Subjects. A total of 50 Stanford undergraduates participated in the study as part of a course requirement.

Materials. Two sets of 12 stimuli were constructed. There were 8 frame and 4 context figures in each condition, as is shown in Figure 3. In Condition A, the density was increased in the region of Figure f1 by the addition of the four stimuli A9, A10, A11, and A12 to the context set. In Condition B, the density was increased for Figure f8 by the addition of B9–B12.

Procedure. The procedure was the same as for the previous experiments. Subjects first looked through a training booklet, then worked through a randomized test booklet, making their ratings of similarity on a 9-point scale on a separate answer sheet.

Results

The matrices of similarity ratings are given in Table 3. The comparison used to test for a density effect was

$$\psi = S_1,^* - S_8,^*,$$

where

$$S_1,^* = S(f1,f2) + S(f1,f3) + S(f1,f4) + S(f1,f5) + S(f1,f6) + S(f1,f7)$$

and

$$S_8,^* = S(f8,f2) + S(f8,f3) + S(f8,f4) + S(f8,f5) + S(f8,f6) + S(f8,f7).$$

The difference in the values of the statistic for the two conditions ($\psi_A - \psi_B$) was 0.120, again positive rather than negative, and again not significant, $t(48) = 0.063$, $p > .05$. The *z*-score analysis also yielded a nonsignificant result, with a difference of 0.459, $t(48) = 0.418$.

Summary of Similarity-Rating Studies

The results of the three experiments were consistent. The predicted relation between local density and rated similarity was not found in any of the studies. That this is not merely due to inadequate power of the test is suggested by the fact that the values calculated for the comparisons were all slightly positive, contrary to the prediction of the density hypothesis. Also, the sample sizes used were comparable to those normally used in studies of rated similarity. Given these results, it seems safe to conclude that stimulus density does not affect the similarity of stimuli to any significant degree under the conditions of typical similarity rating studies.

Objections that local density for the target stimuli may not have been effectively manipulated can be dismissed by an examination of the similarity matrices. In each matrix, the nearest neighbors of that matrix's target stimulus (e.g., Ellipse e4 in Table 1, Condition A) are the added context stimuli. For example, the similarities of Ellipses A10, A11, and A12 to Ellipse e4 in Condition A are 7.63, 7.42, and 7.79. The next closest stimulus to Ellipse e4 in that condition is Ellipse e1, which has a similarity of 6.08 to Ellipse e4. Inspection of the other matrices reveals that the density manipulation was effective in all cases.

Table 3
Rated Visual Similarity of Letter-Like Figures: Experiment 3

Figure	f1	f2	f3	f4	f5	f6	f7	f8	B9	B10	B11	B12
f1	—	1.80	1.88	2.52	2.40	2.64	2.96	2.36	2.16	2.16	2.76	2.96
f2	2.72	—	2.40	3.36	3.36	2.40	2.44	2.96	2.96	4.08	3.48	2.80
f3	2.40	3.96	—	4.12	3.96	3.04	4.08	5.56	5.20	5.68	4.52	4.28
f4	3.48	4.16	4.84	—	3.88	2.96	4.00	3.96	4.36	4.12	3.56	3.40
f5	2.60	3.80	4.28	4.40	—	3.20	3.72	3.36	4.28	3.04	3.08	4.92
f6	3.72	3.52	3.24	3.64	3.64	—	3.76	3.72	3.80	3.68	2.84	4.04
f7	3.56	2.24	4.76	4.48	4.24	4.08	—	5.40	4.68	4.76	6.12	6.12
f8	3.16	3.12	5.80	4.40	4.56	4.60	6.64	—	7.96	7.68	8.04	6.28
A9	6.64	3.28	3.04	3.60	3.16	4.64	4.64	4.36	—	6.72	7.68	4.84
A10	6.60	2.84	2.88	2.88	3.68	5.48	4.68	4.36	5.16	—	6.24	5.12
A11	6.52	3.36	3.40	3.44	3.80	4.60	3.28	3.52	5.28	4.92	—	5.48
A12	6.92	3.12	2.52	3.56	3.44	4.32	4.00	4.40	4.68	5.36	5.44	—

Note. Data above the diagonal is from Condition B (density increased for Figure f8). Data below the diagonal is from Condition A (density increased for Figure f1).

Discrimination and Identification Studies

If density affects the perceived similarity of stimuli, then one might expect to observe consistent effects of density across a variety of proximity tasks. On the other hand, density might affect task-specific components of certain proximity tasks but not others. This might occur, for example, if density did not affect perceived similarity but somehow affected the selection of responses in a particular task. Because tasks such as rating the similarity of two concepts and discriminating or identifying them under degraded conditions might involve very different processes or response strategies, density might affect performance in some tasks but not others. Therefore, Experiments 4–6 were designed to test whether manipulating density experimentally has any effect on confusions between stimuli in discrimination and identification tasks.

Experiment 4

Experiment 4 was designed to test if density influences the probabilities of errors in a "same–different" discrimination task. The study used the same set of figures as Experiment 3.

Method

Subjects. A total of 52 Stanford undergraduates participated in this experiment as part of a course requirement.

Materials. The materials used were the figures used in Experiment 3, shown in Figure 3.

Procedure. Subjects were tested individually. Pairs of stimulus figures were presented on a Megatek 5000 graphics display screen controlled by a Data General Nova computer. The subject was seated before the screen in a dimly lit room. A button board in front of the subject held three keys: One was to initiate trials and the other two were used to record "same" and "different" responses. The subject could use a foot pedal instead of the middle key to initiate trials, if preferred.

The outline of a trial was as follows. A dot appeared in the middle of the screen, indicating that all was ready for the subject to initiate the trial. When the key was pressed, a pair of stimulus figures was flashed very briefly upon the screen and immediately followed by a mask. The subjects' task was to respond "same" if they thought the same figure had appeared on both sides and "different" if they thought that two different figures had appeared. For approximately half of the subjects in each condition, "same" corresponded to the right key, and for the other half to the left key.

Before the actual experiment began, subjects were given two types of training tasks. First, the nature of the same–different task was explained to them, and they were given 70 warm-up trials to familiarize them with the task and use of the response keys. In these practice trials the stimulus set consisted of the plus (+) and minus (−) symbols. The presentation speed was increased every 10 trials during this training phase until the subject was making a significant proportion of errors (greater than 15%). Feedback on correctness was given only during this training phase.

Next, the subject was familiarized with the stimulus set to be used in the actual experiment. A single figure was presented on the screen, and the subject could view the next in the series by pressing the middle key. At the end of the series the first was again presented, so that the subject could go through the set as many times as desired. This procedure was meant to be an electronic equivalent of the training booklet used in the similarity-rating studies. Subjects were told to take approximately 5 min and study the set of figures to be used in the same–different task until they felt confident they could pass a recognition test for the figures in the set. This instruction was given to increase attention to the figures, presumably making the composition of the context set more salient. No recognition test was actually given.

When the subject expressed confidence that he or she knew the set of materials, the experimental task began. The paradigm was the same–different task described earlier. A subject was given two blocks of trials, each consisting of 114 trials: one complete set of the 66 "different" pairs, mixed with 48 "same" pairs. Pairs were presented in random order. A presentation speed was selected on the basis of subject's performance in the training task, with the aim of adjusting the difficulty of the task to result in between 10% and 40% errors. This speed was adjusted between blocks if necessary. Data from two subjects who performed at chance levels on "different" pairs were excluded from the analysis, as were subjects who did not complete both blocks for any reason.

Results

Table 4 shows the average percentage of confusions for stimulus pairs in both conditions (the overall percentage of error for "different" pairs was 25.9% for Condition A, 27.4% for Condition B.)

The percentage of error for a "different" pair (that is, the proportion of times that subjects responded "same" when figures x and y were presented) is a measure of the similarity or confusability of x and y and is denoted $C(x, y)$. According to the density hypothesis, in Condition A the confusability of Figure f1 with all other frame figures should be decreased by the addition of the four highly similar figures. Analogously, in Condition B the confusability of Figure f6 with the other frame figures should be decreased. Accordingly, the same statistic used to test the density hypothesis in Experiment 3 can be used for the test of the present confusion data, substituting the proportions of confusion errors $C(x, y)$ for the similarity ratings $S(x, y)$. For purposes of the analysis, data from separate blocks were treated as independent. Just as in the rating studies, the density hypothesis predicts a negative value for the comparison. The actual value of the comparison was 0.310, which is in the direction opposite to that predicted by the density hypothesis, and is not significant, $t(76) = 1.58$, $p > .05$.

Examination of Table 4 reveals that density was effectively manipulated in this paradigm. The four context figures of Condition A (A9, A10, A11, A12) were the most often confused with the target Figure f4, whereas in Condition B only Figure f7 was confused with the target stimulus f8 as often as one of the added context figures.

Experiment 5

For familiar stimuli, such as letters, density can be manipulated in between-subjects designs by selecting different subsets of the population of stimuli. Experiments 5 and 6 examined the effects of this type of density manipulation on letter confusions in a discrimination task and an identification task.

Method

Subjects. 24 subjects participated in Experiments 5 and 6, receiving either course credit or $4 for their participation.

Materials. The materials used in the experiments were two sets of six uppercase block letters, constructed of line segments. For Condition 1

Table 4
Discrimination Confusions Between Letter-Like Figures and Proportion Correct (PC): Experiment 4

Figure	f1	f2	f3	f4	f5	f6	f7	f8	B9	B10	B11	B12	PC
f1	—	0.111	0.028	0.028	0.111	0.028	0.139	0.083	0.000	0.139	0.083	0.111	0.882
f2	0.095	—	0.139	0.167	0.167	0.111	0.000	0.139	0.056	0.056	0.083	0.000	0.889
f3	0.095	0.048	—	0.250	0.278	0.111	0.306	0.167	0.389	0.333	0.333	0.333	0.583
f4	0.024	0.214	0.286	—	0.306	0.167	0.111	0.111	0.194	0.056	0.083	0.167	0.799
f5	0.119	0.143	0.262	0.214	—	0.194	0.167	0.111	0.111	0.056	0.194	0.306	0.819
f6	0.048	0.167	0.143	0.095	0.095	—	0.139	0.167	0.194	0.056	0.250	0.250	0.882
f7	0.167	0.024	0.286	0.190	0.405	0.119	—	0.472	0.139	0.639	0.250	0.528	0.639
f8	0.024	0.000	0.143	0.167	0.143	0.071	0.381	—	0.361	0.583	0.444	0.639	0.688
A9	0.286	0.024	0.190	0.024	0.071	0.119	0.190	0.214	—	0.333	0.472	0.500	0.521
A10	0.333	0.167	0.262	0.167	0.214	0.190	0.286	0.143	0.190	—	0.389	0.389	0.736
A11	0.190	0.143	0.286	0.095	0.119	0.524	0.167	0.119	0.167	0.357	—	0.500	0.660
A12	0.476	0.143	0.095	0.071	0.286	0.071	0.095	0.143	0.333	0.167	0.119	—	0.611
PC	0.744	0.792	0.607	0.732	0.857	0.792	0.571	0.696	0.839	0.827	0.821	0.613	

Note. Data above the diagonal is from Condition B (density increased for Figure f8). Data below the diagonal is from Condition A (density increased for Figure f1).

the set used was {C, H, O, E, F, P}; for Condition 2 it was {C, H, O, E, U, Q}. Both sets of letters are shown in Figure 4. The context stimuli (F and P) included in Condition 1 were chosen to increase local density for the target letter E, whereas in Condition 2 Q and U were included to increase the local density for O.

Procedure. Each subject participated in both Experiment 5 and 6, with half being given the identification task first, and half being given the discrimination task first. A subject saw the same context set of letters in both experiments. The procedure was similar to that of Experiment 4. Subjects were tested individually in a dimly lit room, and the materials were presented on a Megatek 5000 graphics display screen controlled by a Data General Nova computer. The subject initiated trials with a foot pedal and then responded to the stimulus pair by pressing either the key marked "same" or the one marked "different."

Before the actual experiment began, subjects were shown the set of six letters for their condition, and were instructed to look through the set to familiarize themselves with the letters. For practice with the same–different task, they were given 20 training trials at a 100-ms presentation time. The stimuli used in this training task were the same letters used in the actual experiment. When they had completed these training trials, the testing phase began. A subject was presented with four blocks of trials, each consisting of 120 trials: 60 "different" and 60 "same" pairs. Pairs were presented in a separate random order for each block. The presentation time for stimulus pairs was initially set at 40 ms, but

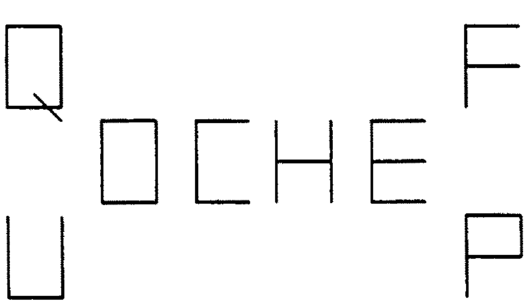

Figure 4. Letter stimuli used in Experiments 5 and 6. (The letters C, H, E, and O were common to both conditions, in Condition 1 the context letters F and P were also included, and in Condition 2 the letters Q and U were included.)

the program controlling the experiment adjusted the presentation time every 10 trials during the first block to arrive at a presentation time that would result in a moderate number of errors (data from this first block were not analyzed).

In a single trial, a subject initiated the trial with the foot pedal, and the stimulus pair appeared briefly, followed by a masking pattern. Simultaneous with the appearance of the poststimulus masking pattern, the set of six letters in the stimulus set appeared, randomly ordered, at the bottom of the subject's screen. This was done both to ensure that the context set remained salient for the subjects and to maximize comparability of the procedure with that of Experiment 6.

Results

Data from the first block were discarded, and the data from the last three blocks were analyzed. The analysis examined the average proportion of times a given pair of letters was confused (i.e., false "same" responses). According to the density hypothesis, in Condition 1 the confusability of the letter E with all other frame letters should be decreased by the addition of the four highly similar letters. Analogously, in Condition 2 the confusability of letter O with the other frame letters should be decreased.

Table 5 shows the confusion matrices for both conditions. The comparison used to test for a density effect was

$$\psi = C_{E,}{}^{*} - C_{O,}{}^{*} = C(E, C) + C(E, H) - C(O, C) - C(O, H),$$

where $C(E, C)$ refers to the proportion of confusion errors (false "same" responses) between the letter E and the letter C. The difference in this measure between conditions was −0.100, which is in the direction predicted by the density hypothesis but is not significant, $t(58) = -0.264, p > .05$.

Experiment 6

Experiment 6 was an identification task using the letter stimuli of Experiment 5. The subjects and materials used were the same as for that experiment.

Table 5
Discrimination Confusions Between Letters and Proportion Correct (PC): Experiment 5

Letter	C	H	E	O	F	P	PC
C	—	.284	.334	.284	.134	.217	.513
H	.300	—	.267	.167	.000	.284	.860
E	.400	.184	—	.050	.084	.100	.593
O	.300	.184	.200	—	.367	.267	.607
Q	.216	.167	.184	.084	—	.267	.820
U	.150	.167	.200	.084	.284	—	.793
PC	.353	.880	.507	.793	.640	.673	—

Note. Data above the diagonal is from Condition 2 (density increased for letter O). Data below the diagonal is from Condition 1 (density increased for letter E).

Method

In this procedure subjects gave their identification responses by means of six response keys. When the stimulus had been presented and while a masking pattern remained on, the set of six letters that were used in that condition appeared across the bottom of the screen, and remained until the subject chose a response. This array served both as a reminder of the context set and as a means of identifying the response keys because the order of letters in the array (randomly assigned for each subject) corresponded to the order of the response keys. For example, if the letter H appeared as the leftmost element of the array, the correct response key for H was the leftmost key.

As in Experiment 5, subjects were shown the set of six letters for their condition and asked to familiarize themselves with the stimulus set. Next they were given 20 training trials at a 100-ms presentation time. When they had completed these training trials, the testing phase began. A subject was presented with four blocks of trials, each consisting of 60 letter presentations. Letters appeared in separate random orders in each block. As in the discrimination task, the presentation time for a stimulus was initially set at 40 ms and adjusted every 10 trials during the first block. Only the data from the last three blocks were analyzed.

Results

Table 6 summarizes the identification errors made for each letter. In order to facilitate comparison with the results of Experiment 5, the identification errors were first analyzed as symmetrized confusions. That is, the two halves of the identification error matrix were averaged to obtain a lower-half confusion matrix. For example, the number of times E was given as the response to H, $P_H(E)$, was averaged with the number of times H was given as the response to E, $P_E(H)$, to obtain a measure of confusability, $C(E, H)$, between the two letters. The density hypothesis may then be tested for this data by the same comparison as for Experiment 5. The difference in the values of the comparison for the two conditions was -1.939, which is in the direction specified by the density hypothesis and is significant, $t(61) = -1.884, p < .05$. This means that the confusability of E with H and C was increased by adding letters (F and P) to the set that were highly similar to E.

Analyzing identification errors as symmetrized confusions does not distinguish between the effects of density of the stimulus (presented) letter and of the response letter. Under the focusing hypothesis discussed in the introduction, the density of one member of the stimulus pair may be more heavily weighted. Because only one letter is presented in an identification task, it seems clear that the density of this letter should be the salient factor. Accordingly, the probability of an error should be affected by the density in the region of the stimulus letter, but not necessarily by the density in the region of the response letter. It is possible to test separately for a stimulus density effect and a response density effect. The stimulus effect can be tested by the statistic

$$\psi_S = P_E(C) + P_E(H) + P_E(O) - P_O(C) - P_O(H) - P_O(E).$$

Note that this statistic includes a comparison of $P_E(O)$ and $P_O(E)$. The difference in the statistic for the two conditions was -0.009, which is not significant, $t(61) = -0.0131, p < .05$.

Similarly, a density effect on the response letter can be tested by

$$\psi_R = P_C(E) + P_H(E) + P_O(E) - P_C(O) - P_H(O) - P_E(O).$$

The difference in value for this statistic was -1.9303, $t(61) = -1.7846$, $p < .05$. Thus, a significant effect of density is obtained, but it is an effect of density in the region of the response letter. The choice probabilities are not affected by the density in the region of the stimulus letter, as predicted by the density hypothesis in conjunction with the focusing hypothesis.

Summary of Discrimination and Identification Studies

Experiments 4–6 examined the influence of stimulus density on confusion errors in perceptual tasks involving letter-like figures and actual letters. For the discrimination tasks (Experiments 4 and 5), there was no significant effect of density on the confusability of stimulus pairs. For the identification task of Experiment 6, no effect of the density of stimuli in the region of the stimulus letter was found, but a *response density* effect was observed. This pattern of results suggests that density does

Table 6
Identification Errors for Letter Stimuli: Experiment 6

Letter	C	H	E	O	F	P
C	.327	.090	.287	.133	.120	.043
H	.007	.803	.070	.013	.063	.043
E	.063	.107	.563	.037	.150	.080
O	.027	.120	.003	.827	.010	.013
F	.027	.090	.167	.037	.570	.110
P	.010	.057	.027	.030	.097	.780
	C	H	E	O	Q	U
C	.412	.027	.397	.061	.048	.055
H	.015	.773	.076	.045	.018	.073
E	.067	.103	.788	.018	.003	.021
O	.036	.073	.021	.403	.127	.339
Q	.003	.009	.000	.000	.964	.024
U	.012	.067	.000	.070	.030	.821

Note. Data in upper part of table is from Condition 1 (density increased for letter E). Data in lower part is from Condition 2 (density increased for letter O).

not influence the perceived similarity of stimulus pairs, although it may influence the process of selecting a response in identification tasks.

Discussion

One way in which Krumhansl (1978) motivated the distance–density model was to point out its relation to range–frequency theory (Birnbaum, 1974; Parducci, 1965, 1974), which describes certain context effects in unidimensional judgment tasks. In range–frequency theory; it is assumed that the spacing of stimuli along the continuum can affect the rating given to a single stimulus. In particular, subjects seem to make these rating category assignments so as to equalize the number of stimuli assigned to each category. This means that if there is a high frequency of stimuli with nearly equal magnitude somewhere along the continuum, stimuli near the high end of this dense region will receive higher ratings than they would otherwise, and stimuli near the low end will receive lower ratings. In a sense, the data show a "spreading out" of the dense region. Krumhansl noted that if such a spreading out of dense regions occurs in the perception of similarity relations, this effect would show up in the proximity data as a density effect. However, it has since been demonstrated (Mellers & Birnbaum, 1982) that the context effects observed in unidimensional judgment are not due to changes in subjects' perception of the stimuli, but rather to a tendency by them to use different response categories more or less equally often. That is, the frequency effect is a response effect, not a stimulus effect. Thus, the range–frequency findings offer no reason to expect an effect of density on perceived similarity.

The pattern of results observed in the present studies supports this view. If density had an influence on perceived similarity, then density effects should be observed consistently across the different tasks. This was certainly not the case: No effect of density was observed for the similarity rating and discrimination tasks. In the identification experiment, in which the tests of a stimulus effect and a response effect can be separated, only the density of the response letter seemed to have an influence. To understand possible mechanisms for response-density effects in identification data, it is necessary to consider the role of a choice process in stimulus identification.

Models of Identification Performance

Most models of letter identification have explicitly recognized the existence of a choice process operating in the selection of a response in such a task. For example, Townsend and Landon (1983) discussed five quantitative models of letter identification and their application to letter recognition data: the constant-ratio rule (CRR; Clarke, 1957; Egan, 1957; Luce, 1959; Shepard, 1957), Luce's (1963) similarity-choice model, Townsend's overlap (1971) and all-or-none (1978) models, and Nakatani's (1972) confusion-choice model. The CRR was originally proposed as a general model of individual choice behavior, the others specifically as models of stimulus identification. Each of these last four models, however, incorporates an explicit choice component based on the CRR. Takane and Shibayama (1986) also compare several models in which the CRR is used to represent the choice component of stimulus identification.

In the models of Townsend (1971, 1978) and Nakatani (1972), the process of stimulus identification is explicitly separated into two sequential subprocesses: stimulus letter perception and response letter selection. Following Nakatani, such models will be referred to as confusion-choice models. In these models, recognition confusions are presumed to be the result of an errorful perceptual process followed by a probabilistic choice mechanism. The errorful recognition process has been variously assumed to result in all-or-none recognition (Townsend, 1971), in a set of pair-wise confusion states (Townsend, 1978), or in a multivariate normal distribution representing the stimulus in a multidimensional stimulus space (Nakatani, 1972). Tests by Townsend and Landon (1982) indicated that Nakatani's gave the best fit of the three confusion-choice models.

By separating the recognition process into sequential processes of stimulus perception and response selection, a confusion-choice model offers a natural account of how response density effects might arise. In such a model, a stimulus is presented, resulting (on a particular trial) in some percept. Given this percept, the subject must choose the best or most likely response from the set of valid alternatives. In order to understand how the composition of the stimulus set, particularly local stimulus density, might affect this response-selection phase, it is useful to review certain results concerning the effect of stimulus-set context on choice probabilities.

Choice Probabilities and Context Effects

Two well-known properties of choice data are particularly relevant to understanding how density might affect response probabilities in certain identification experiments. The first is *regularity.* Regularity requires that adding members to a choice set increase (or leave unchanged) the probability of choosing any one of its members. That is, for choice sets A and B, $A \subset B$, the probability of choosing x from A must be greater than or equal to the probability of choosing x from B: $P(x, A) \geq P(x, B)$. Regularity is a property derivable from most choice models and is usually found to be satisfied in actual choice data (e.g., Becker, DeGroot, & Marshak, 1963; Tversky, 1972).

What is the relation between the regularity property and potential response density effects in identification data? Consider two identification experiments. In the first a particular set of frame stimuli is used. In the second the same stimuli are included, but the density is increased for one target stimulus by the addition to the set of several highly similar context stimuli. As applied to the identification data, regularity requires that the probability of selecting any stimulus as a response cannot be increased by the addition of the context stimuli; it must either decrease or stay the same. In essence, regularity places a directional constraint on any effect of stimulus density achieved by simply adding alternatives to the set, and this constraint is consistent with the directional prediction of the density hypothesis. Note that regularity does not predict that density necessarily affects response probabilities, only that if the effect occurs it must be in a particular direction.

However, another well-established property of choice probabilities does predict a response-density effect in identification tasks. For convenience, this property will be referred to as the *similarity* effect. It was first noted in the well-known counterexample of Debreu (1960; see also Tversky, 1972). The import of these counterexamples is that the more similar members of the set of alternatives are to x, the lower will be the probability of choosing x relative to the probability of choosing any other stimulus y. Such a pattern of choice probabilities violates the property known as independence from irrelevant alternatives. Numerous studies document such violations in choice probabilities (Becker et al., 1963; Debreu, 1960; Townsend & Landon, 1982; Tversky & Russo, 1969).

The similarity effect predicts changes in choice probabilities with changes in local stimulus density. Accordingly, changes would also be expected in the response probabilities of an identification task. Stated loosely, the similarity effect is that the addition of an alternative to a choice set "hurts" similar alternatives more than dissimilar ones. That is, the more similar an added alternative is to stimulus x, the lower will be the probability of selecting response x (no matter what letter is actually presented). This is exactly the effect observed in the identification data of Experiment 6. Adding context letters to increase the density for the letter C had the effect of decreasing the overall probability of responding "C". Similarly, increasing the density for O decreased the overall probability of "O" as a response.

Thus, the assumption of standard models of stimulus identification that identification performance incorporates a choice mechanism for response selection offers an explanation for the finding that density can affect response probabilities. Therefore, the results of the experiments reported here can be seen as consistent with the general outline of these models. However, specific assumptions used in such models may be called into question by the present data. In particular, the specific choice rule used by the models reviewed above, the CRR, is unable to account for the similarity effect in choice data.

Less restrictive choice models can, however, account for the similarity effect. Consequently, it seems useful to explore the usefulness of variants of the confusion-choice model incorporating more sophisticated choice models, such as Tversky's (1972) elimination-by-aspects model. One such discrete-feature version of a confusion-choice model is described in Corter (1987). Another model for choice data that can account for the similarity effect, and that is especially noteworthy in the present application, is the wandering ideal point model (De Soete, Carroll, & DeSarbo, 1986). In this model, a subject is represented as a point in a multidimensional stimulus space. This point is termed the *ideal point* for the subject, and is assumed to follow a multivariate normal distribution across trials. On a particular trial, the subject chooses the stimulus closest to the sampled ideal point. This model, proposed to account for pair-wise choice data, could be generalized in a natural way to identification data. In such an extension each stimulus letter (rather than each subject) would be represented by a multivariate normal ideal point distribution. The presentation of a stimulus letter would sample a point from the corresponding distribution, and the response chosen would be the stimulus with centroid closest to the sampled point. Note that as a psychological account of identification performance this extension of the wandering ideal point model would closely resemble Nakatani's model, although the models would differ in certain specific details.

References

Appelman, I. B., & Mayzner, M. S. (1982). Application of geometric models to letter recognition: Distance and density. *Journal of Experimental Psychology: General, 111,* 60–100.

Balzano, G. J. (1977). *Chronometric studies of the musical interval sense.* Unpublished doctoral dissertation, Stanford University.

Becker, G. M., DeGroot, M. H., & Marshak, J. (1963). Probabilities of choices among very similar objects. *Behavioral Science, 8,* 306–3111.

Birnbaum, M. H. (1974). Using contextual effects to derive psychophysical scales. *Perception & Psychophysics, 15,* 89–96.

Clarke, F. R. (1957). Constant-ratio rule for confusion matrices in speech communication. *Journal of the Acoustical Society of America, 29,* 715–720.

Corter, J. E. (1987). *A discrete-feature confusion-choice model for stimulus identification.* Unpublished manuscript, Teachers College, Columbia University.

Debreu, G. E. (1960). Review of R. D. Luce, *Individual choice behavior: A theoretical analysis. American Economic Review, 50,* 186–188.

De Soete, G. E., Carroll, J. D., & DeSarbo, W. S. (1986). The wandering ideal point model: A probabilistic multidimensional unfolding model for paired comparisons data. *Journal of Mathematical Psychology, 30,* 28–41.

Egan, J. P. (1957). Monitoring task in speech communication. *Journal of the Acoustical Society of America, 30,* 715–720.

Gilmore, G. C., Hersh, H., Caramazza, A., & Griffin, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics, 25,* 425–431.

Keren, G., & Baggen, S. (1981). Recognition models of alphanumeric characters. *Perception & Psychophysics, 29,* 234–246.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review, 85,* 445–463.

Krumhansl, C. L. (1982). Density versus feature weights as predictors of visual identifications: Comment on Appelman and Mayzner. *Journal of Experimental Psychology: General, 111,* 101–108.

Luce, R. D. (1959). *Individual choice behavior.* New York: Wiley.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology,* (Vol. 1, pp. 103–189). New York: Wiley.

Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 582–601.

Nakatani, L. H. (1972). Confusion-choice model for multidimensional psychophysics. *Journal of Mathematical Psychology, 9,* 104–127.

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72,* 407–418.

Parducci, A. (1974). Contextual effects: A range-frequency analysis. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 127–172). New York: Academic Press.

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology, 53,* 94–108.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika, 22,* 325–345.

Shepard, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function. Part I. *Psychometrika, 27,* 125–140.

Shepard, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function. Part II. *Psychometrika, 27,* 219–246.

Takane, Y., & Shibayama, T. (1986). Comparison of models for stimulus recognition data. In DeLeeuw, J., Heiser, W. J., Meulman, J., & Critchley, F. (Eds.), *Multidimensional data analysis* (pp. 119–148). Leiden: DSWO Press.

Torgerson, W. S. (1958). Theory and methods of scaling. New York: Wiley.

Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics, 9,* 40–50.

Townsend, J. T. (1978). A clarification of some current multiplicative confusion models. *Journal of Mathematical Psychology, 18,* 25–38.

Townsend, J. T., & Landon, D. E. (1983). Mathematical models of recognition and confusion in psychology. *International Journal of Mathematical Social Sciences, 4,* 25–71.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281–299.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. B. Lloyd, (Eds.), *Cognition and categorization.* Hillsdale, NJ: Erlbaum.

Tversky, A., & Russo, J. E. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology, 6,* 1–12.

Received August 12, 1986
Revision received January 14, 1987
Accepted February 5, 1987 ■

1988 APA Convention "Call for Programs"

The APA "Call for Programs" for the 1988 annual convention will appear in the October issue of the *APA Monitor.* The 1988 convention will be in Atlanta, Georgia, from August 12 to August 16. Deadline for submission of programs and papers is December 21, 1987. This early deadline is required because the 1988 convention is earlier in August than in the past. Additional copies of the "Call" will be available from the APA Convention Office in October.