Journal of Experimental Psychology: General
1988, Vol. 117, No. 4, 412–414

0096-3445/88/$00.75

On Exemplar-Based Exemplar Representations: Reply to Ennis (1988)

Robert M. Nosofsky
Indiana University

The multivariate theory of similarity discussed by Ennis (1988) entails the assumption that individual category exemplars are themselves represented psychologically as distributions of individual exemplars. The potential utility of this exemplar-distribution approach for theories relating selective attention, similarity, and categorization is briefly discussed.

Nosofsky (1986) proposed and tested a unified quantitative approach to modeling performance in tasks of multidimensional stimulus identification and categorization. The approach was based on the assumption that both identification and categorization decisions are governed by similarity comparisons with individually stored exemplars. By modeling patterns of identification confusions, a multidimensional scaling solution for the exemplars was derived. This solution was then used in conjunction with the exemplar-similarity model to predict performance in four separate categorization conditions with the same set of stimuli. A key component assumption was that selective attention processes may systematically modify distances among exemplars in the multidimensional psychological space.

On the Similarity Gradient

Detailed quantitative analyses of the identification and categorization data (Nosofsky, 1985b, 1986) revealed a "Gaussian" relation between similarity and psychological distance. This Gaussian relation runs counter to Shepard's (1987) proposed universal law of generalization, in which similarity is presumed to be an exponential decay function of psychological distance. (For comparisons between the Gaussian and exponential gradients applied to confusion data, see, e.g., Nosofsky, 1985a, 1985b). Commenting on Nosofsky (1986), Shepard (1986) drew a distinction between failures of perceptual discrimination and the process of generalization. The latter is a cognitive act on the part of an organism in which a new stimulus—clearly discriminable from a previously experienced one—is nevertheless judged to have the same consequences or belong to the same class as the original. Because Nosofsky (1986) modeled performance by using highly confusable stimuli, Shepard (1986) conjectured that the Gaussian similarity gradient may have been reflecting "irreducible noise in the perceptual/memory system" (p. 60)—not the "cognitive" form of similarity intrinsic to generalization.

Ennis (1988) provided an elegant mathematical formulation of Shepard's (1986) proposal stemming from a multivariate theory of similarity (Ennis, Palen, & Mullen, in press; Mullen & Ennis, 1987; for related work, see Ashby & Perrin, 1988; Ashby & Townsend, 1986; Zinnes & MacKay, 1983). As in classic Thurstonian approaches to modeling classification, presentation of a stimulus is assumed to result in some psychological representation. However, because of noise in the perceptual/memory system, a given stimulus will not yield the same internal representation each time it is presented; rather, there will be a distribution of psychological dispersions associated with a stimulus. Thus, instead of representing an exemplar as a single point in a psychological space (as assumed by Nosofsky, 1986), a more appropriate model would represent each exemplar as a distribution of points in the psychological space. Figure 1 provides an illustration in which the psychological dispersions are assumed to be Gaussian distributed (cf. Green & Swets, 1966). Following Shepard (1987), the similarity between any two individual points in these distributions would be an exponential decay function of their psychological distance. But the overall similarity between the exemplars would reflect two components—the Gaussian distributed dispersions corresponding to each exemplar and the exponential similarity function. With highly confusable stimuli, the Gaussian noise would presumably swamp the cognitive similarity function, thereby explaining the pattern of results revealed in Nosofsky's (1985a, 1985b) analyses.

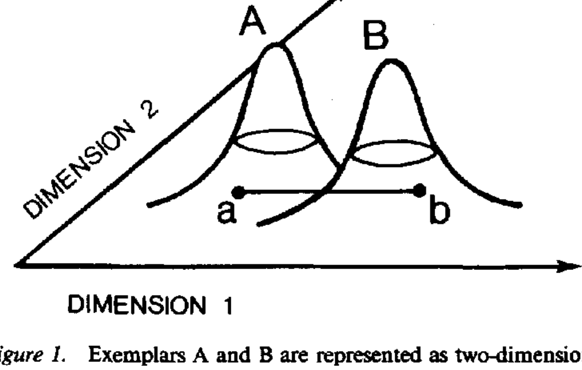

Figure 1. Exemplars A and B are represented as two-dimensional, Gaussian distributed psychological dispersions. The similarity between any two individual dispersions, a and b, would be an exponential decay function of their psychological distance: $s_{ab} = \exp(-d_{ab})$. The overall similarity between Exemplars A and B would reflect both the Gaussian noise and the exponential similarity function.

Preparation of this comment was supported by National Science Foundation Grant BNS 87-19938 to Indiana University.

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

Utility of the Distributional Approach

Although the quantitative accuracy of Ennis's (1988) proposal remains to be tested, it strikes me as highly promising and conceptually well motivated. Nosofsky (1986) operated on the assumption that categories are represented psychologically as collections of individually stored exemplars. Ennis's proposal entails an assumption that is recursive with Nosofsky's (1986): Namely, the exemplars themselves are represented psychologically as collections of individually stored exemplars. That is, the psychological representation of the exemplar corresponds to the distribution of psychological dispersions it evokes. Incorporating this assumption within the framework of Nosofsky's (1986) categorization model would be an entirely natural extension. Instead of summing the similarity of a probe to individual exemplars, one would integrate over the similarity between the probe and the exemplar distributions.

Besides its potential for reconciling Nosofsky's (1985b, 1986) work and Shepard's (1958, 1987) on the form of the similarity gradient, I believe the distributional approach to representing exemplars holds more general utility for theories of similarity and categorization. Ashby and Perrin (1988) illustrated a variety of empirical phenomena dealing with stimulus similarity that are interpretable in terms of a distributional approach. Here I illustrate several additional examples that bear on what was the central theme of Nosofsky's 1986 article: Namely, selective attention processes may systematically modify similarities among exemplars across different categorization conditions. Selective attention was described in Nosofsky's (1986) model in terms of stretching and shrinking of distances in a psychological space (Carroll & Wish, 1974), but a process interpretation was not forthcoming. In terms of the distributional approach, selective attention can be represented by decreased variability of the exemplar distributions along the attended dimensions (cf. Luce, Green, & Weber, 1976). Such a situation is illustrated in Figure 2, which shows two categories that can be perfectly discriminated on the basis of values on Dimension 1. Presumably, people would attend selectively to this dimension, resulting in less variable psychological dispersions. Thus the exemplars would be less similar along Dimension 1, and between-category discriminability would increase.

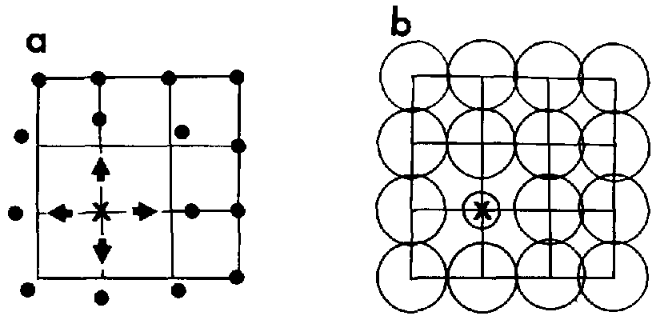

Figure 3. (a) Representation of selective attention to stimulus X in terms of increased distances between that stimulus and its neighbors in the psychological space. (Arrows indicate directions of local polarization around stimulus X.) Note, however, that distances between the neighbors and the remaining stimuli in the set have also been changed. (b) Distributional approach to representing selective attention to the single stimulus.

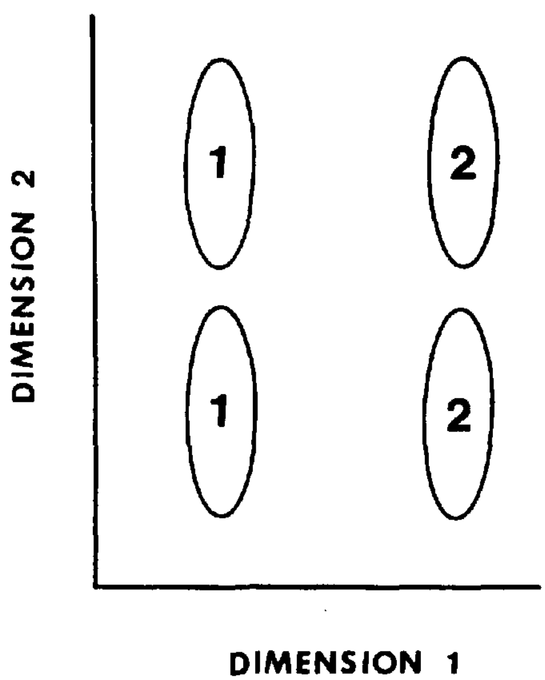

Figure 2. Exemplar distributions are illustrated as ellipses in a two-dimensional psychological space, with the length of the major and minor axes of the ellipses representing the variability of the distributions along each dimension. Exemplars labeled *1* and *2* represent, respectively, members of Category 1 and 2. Dimension 1 is perfectly diagnostic for discriminating the categories. The figure illustrates selective attention to the relevant dimension in terms of reduced variability along that dimension, as in the process model of selective attention proposed by Luce, Green, and Weber (1976).

Besides selective attention to a dimension (as illustrated in Figure 2), other forms of selective attention may be possible. For example, Garner (1978) considered the idea of selective attention to individual stimuli. Figure 3a illustrates an attempt to represent selective attention to a single stimulus in terms of increased distances between that stimulus and its neighbors in the psychological space. The problem with the representation is that distances and structural relations among the remaining stimuli in the set have also perforce been changed. A distributional approach to representing selective attention to the single stimulus is illustrated in Figure 3b and may offer a more attractive alternative. The representation in Figure 3b may also reflect what happens when an individual exemplar is presented with high frequency (Nosofsky, 1988). In addition to increased traces of the exemplar being laid down in memory, these memory traces may become more differentiated from the remaining exemplars in the set (Gibson & Gibson, 1955).

To summarize, the distributional approach to representing exemplars offers intriguing potential for developing increasingly accurate and powerful models of relations among selective attention, similarity, and categorization.

References

Ashby, F. G., & Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review, 95,* 124–150.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154–179.

Carroll, J. D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2, pp. 57–105). San Francisco: Freeman.

Ennis, D. M. (1988). Confusable and discriminable stimuli: Comment on Nosofsky (1986) and Shepard (1986). *Journal of Experimental Psychology: General, 117,* 408–411.

Ennis, D. M., Palen, J., & Mullen, K. (in press). A multidimensional theory of similarity. *Journal of Mathematical Psychology.*

Garner, W. R. (1978). Selective attention to attributes and to stimuli. *Journal of Experimental Psychology: General, 107,* 287–308.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment. *Psychological Review, 62,* 32–41.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Luce, R. D., Green, D. M., & Weber, D. L. (1976). Attention bands in absolute identification. *Perception & Psychophysics, 20,* 49–54.

Mullen, K., & Ennis, D. M. (1987). Mathematical formulation of multivariate Euclidean models for discrimination methods. *Psychometrika, 52,* 235–249.

Nosofsky, R. M. (1985a). Luce's choice model and Thurstone's categorical judgment model compared: Kornbrot's data revisited. *Perception & Psychophysics, 37,* 89–91.

Nosofsky, R. M. (1985b). Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. *Perception & Psychophysics, 38,* 415–432.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 54–65.

Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review, 65,* 242–256.

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. (1987). Toward a universal law of generalization for psychological science. *Science, 237,* 1317–1323.

Zinnes, J. L., & MacKay, D. B. (1983). Probabilistic multidimensional scaling: Complete and incomplete data. *Psychometrika, 48,* 27–48.

Received May 26, 1988
Accepted May 27, 1988 ■

Rayner Appointed Editor of *JEP: Learning, Memory, and Cognition,* 1990–1995

The Publications and Communications Board of the American Psychological Association announces the appointment of Keith Rayner, University of Massachusetts, as editor of the *Journal of Experimental Psychology: Learning, Memory, and Cognition* for a 6-year term beginning in 1990. As of January 1, 1989, manuscripts should be directed to

Keith Rayner
Department of Psychology
Tobin Hall
University of Massachusetts
Amherst, Massachusetts 01003

Manuscript submission patterns for *JEP: Learning, Memory, and Cognition* make the precise date of completion of the 1989 volume uncertain. The current editor, Henry Roediger, will receive and consider manuscripts until December 31, 1988. Should the 1989 volume be completed before that date, manuscripts will be redirected to Rayner for consideration in the 1990 volume.