Family Resemblance and Measures of Central Tendency in an M-dimensional Space as Determinants of Typicality in Categories

Eef Ameel (eef.ameel@psy.kuleuven.ac.be)
Leuven University, Department of Psychology; Tiensestraat 102
B-3000 Leuven, Belgium

Gert Storms (gert.storms@psy.kuleuven.ac.be)
Leuven University, Department of Psychology; Tiensestraat 102
B-3000 Leuven, Belgium

Abstract

Three possible determinants of typicality were examined in categories of common household objects: family resemblance and an exemplar's similarity to a geometrical centroid of the category, either unweighted or weighted for name frequencies. Similarity to the geometrical centroids in 2- to 6-dimensional spaces outperformed family resemblance in predicting typicality. Finally, optimizing the prediction of typicality in an M-dimensional space resulted in an abstract representation far outside the set of objects, suggesting that people seem to use a caricature to judge typicality. The position of this external prototype was most influenced by the typicality of nonmembers.

Introduction

Category boundaries are fuzzy rather than clear-cut. For most concepts, like *furniture*, there exist borderline cases, like *carpet* or *desk lamp*, that are not consistently assigned to the category. These exemplars are very atypical of the category *furniture*, in contrast to other exemplars, like *chair* and *table*, that are very typical members of *furniture*. Members of a category vary in how good an example they are of their category or in how typical they are of their category (Rips, Shoben & Smith, 1973; Rosch, 1973). Analogously, nonmembers of a category vary in how atypical they are of their category (Barsalou, 1985). For example, a *robin* is a better nonmember of *furniture* than a *sleeping bag*. This continuum of category representativeness, is referred to as graded structure. Prototype models for concept representation provide an account for the graded structure of categories. They assume that instances of everyday concepts are compared to an abstract summary or prototype of the concept. Category membership depends on the degree of similarity to this prototype (Rosch, 1978).

A prototype can be defined in several ways: It is usually assumed to be the central tendency of the category (Hampton, 1979, 1993, Smith, Shoben & Rips, 1974), where central tendency refers to any kind of central summary information about a category's instances (e.g. average value of the category instances, median,...). The more similar an instance is to the central tendency of the category, the more typical it is of its category. According to Rosch and Mervis (1975) family resemblances underlie the prototype structure of categories. Family resemblance is defined as an instance's average similarity to other category members and its average dissimilarity to members of contrast categories. The more characteristic features an instance has in common with other category members and the less characteristic features it has in common with members of contrast categories, the higher its family resemblance and the more typical it is of its category.

In contrast to this feature-based approach to similarity, other approaches evolved from the category learning tradition, start from a geometrical representation of a category. Instances of the category are represented as points in an M-dimensional space. Similarity is defined in terms of distances in the geometrical space. Small distances between points correspond to very similar instances, large distances to very dissimilar instances. The prototype corresponds to the point in the M-dimensional space that possesses the central tendency of dimension values, averaging over all of the concept's instances (Nosofsky, 1992; see also Minda & Smith, 2001 and Gardenfors, 2000 for geometric approaches to similarity and categorization).

Some researchers have extended the notion of a prototype to a caricature of the category, i.e. a prototype with more extreme values on the dimensions that distinguish the category from other categories (Goldstone, 1996).

The present study addresses the issue of what prototype model best determines typicality . Barsalou (1985) found different determinants of typicality for common taxonomic categories and goal-derived categories, including an exemplar's similarity to the central tendency of its category. The main question in the present paper concerns the difference in prediction quality between family resemblance and distance-based similarity to a centroid, using categories of common household objects (Malt, Sloman, Gennari, Shi & Wang, 1999).

Study

Three different determinants of typicality were examined: family resemblance and two measures of central tendency in an M-dimensional space: unweighted and weighted for name frequencies. Finally, we also looked at the nature of the abstract representation dictated by the optimized prediction of typicality.

Method

Materials The categories for which we wanted to predict graded structure were derived from a previous naming study with common household objects (Ameel, Storms, Malt & Sloman, in press). In that study, speakers of different language groups, Belgian monolingual speakers of Dutch and French and Dutch-French bilinguals, named two sets of objects, bottles and dishes. For the present study, only four categories were selected from the linguistic category names generated by Dutch-speaking monolinguals for the bottles set. The bottles set contained 73 objects that were selected to be likely to receive the name *bottle* or *jar* in American English, or else to share one or more salient properties with bottles and jars. The selection of the categories was based on the criterion of most frequent dominant name, where the dominant name of an object is the name that was most frequently generated for that object. The four linguistic categories selected were *fles*, *bus pot* and *brik*. Figure 1 shows examples of each of the four categories.

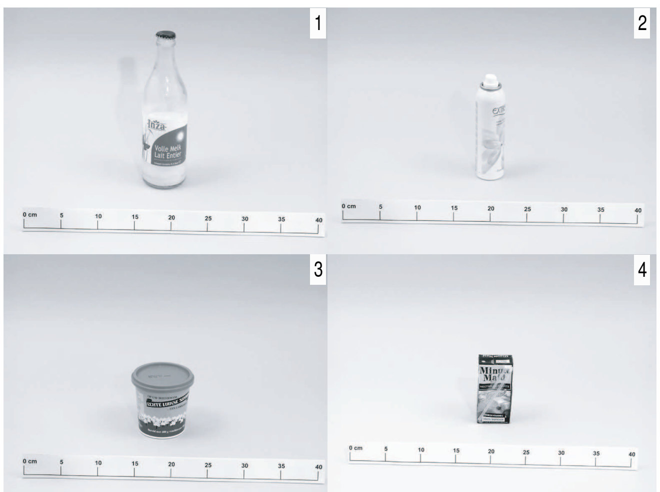

Figure 1: Black-and-white versions of some of the color photographs used in the experiment. Object 1 is a very typical example of the category *fles*, object 2 of the category *bus*, object 3 of the category *pot*, and object 4 of the category *brik*.

Subjects We used the naming and sorting similarity data for the objects of the bottles set that were gathered from thirty-two undergraduate students at the Psychology Department of the University of Leuven as part of the naming study of Ameel et al. (in press). Twenty-eight undergraduate students provided typicality ratings. Finally, 10 undergraduate students participated in a feature generation task. All subjects were Dutch-speaking monolingual Belgians. They participated for course credit.

Procedure In the naming task, the experimenter asked subjects to name the 73 objects, presented on individual pictures. They could give whatever name seemed like the best or most natural name, and they were told that they could give either a single-word name or a name with more than one word (for more details, see Ameel et al., in press). In the sorting task, we asked the subjects to sort the objects into piles based on overall similarity (for more details, see Ameel et al., in press). To obtain typicality ratings, the subjects were asked to rate on a 7-point scale the degree to which each of the objects was a good example of the linguistic categories *fles*, *bus*, *pot* and *brik*. The scale ranged from 1 to 7 with 1 labeled *very atypical* and 7 labeled *very typical*. The task was done on a computer and the objects were presented randomly on the screen. At the top of the screen appeared the category name. At the bottom of each picture appeared the 7-point scale. The subjects filled in one scale number for each object. This procedure was repeated three times, once for each category. The order of presentation of the categories was counterbalanced across subjects. Finally, in the feature generation task, constructed to derive a measure of family resemblance according to Rosch and Mervis (1975), subjects were asked to list attributes of the four linguistic categories *fles*, *bus*, *pot* and *brik*. Following the procedure introduced by Hampton (1979), a set of questions was used for each category in order to encourage subjects to generate as many different properties as they could (for more details, see Hampton, 1979).

Results

The results are presented in three sections. Each section describes a different determinant of typicality.

Family resemblance To derive a measure of family resemblance, we followed the Rosch and Mervis (1975) procedure: For each category, we listed all attributes that were generated at least twice for that category in the feature generation task. A judge assessed for each object whether each of the listed features was applicable to the object. Each attribute received a weight, corresponding to the number of objects that possessed that attribute. The basic measure of family resemblance for an object was the sum of the weighted applicability scores of all the features that had been generated for its category. To examine how well family resemblance predicts typicality, Spearman rank-order correlations between the ranks of the obtained measures of family resemblance and the ranks of the typicality ratings averaged across subjects were performed separately for each of the categories. Note that this measure of family resemblance does not take into account an object's dissimilarity to members of contrast categories. Verbeemen, Vanoverberghe, Storms and Ruts (2001) showed that contrast categories generally do not make an independent contribution to within-category structure. The resulting rank-order correlations between family resemblance and typicality for *fles*, *bus*, *pot* and *brik* were respectively 0.71, 0.14, 0.77 and 0.49. All were significant ($p < .01$), except for the correlation for *bus*[1]. The result for *fles* is comparable to the correlations reported by Rosch and Mervis

[1]The reason for the low correlation of *bus* might be found in the feature generation task. Some subjects generated features of *brik* in the feature generation task of *bus*, since the name *bus* is sometimes used for objects with *brik* as the dominant name. As a result, the set of features generated for *bus* does not reproduce an unequivocal representation of *bus*.

for basic level categories (between 0.69 and 0.94, average of 0.84), while the other correlations are lower. The lower correlations might be due to the lower number of generated frequencies. Except for *bus*, these significant positive correlations confirm that the more attributes an object has in common with other objects of the category, the more it is a representative or typical member of the category.

Centroid prototypes in an M-dimensional space unweighted and weighted for name frequencies The second determinant assumes an underlying M-dimensional geometrical representation in which the exemplars of a category are embedded. Each exemplar is represented by a vector of M coordinates, one on each of the M dimensions. The coordinates of the objects were obtained by performing multidimensional scaling (Borg & Groenen, 1997) on the similarity sorting data that were first transformed to pairwise similarity data. Pairwise similarity was recovered by counting for each of the 2628 ($\frac{73*72}{2}$) possible pairs of objects how many subjects placed that pair of objects in the same pile. A large number of subjects placing the two objects in a pile indicates high perceived similarity between the objects of the pair. Two- to 6-dimensional MDS solutions were computed.

In the category-learning tradition (Nosofsky, 1984, 1992; Minda & Smith, 2001), the prototype of a category is assumed to be a single point in the space representation. As mentioned earlier, this point could be the centroid of all category exemplars, a vector of modal values over all category exemplars, a vector of ideal values, etc (Barsalou, 1985). In the case of an assumed underlying M-dimensional space, these different definitions are reflected in the way the coordinates of the prototype are computed: as the mean of the coordinates of all category exemplars on each dimension, as the modal coordinate on each dimension, as the 'ideal' coordinate on each dimension, etc. The closer an instance is to the prototype, the more typical it is of its category.

The coordinates of the centroids for *fles*, *bus*, *pot* and *brik* were calculated as follows. Two different versions of the centroid prototypes were computed: an unweighted version and a weighted version. In the unweighted version the coordinates of the centroid of the category are based on the average coordinates of all the objects with the category name as dominant name. For instance, the coordinates for the centroid of *fles* were the average of the coordinates of all the objects with *fles* as dominant name. However, by looking at the dominant names, some information present in the data might be lost. Only five objects were unanimously called by the same name. The remaining objects were named by at least two different names. An object that has been called *bus* more often than *fles*, is assigned to the category *bus* and will not be taken into account in the computation of the coordinates of the centroid prototype of *fles*. Therefore, in the weighted version of centroids all the objects that were at least called once with the category name were involved in the computation of the coordinates of the centroids of the four categories. This was done by averaging over the coordinates of each object weighted for the frequency with which each object was called by the category name concerned. Thus, the more an object was called *fles*, the larger was its effect on the averaged coordinates of the centroid of *fles*. The same holds for the other centroids.

To examine how well typicality can be predicted by the two geometrical prototype models, the euclidean distances to the weighted and unweighted centroids of each category, calculated in 2 to 6 dimensions, were correlated with the mean typicality ratings of the objects. The euclidean distance is computed by formula 1:

$$\sqrt{\sum_{i=1}^{n}(x_{ij}-x_{ip})^2}, \tag{1}$$

with n the number of dimensions, x_{ij} the coordinate of object j on dimension i and x_{ip} the coordinate of prototype p on dimension i. Table 1 shows the correlations between the mean typicality ratings and the distances calculated in 2 to 6 dimensions for the unweighted centroids. Table 2 contains the corresponding correlations for the weighted centroids.

Table 1: Correlations between typicality and distance to the unweighted centroids.

Number of dimensions	prototypes *fles*	*bus*	*pot*	*brik*
2	-0.76	-0.58	-0.66	-0.67
3	-0.76	-0.67	-0.76	-0.89
4	-0.83	-0.69	-0.79	-0.91
5	-0.85	-0.66	-0.79	-0.93
6	-0.87	-0.75	-0.85	-0.94

Table 2: Correlations between typicality and distance to the weighted centroids.

Number of dimensions	prototypes *fles*	*bus*	*pot*	*brik*
2	-0.75	-0.57	-0.59	-0.67
3	-0.75	-0.62	-0.69	-0.89
4	-0.85	-0.66	-0.79	-0.93
5	-0.85	-0.61	-0.71	-0.93
6	-0.87	-0.67	-0.74	-0.94

As seen in Table 1 and 2, all correlations are negative: the smaller the distance from an object to the centroid of a category, the more typical the object is for the category. The correlations between typicality and the distance to the centroid generally increase from two to six dimensions for the four categories. The correlations between typicality and distance to the unweighted centroid do not differ significantly from the corresponding correlations between typicality and distance to the weighted centroid for the four categories. Apparently, the extra information provided by the name frequencies does not help to construct a better prototype.

When comparing the correlations for these centroids to the family-resemblance-based correlations, it is clear that the correlations for the centroids exceed the family-resemblance-based correlations, except for *fles* where only the correlation in six dimensions is slightly though not significantly higher than the family-resemblance-based correlation. In general, a centroid prototype represented in a multidimensional space appears to explain more of the graded structure of categories than a family-resemblance-based prototype.

Optimal prediction The centroid conception of a prototype described above may not be the optimal prototype location in the multidimensional representation to account for the rated typicality. Therefore, we looked for the prototype that guarantees a maximal or optimal (negative) correlation between the typicality ratings and the distances to that prototype. This optimization problem was solved by using the steepest ascent method. This method is shortly explained. M coordinates need to be estimated, corresponding to the number of dimensions of the geometrical representation. So, the vector Θ of 'parameters to be estimated' can be written as $\{x_{1p}, .., x_{np}\}$ with n the number of dimensions. The vector of coordinates of the category centroid is taken as the starting point $x^{(0)}$. Next, the correlation is computed between typicality and the distance to point $x^{(0)}$ as well as the gradient of the function at point $x^{(0)}$. From that point, as the name of the algorithm suggests, a step is taken in the direction of the steepest ascent, which is the direction of the gradient of the function at the point. This results in a vector $x^{(1)}$, with a higher correlation than the starting vector. This procedure is repeated (K times), until vector $x^{(K+1)}$ does not produce a higher correlation than vector $x^{(K)}$.

The method of the steepest ascent generates iterating points $x^{(k+1)}$ with formula 2:

$$x^{(k+1)} = x^{(k)} + \beta * \nabla(f(x^{(k)})). \quad (2)$$

The variable β indicates the length of the step. $\nabla(f(x^{(k)}))$ is the gradient of the function f, where f is the correlation between typicality and distance to $x^{(k)}$. The length of the step β is chosen such that $f(x^{(k+1)}) > f(x^{(k)})$. Iterating points $x^{(k+1)}$ are generated until the gradient of $\nabla(f(x^{(k+1)}))$ equals zero. At that point, the maximum has been reached, the optimal estimated parameters have been found.

This optimization algorithm was performed for prototypes in 2 to 6 dimensions. We only discuss the results of the 2-dimensional optimal points, since the results in two dimensions are similar to the results in more dimensions and since the solution in two dimensions can easily be visualized in contrast to solutions in more dimensions.

Figure 2 shows the 2-dimensional MDS solution. Each of the four categories is represented. Further, the weighted centroids for the four categories, respectively, *FLES_ct*, *BUS_ct*, *POT_ct* and *BRIK_ct* are pictured, as well as their optimal points, respectively, *FLES_opt*, *BUS_opt*, *POT_opt* and *BRIK_opt*.

For *fles*, the optimal correlation between typicality and distance to the best predicting point in two dimensions was -0.81. This correlation is significantly higher than the correlations found for the weighted and unweighted centroids of *fles* in two dimensions (see Table 1 and 2, $p < .05$).

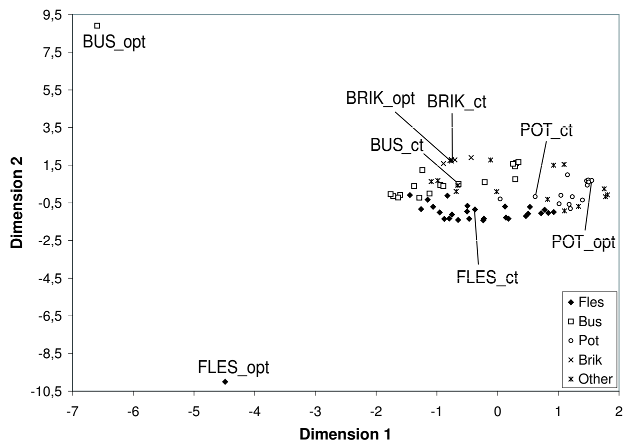

Figure 2: 2-dimensional MDS representation of the bottles. *FLES_ct*, *BUS_ct*, *POT_ct* and *BRIK_ct* represent the central tendencies for *fles*, *bus*, *pot* and *brik*. *FLES_opt*, *BUS_opt*, *POT_opt* and *BRIK_opt* represent the optimal prototypes for *fles*, *bus*, *pot* and *brik*.

Figure 2 shows that the optimal point of *fles*, *FLES_opt*, moves away from the centroid *FLES_ct* and from the set of objects. It is situated outside the cloud of points representing the objects. Its coordinates on the two dimensions reach more extreme values than the coordinates of the exemplars of the category *fles*, indicated with a black diamond in Figure 2. It seems that people are using a caricature (more extreme on the dimensions than the usual typical objects) to judge typicality.

For *bus*, the optimal correlation reached -0.65, again higher than the correlations between typicality and distance to the centroids of *bus* (see Table 1 and 2, marginally significant, $p < .1$). Like for *fles*, the optimal point for *bus*, *BUS_opt*, seems to be a caricature of the category with more extreme values on the two dimensions (see Figure 2) than the typical exemplars, indicated by a white square in Figure 2. Also for MDS solutions in 3- to 6-dimensions, the highest correlations for *fles* and *bus* were obtained with optimal points located far outside the set of objects (e.g. a correlation of -0.95 for the 5-dimensional solution of *fles*).

For *pot* and *brik*, the optimal correlations reached -0.73 and -0.68. Again these values exceed the values found for the centroids of *pot* and *brik* in two dimensions (see Table 1 and 2). This time, the optimal points, *POT_opt* and *BRIK_opt* are not situated outside the cloud of objects. However, compared to the position of their respective centroids *POT_ct* and *BRIK_ct*, they do move somewhat away from the object set. The mean distance of the objects to the optimal point POT_opt is significantly larger than the mean distance of the objects to the centroid POT_ct ($t(72) = 8.10$, $p < .0001$). The same holds for *brik* ($t(72) = 19.49$, $p < .0001$). Although the optimal prototypes for *pot* and *brik* do not seem to occupy an extreme position away from the complete set of objects, both POT_opt and $BRIK\_opt$ lie at the boundary of the convex hull that covers the exemplars of their category. Similar results were obtained for MDS representations in 3- to 6-dimensions.

A plausible reason for the fact that POT_opt and $BRIK\_opt$ did not move to an extreme position, while $FLES\_opt$ and BUS_opt did so, might be the difference in representativeness of the object set for the different categories. There are two indications for this difference in representativeness. First, overall, the set of objects is significantly more typical for the categories *fles* and *bus* than for the categories *pot* and *brik*. This can be seen in the fact that there are more objects with *fles* ($n = 25$) or *bus* ($n = 16$) as the dominant name than there are objects with *pot* ($n = 13$) or *brik* ($n = 4$) as the dominant name. Second, for *fles* and *bus*, the stimulus set represents the complete continuum of typicality, ranging from very good exemplars to intermediate to very bad exemplars. In contrast, for *pot* or *brik*, the stimuli in the set are either extremely good exemplars, or extremely bad exemplars with no intermediate cases. For instance, the few objects with *brik* as the dominant name receive extremely high mean typicality ratings (average of 5.83 on a scale of 1 to 7, SD = 0.48), while the other 69 objects -the nonmembers- receive extremely low mean typicality ratings (1.68 on average, SD = 0.79). This means that the four exemplars of the category *brik* will have a relatively large influence on the computation of the optimized correlation, in contrast to the exemplars of the contrast categories (i.e. all 69 objects named by a dominant name different from *brik*). The optimization algorithm will focus on the four objects called *brik* and try to find the point for which the distances to these four objects are minimal. This results in a point that comes very close to the central tendency of the four objects, since this point guarantees that the squared deviations from it are minimal. A similar reasoning explains why the optimal point for *pot* does not move farther away from the cloud of data points.

For *fles* and *bus*, the mean typicality ratings of the nonmembers are significantly larger than the means of the nonmembers for *pot* and *brik* (1.91 for *fles* and 1.96 for *bus* versus 1.68 for *pot* and *brik*, $t(288) = 3.71$, $p < .01$). These larger mean typicality ratings of nonmembers have an extreme influence on the position of the optimal point: They 'catapult' the point away from the centroid that is fully determined by the position of the members.

Figure 3 visualizes the effect of the mean typicality of the nonmembers of a category on the position of the optimal point relative to the centroid. The mean typicality ratings of nonmembers for the four categories are plotted against the observed distances between the optimal point and the centroid. The (regression) line connects the distances predicted by the mean typicality of nonmembers for the different categories.

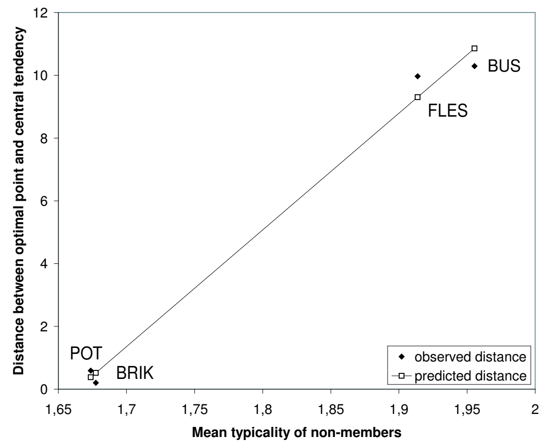

Figure 3: Plot of mean typicality ratings of nonmembers for *fles*, *bus*, *pot* and *pot* against observed distances between optimal point and central tendency and distances predicted by the mean typicality of nonmembers.

As can be seen from Figure 3, the mean typicality ratings of nonmembers do a very good job in predicting the shift of the optimal prototype. The larger the mean typicality rating of nonmembers, the larger is the shift away from the central tendency. However, the number of data points ($n = 4$) on which this regression is based, is too small to draw reliable conclusions. Many more categories are needed to test this regression effect. Also the possible additional contribution of other factors (e.g. the mean typicality of the nonmembers *relative* to the mean typicality of members, that might explain the somewhat larger shift for *pot* compared to *brik* and for *bus* compared to *fles*) can only be tested on a larger set of data. Further analyzes are planned with the typicality data of French-speaking monolinguals and French-Dutch bilinguals in order to get a more considerable data set size and to allow us to test these effects in a more reliable way.

A set of 'more typical' nonmembers for *brik* and *pot* would probably lead to an optimal point outside the cloud of points. However, in real life, people might also exclusively encounter very typical members and very atypical nonmembers of *brik* and *pot*. Objects of these categories have a specific shape determined by only a few covarying features (e.g. for *brik*: rectangular shape, specific proportion of height to width; for *pot*: specific proportion of height to width). Intermediate positions (atypical members and typical nonmembers) are rather rare. In contrast, for *fles* and *bus* there does exist a considerable group of objects occupying the intermediate positions on the continuum of typicality. So, the position of the optimal prototype depends on the degree of graded structure inherent to the category, which is expressed in the typicality of nonmembers.

Conclusion

The present study examined different determinants of the graded structure of categories. We found that geometrical centroids (unweighted and weighted for name frequencies) in higher-dimensional spaces did a better job in predicting typicality than family resemblance. However, an even better prediction was obtained with an external prototype whose position was based on the typicality ratings of nonmembers of the category. Further research is required to examine whether these findings can be generalized to other categories.

Acknowledgments

The first author is a research assistent of the Fund for Scientific Research-Flanders.

References

Ameel, E., Storms, G., Malt, B., & Sloman, S. A. (in press). How bilinguals solve the naming problem. *Journal of Memory and Cognition*.

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*(4), 629-654.

Borg, I., & Groenen, P. (1997). *Modern Multidimensional Scaling: Theory and applications.* New York: Springer-Verlag, New York.

Gardenfors, P. (2000). *Conceptual spaces: The geometry of mind.* Bradford Books: MIT Press.

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory and Cognition, 24*(5), 608-628.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 441-461.

Hampton, J. A. (1993). Prototype models of concept representations. In I. Van Mechelen, J. A. Hampton, R. S. Michalski & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis.* London: Academic Press

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language, 40*, 230-262.

Minda J.P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 755-799.

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. (1992). Exemplars, prototypes, and similarity rules. In A. Healy, S. Kosslyn, and R. Shiffrin (Eds.), *From learning theory to connectionist theory: Essays in honor of William K. Estes, Volume 1.* Hillsdale, NJ: Erlbaum.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Jorunal 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.* New York: Academic Press.

Rosch, E. H. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization.* Erlbaum, Hillsdale, NJ.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*, 214-241.