

Similarity, Frequency, and Category Representations

Robert M. Nosofsky
Indiana University

This article studies the joint roles of similarity and frequency in determining graded category structure. Perceptual classification learning experiments were conducted in which presentation frequencies of individual exemplars were manipulated. The exemplars had varying degrees of similarity to members of the target and contrast categories. Classification accuracy and typicality ratings increased for exemplars presented with high frequency and for members of the target category that were similar to the high-frequency exemplars. Typicality decreased for members of the contrast category that were similar to the high-frequency exemplars. A frequency-sensitive similarity-to-exemplars model provided a good quantitative account of the classification learning and typicality data. The interactive relations among similarity, frequency, and categorization are considered in the General Discussion.

Among the most well-established findings in the categorization literature is that categories have "graded structures" (Rips, Schoben, & Smith, 1973; Rosch, 1973, 1978; Rosch & Mervis, 1975; Smith & Medin, 1981). Rather than all instances of a category being "equal," it appears that certain instances are better examples than others. For example, people reliably rate a robin as a better example of the category *birds* than they rate a penguin. Various experimental operations converge on the view that categories have graded structures, including typicality ratings, errors in classification learning, reaction time in speeded classification, and exemplar production.

Why are some instances of a category better examples than others? There has been widespread agreement since the work of Rosch and Mervis (1975) that a major determinant of graded category structure involves stimulus similarity. Generally speaking, the more similar an instance is to the other members of its category and the less similar it is to members of contrast categories, the higher will be the typicality rating given to that instance. So, for example, whereas robins are highly similar to numerous other instances of the category *birds*, penguins are relatively dissimilar to other *bird* instances.

Another variable that may play an important role in determining graded category structure is stimulus frequency. It seems plausible that as the frequency with which a person experiences an instance as an example of a category increases, the "goodness" of that instance as an example of the category will also increase. Thus, robins may be rated as highly typical birds because they are frequently experienced examples of the category *birds*.

In a seminal study investigating the determinants of graded category structure, Rosch, Simpson, and Miller (1976, Exper-

iment 2) conducted a category learning condition in which frequency was inversely related to similarity. Stimuli that were highly similar to other members of the category were presented less frequently than were stimuli that were relatively dissimilar to other members of the category. In transfer tests, the low-frequency, high-similarity items were rated as better examples of the category than were the high-frequency, low-similarity items. This result pointed clearly to the importance of similarity structure in determining typicality, and led Rosch et al. to question the importance of frequency information. Nevertheless, Rosch et al. acknowledged in their General Discussion, "Of course, frequency (repetition) is a powerful effect in learning. If the structural relation between items were held constant in the present experimental design, and frequency alone were to be varied, there is no doubt that frequency effects similar to our typicality effects could have been demonstrated" (Rosch et al., 1976, p. 501).

More recent research suggests that some form of familiarity or frequency may indeed play an important role in determining graded category structure.¹ Correlational studies using natural category terms indicate that familiar exemplars are judged as being more typical than unfamiliar ones (Ashcraft, 1978; Barsalou, 1985; Hampton & Gardner, 1983; Malt & Smith, 1982; Schwanenflugel & Rey, 1986). Barsalou (1985, p. 631) suggested and provided evidence for the idea that the relevant variable is not overall familiarity but rather *frequency of instantiation*, which he defined as "... someone's subjective estimate of how often they have experienced an entity as a member of a *particular category*." For example, although people are more familiar overall with chairs than with logs, this greater overall familiarity would not lead people to judge chairs as better examples of the category *firewood* than are logs. Logs are probably experienced more frequently than are chairs as examples of the category *firewood*.

¹ Although I use the terms *familiarity* and *frequency* here interchangeably, there are clearly important distinctions between the two constructs. For example, low-frequency exemplars may still be judged as highly familiar if their memory traces are highly "available" (Tversky & Kahneman, 1973). It also seems important to distinguish between "subjective" frequency and familiarity.

This work was supported by National Science Foundation Grant BNS 85-19573 to Indiana University.

I would like to thank Lawrence Barsalou, Lyle Bourne, Jr., William Estes, and James Hampton for their comments and criticisms regarding earlier versions of this article.

Correspondence concerning this article should be addressed to Robert M. Nosofsky, Department of Psychology, Indiana University, Psychology Building, Bloomington, Indiana 47405.

Effects of exemplar frequency on graded category structure have also been observed in classification learning experiments. Knapp and Anderson (1984) had subjects learn three categories of differing size (1, 6, and 24 unique exemplars). Individual exemplars were presented with frequency inversely proportional to the size of their respective categories. Learning was best for the highest frequency exemplars, that is, those exemplars from the smallest category, and worst for the lowest frequency exemplars. Because all exemplars within any given category were presented with equal frequency, however, Knapp and Anderson's experiment was not designed to detect changes in the local, graded structure of categories. Estes (1986b) found increased classification accuracy for individual, probabilistically generated exemplars that were repeated over the course of category learning. However, differential frequency was not explicitly manipulated in his study and may have been partially confounded with other variables. And in a classification learning study that comes closest to the present one, Barsalou (1981, Experiment 4) manipulated individual exemplar frequencies and similarity relations as orthogonal variables and found clear effects of both on subjects' postacquisition typicality ratings.

The present article adopts as a working hypothesis the idea that both similarity structure and frequency are determinants of graded category structure. The empirical goal is to study the manner in which these variables interact by manipulating the frequency with which items of varying degrees of similarity are presented during category learning. The experiments will allow for a more detailed and controlled study of the role of individual exemplar frequencies and similarities than has been possible with previously published work. The theoretical goal of the study involves the testing of a quantitative model within which to interpret the joint roles of similarity and frequency information as reflected in people's category representations.

Exemplar-Based Representations of Similarity and Frequency

The proposed model for interpreting the joint roles of similarity and frequency information is a version of the context model of classification developed by Medin and Schaffer (1978) and generalized by Nosofsky (1984, 1986). According to the context model, people represent categories by storing individual exemplars in memory. Classification decisions are based on similarity comparisons to the stored exemplars. Formally, the model states that the probability of classifying stimulus i into Category 1, $P(R_1 | S_i)$, is given by

$$P(R_1 | S_i) = \frac{b_1 \sum_{j \in C_1} \eta_{ij}}{b_1 \sum_{j \in C_1} \eta_{ij} + (1 - b_1) \sum_{k \in C_2} \eta_{ik}}, \quad (1)$$

where b_1 ($0 \leq b_1 \leq 1$) represents the bias for making Category Response 1, and where η_{ij} denotes the similarity between exemplars i and j . The expression in the numerator of Equation 1 is the bias for making Category Response 1 multiplied by the sum of similarities of stimulus i to all exemplars j belonging to Category 1. This expression gives the "strength" of making a Category 1 response given presentation of stimulus i . This strength is then divided by the sum of strengths

for both categories (assuming a two-category experiment) to determine the categorization probability.

Two alternative interpretations of the context model need to be distinguished. By one interpretation, the summing of similarities in Equation 1 takes place over all *distinct* exemplars in a category: The exemplars are viewed as "types," with multiple presentations of the same exemplar giving rise to a single representation in memory. A second interpretation is that the summing of similarities takes place over the complete set of presented exemplars: The exemplars are viewed as "tokens," with multiple presentations of the same exemplar leading to multiple representations in memory. The "type" interpretation is a frequency-insensitive model, whereas the "token" interpretation is frequency sensitive. To make explicit these alternatives in this article, Equation 1 will be understood to represent the frequency-insensitive model. Assuming that each exemplar presentation results in a unique memory trace (e.g., Hintzman & Block, 1971), the frequency-sensitive version is formalized as

$$P(R_1 | S_i) = \frac{b_1 \sum_{j \in C_1} N_j \eta_{ij}}{b_1 \sum_{j \in C_1} N_j \eta_{ij} + (1 - b_1) \sum_{k \in C_2} N_k \eta_{ik}}, \quad (2)$$

where N_j is the relative frequency with which exemplar j is presented during training. Estes (1986a, 1986b) assumed this frequency-sensitive version in recent theoretical and empirical work comparing the context model and feature-frequency models. The frequency-sensitive context model serves to guide the present research.

In summary, the working hypotheses motivating this research are that similarity and frequency jointly determine graded category structure, and that the manner in which these variables interact may be interpreted in terms of the frequency-sensitive context model. The conceptual underpinning of this approach is that people make classification decisions on the basis of similarity comparisons to stored exemplars. Furthermore, frequency information is represented naturally in the model in terms of the differential frequency with which individual exemplar traces are stored in memory. In the experiments to be reported, these hypotheses are tested by conducting category learning conditions in which presentation frequency of individual exemplars is manipulated. Attempts are made to provide quantitative accounts of the category learning data and of typicality rating data obtained in postacquisition transfer tests.

Experiment 1

In this experiment, subjects learned to classify Munsell colors varying along the dimensions of brightness and saturation into two categories. The locations of the stimuli in the color space and the category structure that was used are illustrated in Figure 1. The two-dimensional scaling solution that is shown was derived on the basis of pair-wise confusion errors observed during an identification learning condition (Nosofsky, 1987).

The categories divide roughly into colors that are "pinkish" (Category 2) and colors that are "brownish" (Category 1). In addition to the categories having a fairly natural description, the dimensional organization of the stimuli is simple enough

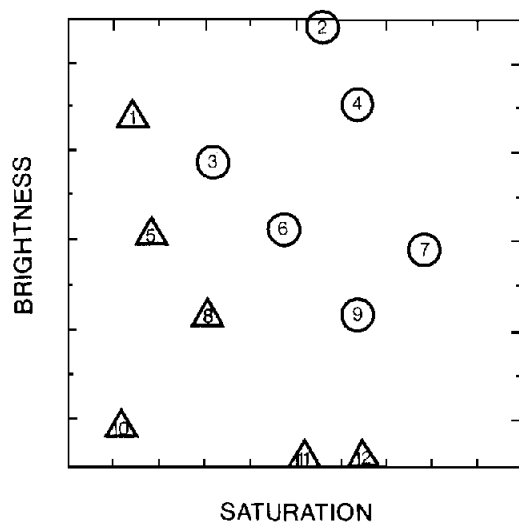


Figure 1. Category structure tested in Experiments 1 and 2. (Stimuli enclosed by triangles = members of Category 1; Stimuli enclosed by circles = members of Category 2)

to allow for a precise quantification of similarity relations. Furthermore, there appear to be varying degrees of category "goodness" associated with the individual exemplars. For example, Colors 2, 4, and 7 clearly are good exemplars of Category 2, whereas Colors 3, 6, and 9 are relatively poor exemplars, lying close to the category boundary.

In Experiment 1 presentation frequencies of good Exemplars 2 and 7 were manipulated. Condition B was a baseline condition in which all exemplars were presented with equal frequency, whereas in Condition E2 (E7), Exemplar 2 (7) was presented approximately five times as often as each of the other exemplars.

If differential frequency information is reflected in people's category representations in the manner suggested earlier, then the first main prediction is that classification accuracy and typicality ratings for the manipulated exemplars should increase with their presentation frequency. A second main prediction stemming from the frequency-sensitive exemplar model is that classification accuracy and typicality ratings should also increase for members of Category 2 that are very similar to the high-frequency exemplars. As can be seen in Figure 1, Exemplar 4 is a close neighbor of Exemplar 2, and Exemplar 9 is a close neighbor of Exemplar 7. The expectation is that category "goodness" will increase primarily for Exemplars 2 and 4 in Condition E2 and for Exemplars 7 and 9 in Condition E7.

Method

Subjects

150 subjects were hired for participation in the experiment. Most subjects were undergraduates at Indiana University. 50 subjects participated in Conditions B, E2, and E7, respectively.

Stimuli

The stimuli were 12 color chips manufactured by the Munsell color company. According to the Munsell color system, the colors were of a constant red hue (5R) but varied in brightness (value) and saturation (chroma). The Munsell value/chroma specifications for Colors 1–12 shown in Figure 1 were 7/4, 7/8, 6/6, 6/10, 5/4, 5/8, 5/12, 4/6, 4/10, 3/4, 3/8, and 3/10.

The colors were mounted behind 5/16-in. (0.79-cm) diameter circles punched in the middle of white index cards. To ensure that subjects were attending to the actual colors and not to incidental markings on the cards, four different tokens were used for each color. A total of 19 tokens of Color 2 (7) were used in Condition E2 (E7).

Procedure

Learning phase. On any given trial, a subject viewed a color and guessed its category assignment (Figure 1). The subject entered the response on a score sheet and then turned over the card to view the correct answer. A code number was also entered on the back of the card to identify the stimulus for the experimenter. The subject entered this code number on the score sheet next to the category response and then restudied the stimulus–response mapping.

Subjects were tested for three blocks of trials. There were 48 trials in each block of Condition B and 63 trials in each block of Conditions E2 and E7. The experimenter shuffled the deck of cards between each block.

Transfer phase. Subjects were presented with each of the 12 colors in a random order. The subject classified the color into either Category 1 or 2, rated on a scale from 1 (lowest) to 10 (highest) how confident he or she was about being correct, and then rated on a scale from 1 to 10 how typical or how good an example the color was of its category. The instructions emphasized that even if a subject was highly confident that a color belonged to a given category, one could still judge it as a poor example of the category. It was hoped that the distinction between classification confidence and judged typicality would be reinforced by collecting both types of ratings. Subjects were then presented with all pairs of colors from Category 2 and were instructed to choose the better example of Category 2 in each pair. Presentation order of the pairs and left–right placement of the colors within each pair were randomized.

Results

Classification Learning

Table 1 shows the mean proportion of classification learning errors for each color in each condition. The distributions of errors were similar across conditions, with colors closest to the category boundary (1, 3, 6, 8, 9, and 12) having the most errors, and colors far from the boundary (2, 4, and 10) the least. Color 2 had a smaller proportion of errors in Condition E2 than in the other conditions, average $t(98) = 4.10$, $p < .001$, whereas Color 7 had a smaller proportion of errors in Condition E7, average $t(98) = 4.08$, $p < .001$. The trends for Colors 4 and 9 were in the predicted directions but were small in magnitude and not statistically significant.

Typicality Ratings

The results for confidence and typicality ratings were essentially identical and so I report only the typicality data. The

Table 1
Proportion of Classification Learning Errors for Each Color:
Experiment 1

Color	Condition		
	B	E2	E7
1	.318	.296	.308
2	.123	.026	.147
3	.513	.460	.555
4	.113	.067	.103
5	.175	.114	.160
6	.337	.328	.384
7	.130	.181	.039
8	.162	.116	.131
9	.372	.409	.345
10	.097	.050	.066
11	.143	.103	.146
12	.272	.223	.267

typicality ratings were converted to a 20-point scale by transforming all ratings associated with Category 1 responses to negative values. Thus, -10 represents the rating associated with the most typical members of Category 1, whereas $+10$ represents the rating associated with the most typical members of Category 2. The mean typicality ratings for each color in each condition are shown in Table 2. There were large differences in the typicality ratings associated with the different stimuli. As expected, Colors 2, 4, and 7 were rated as highly typical members of Category 2, whereas Colors 3, 6, and 9 were rated as less typical. Color 10 was rated as the most typical exemplar of Category 1, whereas Color 1 was rated least typical. More interesting, the pattern of typicality ratings interacted with learning conditions. In Condition E2, Colors 2 and 4 were rated more typical of Category 2 than was Color 7, whereas the reverse pattern was observed in Condition E7. The typicality ratings for Color 9, which was highly similar to Color 7, also increased in Condition E7.

The reliability of these observations is confirmed by statistical tests. Analyses of variance on the typicality ratings revealed a significant effect of colors, $F(11, 1617) = 266.02$, $MS_e = 21.57$, $p < .01$; and a more important Condition \times

Table 2
Mean Typicality Ratings: Experiment 1

Color	Condition		
	B	E2	E7
1	-1.04	-2.52	-2.54
2	7.96	9.70	6.32
3	.74	.28	.18
4	8.38	9.20	7.02
5	-5.50	-5.16	-6.18
6	2.02	4.00	2.88
7	7.22	6.90	9.10
8	-6.86	-6.40	-6.28
9	1.72	2.34	4.60
10	-8.76	-8.40	-8.82
11	-6.78	-7.00	-4.80
12	-4.84	-5.34	-2.00

Note. -10 = Category 1 most typical; $+10$ = Category 2 most typical.

Color interaction, $F(22, 1617) = 2.92$, $MS_e = 21.57$, $p < .01$. Separate t tests revealed that the typicality ratings for Colors 2, 4, 7, and 9 all changed significantly across Conditions E2 and E7, average $t(98) = 4.13$, $p < .001$. These results are consistent with the view that similarity and frequency jointly determine graded category structure. Typicality ratings increased for exemplars that were presented with high frequency, and also increased for colors that were very similar to the high-frequency exemplars.

With the exception of Color 12, typicality ratings for the remaining exemplars were fairly stable across conditions. Color 12 was rated as a significantly less typical member of Category 1 in Condition E7 than in Conditions B and E2, average $t(98) = 2.80$, $p < .01$. Although not the focus of Experiment 1, this result is consistent with another prediction of the frequency-sensitive exemplar model, namely that typicality should decrease for members of contrast categories that are similar to high-frequency exemplars of a target category. (As will be seen in ensuing theoretical analyses, subjects had a tendency to attend selectively to the saturation dimension during their classification learning. This selective attention tendency results in Color 12 becoming more similar to Color 7 than is illustrated in Figure 1.) The hypothesis of decreasing typicality for members of contrast categories that are similar to high-frequency exemplars is tested more directly in Experiment 2.

Typicality Pair Comparisons

The results for the typicality pair comparisons are shown in Table 3. The entry in cell (i, j) gives the frequency with which the color in row i was judged as a better example of Category 2 than was the color in column j . The pattern of results for the typicality pair comparisons mirrors the pattern observed for the typicality ratings. Colors 2, 4, and 7 dominated Colors 3, 6, and 9 across all three conditions. More interesting, manipulating the frequency of Exemplars 2 and 7 had a major effect on the pattern of typicality pair comparisons. In Condition E2, Colors 2 and 4 were judged as better examples of Category 2 than was Color 7. The reverse pattern was observed in Condition E7. Color 9 also made consistent gains in category goodness across Conditions E2 and E7 (except, of course, relative to Color 7). Furthermore, Color 4 made gains relative to Color 2 across Conditions E2 and E7. This result would be expected for two reasons. First, decreasing the frequency of Color 2 should "hurt" Color 2 more than Color 4; and second, increasing the frequency of Color 7 should "help" Color 4 more than Color 2 because Color 4 is more similar to Color 7 than is Color 2 (see Figure 1). To confirm the reliability of these observations, chi-square tests of independence were conducted for each pair of colors across Conditions E2 and E7. The changes in pair-comparison judgments were statistically significant for numerous pairs of colors, and were particularly large for Pairs 2-4, 2-7, and 4-7, average $\chi^2(1, N = 100) = 22.1$, $p < .001$. The results of the typicality pair-comparison tests provide further evidence that similarity and frequency jointly influence graded category structure.

Table 3
Frequency With Which Row Stimulus Was Selected as a Better Example of Category 2 Than Was Column Stimulus: Experiment 1

Color and condition	Color					
	2	3	4	6	7	9
2						
B	—	45	34	44	33	43
E2	—	49	43	49	42	48
E7	—	43	26	42	17	36
3						
B	5	—	8	19	12	21
E2	1	—	1	16	12	21
E7	7	—	1	15	1	12
4						
B	16	42	—	44	27	43
E2	7	49	—	47	38	48
E7	24	49	—	45	12	41
6						
B	6	31	6	—	6	23
E2	1	34	3	—	6	24
E7	8	35	5	—	0	16
7						
B	17	38	23	44	—	48
E2	8	38	12	44	—	46
E7	33	49	38	50	—	49
9						
B	7	29	7	27	2	—
E2	2	29	2	26	4	—
E7	14	38	9	34	1	—

Note. Entry in cell (i, j) plus entry in cell (j, i) equals 50.

Theoretical Analysis

In this section quantitative tests are made of the proposed exemplar approach to representing the joint roles of similarity and frequency information.

Classification

For the frequency-sensitive exemplar model, all the N_j s in Equation 2 are set equal to one, except that N_2 and N_7 are set equal to 19/4 in Conditions E2 and E7, respectively. In the frequency-insensitive model (Equation 1), N_2 and N_7 are held fixed at one.

To apply the models, a method is needed for computing the interexemplar (η_{ij}) similarity values in Equations 1 and 2. For the continuous-dimension stimuli used in this experiment, it is natural to apply the multidimensional scaling approach (Shepard, 1958b, 1962; Torgerson, 1958). The similarity between each pair of exemplars i and j (η_{ij}) is assumed to be a monotonically decreasing function of their distance in the psychological space (d_{ij}), $\eta_{ij} = f(d_{ij})$. The distance between each pair of exemplars is computed by using a weighted Euclidean metric:

$$d_{ij} = \sqrt{w_1(x_{i1} - x_{j1})^2 + (1 - w_1)(x_{i2} - x_{j2})^2}, \quad (3)$$

where x_{ik} is the psychological value of exemplar i on dimension k and w_1 ($0 \leq w_1 \leq 1$) is the weight given to Dimension 1 (saturation) in computing overall distance. Previous research suggests that the Euclidean metric provides an accurate de-

scription of psychological distance relations for stimuli varying along "integral" dimensions, such as the present Munsell colors (Garner, 1974; Shepard, 1958b; Shepard & Chang, 1963). The weight parameter in Equation 3 is intended to reflect the role of selective attention strategies that may operate during classification learning (Medin & Schaffer, 1978; Nosofsky, 1987; Reed, 1972). Note that the x_{ik} coordinates in Equation 3 are given by the multidimensional scaling solution for the colors derived in Nosofsky's (1987) identification learning study (see Nosofsky, 1987, Table 3).

The distance d_{ij} is converted to a similarity measure using an exponential decay function:

$$\eta_{ij} = e^{-cd_{ij}}, \quad (4)$$

Previous research indicates that the exponential decay function describes accurately the relation between similarity and psychological distance in classification learning situations (Nosofsky, 1987; Shepard, 1958a, 1984, 1986). The parameter c ($0 \leq c \leq \infty$) in Equation 4 is a scale factor reflecting overall discriminability in the psychological space.

In summary, the distance between each pair of exemplars is computed by using Equation 3; this distance is converted to a similarity measure by using Equation 4; the derived similarity values are then substituted into Equation 2 (or 1) to predict the classification probabilities. The model uses three parameters: the category response bias parameter b_1 in Equation 2; the attention weight w_1 in Equation 3; and the scale parameter c in Equation 4. In the present situation, this three-parameter model is used to predict 36 classification probabilities, 12 in each of Conditions B, E2, and E7.

The frequency-sensitive exemplar model was fitted simultaneously to the Block 3 classification learning data obtained in Conditions B, E2, and E7 by using a maximum-likelihood criterion (see Nosofsky, 1985, 1987, for discussions of the merits and logic behind this model-fitting approach). The maximum-likelihood parameters and summary fits are presented in Table 4 and the predicted and observed Category 2 response probabilities are given in Table 5. The three-parameter model provides a good description of the data, accounting for 97.9% of the response variance. The model predicts correctly the trends of increasing classification accuracy for Colors 2 and 4 in Condition E2 and Colors 7 and 9 in Condition E7. However, even this frequency-sensitive model underestimates correct classification probabilities for the high-frequency exemplars (2 and 7).² Note that the model predicts decreasing classification accuracy for Color 12 in Condition E7. Although this trend was not observed in the classification learning data, it showed up in the postacquisition typicality ratings.

The frequency-insensitive exemplar model performed markedly worse than the frequency-sensitive model ($-\ln L = 280.0$). Not surprisingly, a major shortcoming of the model

² The deviations between predicted and observed probabilities for the high-frequency exemplars made large contributions to $-\ln L$ because they are based on nearly five times the number of observations as other cells and because probabilities near unity have very small standard errors.

Table 4
Maximum-Likelihood Parameters and Summary Fits for the Frequency-Sensitive Exemplar Model in Block 3 of Classification Learning

Experiment	Parameter			Fit		
	<i>c</i>	w_1	b_1	-ln L	SSE	% var
1	1.23	.67	.61	193.9	.099	97.9
2	1.33	.67	.62	199.7	.124	97.6

Note. ln L = log-likelihood; SSE = sum of squared deviations between observed and predicted classification probabilities; % var = percentage of variance accounted for. SSE and % var are presented as auxiliary measures. The criterion of fit was maximum likelihood.

was that it largely underpredicted correct classification probabilities for Color 2 in Condition E2 and Color 7 in Condition E7.

Because manipulating presentation frequencies of the individual exemplars results in changes in a priori probabilities of the categories, a plausible hypothesis is that there may have been changes in category response bias across the three conditions (Busemeyer, Dewey, & Medin, 1984; Parducci, 1974). To test this possibility, five-parameter versions of the models were fitted to the data in which the category bias parameter was allowed to vary across conditions. For the frequency-sensitive exemplar model, increasing the number of free parameters yielded essentially no improvement in fit, and the bias parameter remained essentially constant at $b_1 = .61$ across the three conditions. The additional bias parameters slightly improved the fit of the frequency-insensitive exemplar model, but it still performed markedly worse than the frequency-sensitive model. The important point is that manipulating presentation frequencies of the individual exemplars appears to have led to changes in local classification probabilities rather than global changes in overall category response bias.

A few comments are in order regarding interpretation of the best-fitting exemplar model parameters. The scale parameter *c* is simply an overall sensitivity parameter. Its estimated value in the present classification conditions is close to the value observed in Nosofsky's (1987) identification condition at a similar stage in learning. The value $w_1 = .67$ indicates that subjects had a tendency to differentially weight the saturation dimension relative to the brightness dimension in their classification learning. It should be acknowledged that this tendency was unanticipated because the saturation and brightness dimensions appear to be equally relevant in defining the "pink-brown" category structure. Finally, the value $b_1 = .61$ indicates a bias toward making Category 1 responses relative to Category 2 responses (after the effect of similarity comparisons to the stored exemplars). In Nosofsky's (1987) identification learning study, it was found that individual bias parameters associated with the "brownish" stimuli were generally larger than those associated with the "pinkish" stimuli (see Nosofsky, 1987, Table 3). The Category 1 bias in the present experiment may be reflecting this differential "stimulus bias."

As a source of comparison, central-tendency prototype models were fitted to the classification learning data (Franks & Bransford, 1971; Reed, 1972; see Nosofsky, 1987, for a precise formalization of the models in the present context).

Both frequency-insensitive and frequency-sensitive versions were formulated. In the frequency-sensitive version, the category central tendency is computed by using a weighted average over the coordinate values associated with each of the individual exemplars. Thus, the central tendency is shifted in the direction of the highest-frequency exemplars. Both the frequency-sensitive and frequency-insensitive prototype models performed considerably worse than the frequency-sensitive exemplar model (average $-\ln L = 286.6$). In line with previous research, the exemplar model appears to provide a better quantitative account of classification learning than do prototype models, at least in situations in which category sizes are small and experience with individual exemplars is extensive (Medin & Schaffer, 1978; Nosofsky, 1986, 1987).

Typicality Ratings and Pair Comparisons

The Pearson product-moment correlations between the frequency-sensitive exemplar model predicted Category 2 response probabilities and the mean typicality ratings were .953,

Table 5
Predicted and Observed Category 2 Response Probabilities in Block 3 of Classification Learning

Color	Experiment 1: Condition			Experiment 2: Condition		
	B	E2	E7	B	E6(3)	E6(5)
1						
PP	.130	.163	.135	.111	.137	.158
OP	.210	.251	.259	.224	.249	.229
2						
PP	.941	.983	.945	.953	.956	.959
OP	.925	.996	.930	.951	.915	.925
3						
PP	.597	.636	.608	.610	.666	.704
OP	.480	.640	.550	.523	.640	.655
4						
PP	.938	.962	.949	.950	.956	.960
OP	.915	.995	.970	.975	.955	.971
5						
PP	.127	.143	.134	.110	.153	.188
OP	.105	.095	.110	.071	.157	.154
6						
PP	.725	.751	.754	.743	.870	.909
OP	.714	.767	.735	.804	.873	.957
7						
PP	.898	.905	.970	.914	.920	.925
OP	.940	.865	.999	.944	.931	.921
8						
PP	.148	.162	.166	.128	.213	.275
OP	.106	.095	.099	.060	.114	.127
9						
PP	.727	.737	.803	.745	.779	.802
OP	.698	.668	.754	.694	.715	.780
10						
PP	.027	.029	.031	.019	.032	.043
OP	.030	.035	.020	.020	.050	.020
11						
PP	.084	.089	.120	.070	.096	.119
OP	.080	.045	.060	.049	.120	.045
12						
PP	.102	.106	.166	.085	.104	.120
OP	.205	.180	.203	.195	.186	.096

Note. PP = predicted probability; OP = observed probability.

.975, and .960 in Conditions B, E2, and E7, respectively. The corresponding Spearman rank-order correlations were .902, .967, and .921.

I do not attempt to provide quantitative predictions of the typicality pair comparisons in this article. As an initial gauge of the adequacy of the frequency-sensitive exemplar model, however, one can compare the direction of each pair-comparison judgment to the classification-predicted direction. For example, the exemplar model predicts that in the baseline Condition B, Color 2 will be classified in Category 2 with probability .941, whereas Color 3 will be classified in Category 2 with probability only .597. In agreement with this prediction, Color 2 was judged as a better example of Category 2 than was Color 3 by 45 of the 50 subjects in Condition B. The exemplar model predicts correctly the direction of all 15 pair comparisons in Condition B, 14 of 15 pair comparisons in Condition E2, and 14 of 15 pair comparisons in Condition E7. Furthermore, the single discrepancies in Conditions E2 and E7 are small in magnitude. For example, in Condition E2 the predicted Category 2 response probabilities for Colors 6 and 9 were .751 and .737, respectively; Color 6 was chosen as a better example of Category 2 than was Color 9 by 24 of the 50 subjects. At least at a qualitative level, the exemplar model predicts quite well the pattern of typicality pair-comparison data.

Discussion

The results of Experiment 1 support the suggestion that similarity and frequency jointly determine graded category structure. Classification accuracy and typicality ratings increased for high-frequency exemplars, and also increased for category members that were similar to the high-frequency exemplars. The theoretical analysis provided support for the exemplar-based approach to representing the joint influence of these variables.

A potential problem of interpretation concerns the relation between degree of classification learning of individual exemplars and the typicality judgments. As presentation frequency increases, there is presumably a higher probability of learning the association between a given exemplar and its category assignment. Could changes in typicality judgments simply have been reflecting differences in the probability with which individual exemplar-category associations had been formed? At least under the present experimental conditions, such an explanation taken by itself seems implausible. The critical exemplars in Experiment 1 were "good" exemplars that had extremely high probabilities of correct classification by the end of learning, regardless of whether they had been presented with high frequency. Across all conditions, the average number of subjects who misclassified Exemplars 2, 4, and 7 in the transfer phase were 1.67, 0.67, and 2.67, respectively. Although the probability of forming correct exemplar-category associations will certainly influence typicality judgments, it is unlikely that this was the sole controlling factor in the present experiment. Presentation frequency influenced the pattern of typicality data even for well-learned exemplars.

Experiment 2

Whereas Experiment 1 manipulated presentation frequencies of "good" exemplars, Experiment 2 manipulates presentation frequency of a relatively "poor" exemplar, namely Color 6. Across three conditions, Color 6 is presented with relative frequencies approximately 1:1, 3:1, and 5:1. One purpose of the experiment is to test whether presentation frequency plays a role for relatively atypical exemplars. I also test another prediction of the frequency-sensitive, similarity-to-exemplars model, namely that classification accuracy and typicality judgments should *decrease* for members of contrast categories that are similar to high-frequency exemplars. As shown in Figure 1, Color 8 of Category 1 is a close neighbor of Color 6. The expectation is that Color 8 will be rated as a poorer example of Category 1 as presentation frequency for Color 6 increases. The other predictions are analogous to those for Experiment 1. Color 6 should be rated as a more typical member of Category 2 as its presentation frequency increases. Colors 3 and 9 are fairly similar to Color 6 (although not as similar as the 2-4 and 7-9 couplets studied in Experiment 1), and so one might observe increases in typicality for these relatively poor exemplars as well.

Method

Subjects

A total of 150 subjects, most of whom were undergraduates at Indiana University, were hired for participation in the experiment; 50 subjects participated in each of three conditions.

Stimuli

The stimuli were the same as in Experiment 1. Four tokens of each color were used in Conditions B, E6(3), and E6(5); 12 and 19 tokens of Color 6 were used in Conditions E6(3) and E6(5), respectively. Note that Condition B was an independent replication of the baseline condition tested in Experiment 1.

Table 6
Proportion of Classification Learning Errors for Each Color: Experiment 2

Color	Condition		
	B	E6(3)	E6(5)
1	.320	.332	.288
2	.084	.155	.141
3	.471	.420	.415
4	.081	.118	.119
5	.164	.195	.190
6	.239	.208	.123
7	.120	.167	.157
8	.132	.168	.167
9	.390	.328	.336
10	.088	.115	.087
11	.134	.169	.108
12	.275	.254	.182

Table 7
Mean Typicality Ratings: Experiment 2

Color	Condition		
	B	E6(3)	E6(5)
1	-3.08	-2.36	-2.80
2	7.48	7.40	6.70
3	1.32	1.14	1.76
4	8.34	7.12	7.58
5	-6.84	-5.90	-5.58
6	3.18	4.34	5.76
7	7.64	7.00	6.68
8	-7.04	-5.30	-3.84
9	2.68	3.24	4.78
10	-9.34	-8.72	-8.72
11	-6.06	-5.66	-8.18
12	-4.10	-3.02	-6.38

Note. -10 = Category 1 most typical, +10 = Category 2 most typical.

Procedure

All other aspects of the procedure were the same as for Experiment 1.

Results

Classification Learning

The proportion of classification learning errors for each color in each condition is shown in Table 6. The main results of interest are that classification errors for Color 6 decreased across Conditions B, E6(3), and E6(5), and to a smaller extent decreased for Colors 3 and 9. There was a small increase in the proportion of errors for Color 8 across conditions. Although these trends are in the predicted directions, separate *t* tests using the Conditions B and E6(5) data revealed that only the changes for Color 6 were statistically significant, $t(98) = 3.83, p < .001$.

Typicality Ratings

The mean typicality ratings are shown in Table 7. The general pattern of ratings for individual stimuli within each condition mirrors the pattern observed in Experiment 1, with Colors 2, 4, 7, and 10 being rated as the most typical members of their categories, and Colors 1, 3, 6, 8, 9, and 12 (close to the category boundary) being rated the least typical. Mean typicality ratings also interacted with learning conditions. As presentation frequency for Color 6 increased, Colors 6, 3, and 9 were rated as more typical members of Category 2, whereas Color 8 was rated as a less typical member of Category 1. Results of *t* tests revealed that the changes in typicality ratings across Conditions B and E6(5) were statistically significant for Colors 6, 8, and 9, average $t(98) = 2.68, p < .01$, but not for Color 3, $t(98) = .38, p > .70$. The *t* tests also revealed that Colors 11 and 12 were rated as significantly better examples of Category 1 in Condition E6(5) than in Condition B, average $t(98) = 2.28, p < .05$. This latter result was unanticipated, but

may be reflecting some form of "response contrast" effect (Helson, 1964; Parducci, 1974).

Typicality Pair Comparisons

The results of the typicality pair-comparison task are shown in Table 8. Once again, "good" Colors 2, 4, and 7 tend to dominate "poor" Colors 3, 6, and 9. The more interesting result is that typicality preferences for poor Exemplars 6, 3, and 9 increased relative to the good Exemplars 2, 4, and 7 as presentation frequency of Color 6 was increased. Chi-square tests of independence conducted for each pair of colors across Conditions B and E6(5) showed a significant effect of conditions for numerous pairs of colors. Unexpectedly, however, typicality preference for Color 6 did not increase relative to its neighbors Colors 3 and 9.

Theoretical Analysis

Classification

The three-parameter, frequency-sensitive exemplar model was fitted simultaneously to the Block 3 classification data obtained in Conditions B, E6(3), and E6(5) by using a maximum-likelihood criterion. The best-fitting parameters and summary fits are shown in Table 4 and the predicted and observed Category 2 response probabilities are shown in Table 5. The model accounts for 97.6% of the response variance,

Table 8
Frequency With Which Row Stimulus Was Selected as a Better Example of Category 2 Than Was Column Stimulus: Experiment 2

Color and Condition	Color					
	2	3	4	6	7	9
2						
B	—	48	29	45	29	44
E6(3)	—	42	18	40	24	36
E6(5)	—	38	19	30	21	32
3						
B	2	—	1	17	9	22
E6(3)	8	—	7	16	8	19
E6(5)	12	—	9	16	10	22
4						
B	21	49	—	48	28	46
E6(3)	32	43	—	41	26	37
E6(5)	31	41	—	36	23	36
6						
B	5	33	2	—	6	30
E6(3)	10	34	9	—	6	25
E6(5)	20	34	14	—	15	30
7						
B	21	41	22	44	—	49
E6(3)	26	42	24	44	—	42
E6(5)	29	40	27	35	—	39
9						
B	6	28	4	20	1	—
E6(3)	14	31	13	25	8	—
E6(5)	18	28	14	20	11	—

Note. Entry in cell (*i,j*) plus entry in cell (*j,i*) equals 50.

General Discussion

Summary

and predicts correctly the trends of increasing classification accuracy for Colors 3, 6, and 9 across Conditions B through E6(5), and decreasing classification accuracy for Color 8. Note that the values of the best fitting parameters are very close to those estimated for the conditions in Experiment 1. Thus, the frequency-sensitive exemplar model traces the changes in classification patterns across conditions with a fair degree of parameter invariance. As in Experiment 1, allowing the category bias parameter to vary across conditions did not significantly improve the fit of the frequency-sensitive exemplar model. And once again, the frequency-sensitive exemplar model considerably outperformed all the alternative models in terms of overall fit.

Typicality Ratings and Pair Comparisons

The Pearson product-moment correlations between the exemplar model predicted Category 2 response probabilities and observed typicality ratings were .972, .964, and .982, in Conditions B, E6(3), and E6(5), respectively. The corresponding Spearman rank-order correlations were .883, .918, and .958. The exemplar model correctly predicts the direction of 14 of 15 pair comparisons in Condition B, 12 of 13 pair comparisons in Condition E6(3) (two ties), and 14 of 15 pair comparisons in Condition E6(5). (The single discrepancies in each condition are all small in magnitude.)

Note that the exemplar model predicts correctly that people will judge Color 6 as a poorer example of Category 2 than Colors 2, 4, and 7, although Color 6 was presented five times as often as these other colors. Theoretically, Color 6 would have had to have been presented approximately 12 times as often as these other colors before it achieved "equal typicality status."³

Discussion

The results of Experiment 2 extend those of Experiment 1 by showing that frequency manipulations can influence typicality gradients associated with relatively poor exemplars as well as good exemplars, and can also influence typicality gradients of contrast categories. The principle interpretation of the present results is that increasing the presentation frequency of Color 6 led to increasing the relative frequency with which this exemplar was stored in memory. Another possibility is that the frequency manipulations led to more differentiated perceptual representations in the region of Color 6 (e.g., Gibson & Gibson, 1955; Krumhansl, 1978; Nosofsky, 1987). Or, perhaps the boundary separating Categories 1 and 2 became sharper. It is unclear, however, why increasing perceptual differentiation in the region of Color 6 would lead to Color 8 being judged as a *worse* example of Category 1. Although increasing perceptual differentiation as a function of frequency or category density is likely to be part of the story, other factors also appear to be operating in the present experiments.⁴

The hypothesis motivating this research was that similarity and frequency information are jointly reflected in people's category representations and that both variables influence graded category structure. This hypothesis was tested by conducting classification learning experiments in which presentation frequency of exemplars was manipulated. The exemplars had varying degrees of similarity to other members of the target and contrast categories. The main qualitative results supported the hypothesis. Classification accuracy and typicality ratings increased for exemplars that were presented with high frequency and increased for members of the target category that were very similar to the high-frequency exemplars. Classification accuracy and typicality ratings decreased for members of the contrast category that were similar to the high-frequency exemplars.

The second focus of this research was to test a model for interpreting the joint roles of similarity and frequency information. The conceptual underpinning of the approach is the assumption that people learn categories by storing individual exemplars in memory. Classification decisions are based on similarity comparisons to the stored exemplars. Frequency information is represented naturally in the model in terms of the differential frequency with which individual exemplars are stored in memory. The frequency-sensitive exemplar model provided a good quantitative account of the classification learning data and of postacquisition typicality ratings. Quantitative comparisons favored the predictions of the model over those of a frequency-sensitive prototype model, and frequency-insensitive exemplar and prototype models. A number of other classification models also appear well equipped to handle the joint roles of similarity and frequency, but leave unspecified the precise format of the stimulus representation needed for making quantitative predictions (e.g., Eich, 1982; Hintzman, 1986; Knapp & Anderson, 1984). The data reported in this article should provide a fertile testing ground for these alternative models.

Unresolved Issues and Questions

Although this work provides preliminary support for the exemplar approach to modeling the roles of similarity and

³This computation assumes a continued linear relation between actual frequency and represented frequency of the exemplars. Busemeyer, Dewey, and Medin (1984, p. 646) suggested, however, that the relation between represented and actual frequency may be negatively accelerated, in which case an even greater relative frequency would be required for Color 6 before it achieved equal typicality status.

⁴The notion of increasing perceptual differentiation may help explain why even the frequency-sensitive exemplar model underpredicted correct classification probabilities for the high-frequency exemplars in Experiment 1 (see Table 5). A fully specified model will also need to account for memorial sequential effects induced by the differential presentation frequencies.

frequency in classification learning, it is important to point out limitations of the approach and new questions that are raised.

Frequency

First, it seems useful to distinguish between different kinds of frequency information. Barsalou (1985) distinguished between the overall subjective frequency with which a person has experienced an object and the subjective frequency with which the person has experienced the object as an example of a particular category. The present research was concerned with the latter kind of category-instantiated frequency. Although correlational work has been reported, the potential effect of overall frequency and familiarity, independent of category assignment, has apparently not yet been studied in an experimentally controlled manner. It also seems important to distinguish between the frequency with which a specific exemplar is experienced and the frequency with which a class of exemplars is experienced. For example, one's judgment of how typical rabbits are of the category *rodents* may differ dramatically depending on whether the person has frequently experienced one particular rabbit (e.g., his or her pet) or a large number of different rabbits.

Although useful for an initial investigation, the baseline model tested in this research undoubtedly gives an oversimplified picture of the relation between psychological and actual frequency. As Tversky and Kahneman (1973) have made well known, for example, judged frequency may vary dramatically depending on the retrievability and availability of the instances. Other issues of concern include the nature of memory decay of individual exemplars over time, and sequential effects induced by differential presentation frequencies. A fully specified model will need to incorporate these factors when characterizing the role of frequency information in classification learning.

Similarity and Frequency

The present research emphasized the joint, interactive roles of similarity and frequency in determining graded category structure. It is important to realize, however, that similarity and frequency can exert mutual influence on one another, thereby making more intricate the relation between similarity, frequency, and categorization. It has been suggested, for example, that increasing the frequency of a given stimulus may lead to increasing perceptual differentiation in the region of that stimulus. In previous research, Medin and his associates (Medin, Dewey, & Murphy, 1983; Medin & Schaffer, 1978; Medin & Smith, 1981) and Nosofsky (1984, 1986, 1987) emphasized the role of selective attention in modifying similarity relations among exemplars. Differential frequency information may lead to certain selective attention strategies benefiting classification performance more than others. Thus, frequency may modify similarity relations because of the influence of selective attention. Luce, Green, and their associates have provided evidence for a selective attention mechanism that monitors local regions of psychological dimensions

(Luce, Green, & Weber, 1976). Changes in sensitivity resulting from sequential dependencies and differential presentation frequencies have been interpreted in terms of systematic shifts of this band of selective attention (Luce & Nosofsky, 1984; Luce, Nosofsky, Green, & Smith, 1982; Nosofsky, 1983).

Just as frequency may influence similarity, so may similarity relations influence frequency judgments. Kahneman and Tversky (1972) demonstrated that people often judge the probability of an event according to how similar it is to the essential characteristics of its population. Use of this "representativeness" heuristic can lead to systematic biases in probability judgment, such as ignoring base rates. A related effect of the influence of similarity on frequency judgments was observed in the present study. Following the postacquisition typicality judgments in Experiment 1, subjects were presented with all exemplars of Category 2 in a pseudorandom arrangement. They were asked to rank the exemplars according to the frequency with which they had been presented during the classification learning phase. As expected, Colors 2 and 7 were ranked first in Conditions E2 and E7, respectively. The more interesting result is that Color 4, which was highly similar to Color 2, was ranked second in Condition E2; whereas Color 9, which was highly similar to Color 7, was ranked second in Condition E7. (The differences in these rank orderings were highly reliable.) The similarity relations among the stimuli apparently led to the development of differential "frequency-generalization" gradients. An interesting question for future research concerns the locus of the effect: For example, does it reside in the storage of differential frequency information or in decision factors operating at the time of retrieval?

Relations Between Classification and Typicality

A noteworthy feature of the results reported in this article was the close correspondence between the context model predicted classification probabilities and the observed typicality pair comparisons. Because the context model assumes that classification is determined by relative degree of within-category to between-category similarity, the suggestion is that both kinds of similarity relations influenced the typicality judgments. Indeed, between-category similarity was even more important than within-category similarity in the baseline (equal-frequency) conditions tested in the present experiments. Colors 2, 4, and 7 were "good" exemplars mainly because they were less similar to members of the contrast category than were "poor" Exemplars 3, 6, and 9. The suggestion that both within-category and between-category similarity relations influence typicality converges with earlier conclusions reached by Rosch and Mervis (1975).

Previous studies have sometimes shown some divergence between classification probability and typicality ratings. Bourne (1982), for example, reported a concept-learning experiment in which the assignment of a "prototype" stimulus to either the target or contrast categories was manipulated probabilistically across conditions. Interestingly, there were some conditions in which the "prototype" was rated as a better example of the target category than was a comparison stimulus, yet was classified by subjects into the target category

with lower probability than the comparison stimulus. Nosofsky (1988) showed that a single-parameter version of the frequency-sensitive context model could account qualitatively for Bourne's classification and typicality data, as long as it was assumed that typicality judgments were based only on summed within-category similarity, rather than on relative degree of within-category to between-category similarity. "Typicality" is an open-ended construct, and the interpretation given to it by people apparently may vary depending on instructions and experimental conditions. Nevertheless, typicality judgments often admit of a high degree of regularity and structure, and can provide useful clues into the nature of people's category representations.

Limits to Generalizability

The present study was concerned with graded category structure as it arises in tasks of perceptual classification learning. Whether similar results will be obtained in other domains remains an open question. Barsalou (1985), for example, provided evidence that similarity relations among exemplars are relatively unimportant in determining the graded structure of "goal-derived" categories such as "things to do for weekend entertainment." Of particular interest is the question of the generalizability of these results to the domain of "conceptual" categorization, where the influence of real world knowledge and implicit "theories" seems crucial (Murphy & Medin, 1985). Unfortunately, it is difficult to exercise careful experimental control on the relevant variables in this arena. The complex psychological structures that underlie most "conceptual" domains and that determine interobject similarity relations generally remain unspecified. Nevertheless, it is reasonable to posit that just as in perceptual classification, the fundamental variables of frequency and similarity function along with other variables in determining the graded structure of many conceptual categories.

References

- Ashcraft, M. H. (1978). Property norms for typical and atypical items from 17 categories: A description and discussion. *Memory & Cognition*, 6, 227-232.
- Barsalou, L. W. (1981). *Determinants of graded structure in categories*. Unpublished doctoral dissertation, Stanford University.
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 629-654.
- Bourne, L. (1982). Typicality effects in logically defined categories. *Memory & Cognition*, 10, 3-9.
- Busemeyer, J. R., Dewey, G. I., & Medin, D. L. (1984). Evaluation of exemplar-based generalizations and the abstraction of categorical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 638-648.
- Eich, J. M. (1982). Composite holographic associative recall model. *Psychological Review*, 89, 626-661.
- Estes, W. K. (1986a). Array models for category learning. *Cognitive Psychology*, 18, 500-549.
- Estes, W. K. (1986b). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General*, 115, 155-174.
- Franks, J. J., & Bransford, J. D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology*, 90, 65-74.
- Garner, W. R. (1974). *The processing of information and structure*. New York: Wiley.
- Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment. *Psychological Review*, 62, 32-41.
- Hampton, J. A., & Gardiner, M. M. (1983). Measures of internal category structure: A correlational analysis of normative data. *British Journal of Psychology*, 74, 491-516.
- Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.
- Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93, 411-428.
- Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, 88, 297-306.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- Knapp, A. G., & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 616-637.
- 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.
- Luce, R. D., Green, D. M., & Weber, D. L. (1976). Attention bands in absolute identification. *Perception & Psychophysics*, 20, 49-54.
- Luce, R. D., & Nosofsky, R. M. (1984). Attention, stimulus range, and identification of loudness. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 3-25). Hillsdale, NJ: Erlbaum.
- Luce, R. D., Nosofsky, R. M., Green, D. M., & Smith, A. F. (1982). The bow and sequential effects in absolute identification. *Perception & Psychophysics*, 32, 397-408.
- Malt, B. C., & Smith, E. E. (1982). The role of familiarity in determining typicality. *Memory & Cognition*, 10, 69-75.
- Medin, D. L., Dewey, G. I., & Murphy, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 607-625.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207-238.
- Medin, D. L., & Smith, E. E. (1981). Strategies and classification learning. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 241-253.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316.
- Nosofsky, R. M. (1983). Shifts of attention in the identification and discrimination of intensity. *Perception & Psychophysics*, 33, 103-112.
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 104-114.
- Nosofsky, R. M. (1985). Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. *Perception & Psychophysics*, 38, 415-438.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 115, 39-57.
- Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 87-109.

- Nosofsky, R. M. (1988). *Exemplar-based accounts of relations between classification, recognition, and typicality*. Manuscript submitted for publication.
- Parducci, A. (1974). Contextual effects: A range frequency analysis. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 128-141). New York: Academic Press.
- Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382-407.
- Rips, L. J., Schoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 12, 1-20.
- Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111-144). New York: Academic Press.
- Rosch, E. H. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.
- Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.
- Rosch, E. H., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception & Performance*, 2, 491-502.
- Schwanenflugel, P. J., & Rey, M. (1986). The relationship between category typicality and concept familiarity: Evidence from Spanish- and English-speaking monolinguals. *Memory & Cognition*, 14, 150-163.
- Shepard, R. N. (1958a). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, 65, 242-256.
- Shepard, R. N. (1958b). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *Journal of Experimental Psychology*, 55, 509-523.
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function: I. *Psychometrika*, 27, 124-140.
- Shepard, R. N. (1984, November). *Similarity and a law of universal generalization*. Paper presented at the 25th annual meeting of the Psychonomic Society, San Antonio, TX.
- Shepard, R. N. (1986). Discrimination and generalization in identification and classification: Comment on Nosofsky. *Journal of Experimental Psychology: General*, 115, 58-61.
- Shepard, R. N., & Chang, J. J. (1963). Stimulus generalization in the learning of classifications. *Journal of Experimental Psychology*, 65, 94-102.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: Wiley.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

Received October 6, 1986

Revision received March 9, 1987

Accepted March 16, 1987 ■

Cutting Appointed Editor of *JEP: Human Perception and Performance*, 1989-1994

The Publications and Communications Board of the American Psychological Association announces the appointment of James E. Cutting, Cornell University, as editor of the *Journal of Experimental Psychology: Human Perception and Performance* for a 6-year term beginning in 1989. The current editor, William Epstein, will be receiving submissions through September 30, 1987. At that point, the 1988 volume will have been filled, and all submissions after that should be sent to James Cutting. Therefore, as of October 1, 1987, manuscripts should be directed to:

James E. Cutting
Department of Psychology
Uris Hall
Cornell University
Ithaca, New York 14853-7601
