Memory & Cognition
2010, 38 (7), 962-968
doi:10.3758/MC.38.7.962

Geometric and featural representations in semantic concepts

Wolf Vanpaemel, Timothy Verbeemen, and Matthew Dry
University of Leuven, Leuven, Belgium

Tom Verguts
Ghent University, Ghent, Belgium

and

Gert Storms
University of Leuven, Leuven, Belgium

We explore the adequacy of two types of similarity representation in the context of semantic concepts. To this end, we evaluate different categorization models, assuming either a geometric or a featural representation, using categorization decisions involving familiar and unfamiliar foods and animals. The study aims to assess the optimal stimulus representation as a function of the familiarity of the stimuli. For the unfamiliar stimuli, the geometric categorization models provide the best account of the categorization data, whereas for the familiar stimuli, the featural categorization models provide the best account. This pattern of results suggests that people rely on perceptual information to assign an unfamiliar stimulus to a category but rely on more elaborate conceptual knowledge when assigning a familiar stimulus.

One of the fundamental problems for cognitive science is understanding how people mentally represent stimuli. Many well-established approaches for determining stimulus representations from empirical data rely on similarity between pairs of stimuli. Two of the most developed and widely used approaches for deriving a stimulus representation based on similarity are the geometric and featural approaches. In the geometric approach (Shepard, 1957), a stimulus is represented as a point in a multidimensional coordinate space, corresponding to continuous values on an (often limited) number of dimensions. The similarity between two stimuli is inversely related to the distance between the points in the psychological space, implying that similar stimuli lie close together, whereas dissimilar stimuli lie far apart. Geometric representations can be generated from proximity data using a variety of multidimensional scaling (MDS) algorithms (see, e.g., Borg & Groenen, 2005; Carroll & Arabie, 1980; Lee, 2001).

In the featural approach (Shepard & Arabie, 1979; Tversky, 1977), a stimulus is represented in terms of the presence or absence of a set of discrete, often binary features. The similarity between two stimuli is a function of feature overlap, whereby common features increase overall similarity, and distinctive features decrease overall similarity. Commonly used methods for extracting featural representations from proximity data include additive clustering and additive tree algorithms (e.g., Lee, 2002; Navarro & Griffiths, 2008; Tenenbaum, 1996; Van Mechelen & Storms, 1995).

It has been argued that different approaches are suitable for different stimulus domains (see, e.g., Tversky, 1977). The present study addresses the relative appropriateness of the geometric and featural approaches, focusing on familiar and unfamiliar stimuli in semantic concepts.

EVALUATING GEOMETRIC AND FEATURAL REPRESENTATIONS

One straightforward way to investigate the appropriateness of the different approaches to stimulus representation is to directly compare both approaches on their ability to account for empirically observed similarity data. For example, Pruzansky, Tversky, and Carroll (1982) reanalyzed 20 previously published similarity data sets, dividing them into two groups depending on the nature of the stimuli: conceptual (e.g., vegetables) or perceptual (e.g., polygons). For 10 out of 11 studies involving conceptual stimuli, similarity data proved to be best accounted for by a featural representation. In contrast, 7 out of 9 studies involving perceptual stimuli showed a clear advantage for a (low-dimensional) geometric representation. It was concluded that the best-suited representation for high-level conceptual stimuli is a featural one, whereas low-level perceptual stimuli are best represented using geometric

W. Vanpaemel, wolf.vanpaemel@psy.kuleuven.be

© 2010 The Psychonomic Society, Inc.

similarity. A similar observation, but phrased in terms of criteria related to nearest-neighbor analyses rather than fit criteria, was made by Tversky and Hutchinson (1986), based on 100 previously published similarity data sets.

A second way to gain insight into the appropriateness of both approaches stems from the fact that models of mental stimulus representation not only have some status as models of human conceptual structure in and of itself, but also underlie higher level models of cognition, decision making, and perception. For example, most models of categorization rely on an underlying stimulus representation to account for categorization decisions (e.g., ALCOVE, Kruschke, 1992; GCM, Nosofsky, 1986; SUSTAIN, Love, Medin, & Gureckis, 2004). Therefore, in the context of categorization, the appropriateness of the geometric and featural approaches can also be assessed by comparing different categorization models, relying on either a geometric or a featural representation, on their ability to account for empirically observed categorization data. A representation model can provide an excellent account of similarity data, but a categorization model relying on this representation can still fail dramatically in accounting for categorization data that are modeled on top of the representation, a point most effectively demonstrated by Lee and Navarro (2002), focusing on ALCOVE.

Similarity Representation and Categorization

The present article relies on this second strategy to gain insight into the relative appropriateness of geometric and featural representations. Using categorization models to assess the suitability of geometric and featural representations requires the comparison of a categorization model relying on a geometric representation with a categorization model relying on a featural representation. Many existing models of category learning differ in whether they assume a geometric or a featural representation, but these models tend to differ in more aspects than simply their stimulus representation, blurring the comparison of interest. Therefore, to reach our goal, we start from existing models of categorization that rely on a geometric representation and then develop their featural counterpart. Keeping every assumption but the stimulus representation constant assures the fairest comparison between the different stimulus representations.

In constructing the featural versions of the geometric categorization models, we adopt a similar approach to Lee and Navarro (2002), who developed a featural version of ALCOVE. To increase the generality of our investigation, we apply this strategy to two different, widely studied geometric categorization models, the generalized context model (GCM; Nosofsky, 1986) and the geometric prototype model (GPM; Reed, 1972, also referred to as the MDS-based prototype model).

Semantic Concepts

In contrast to Lee and Navarro (2002), who relied on artificial stimuli and categories, we focus on semantic concepts, such as fruits and vegetables. One consequence of considering semantic concepts is that it is possible to make a distinction between familiar and unfamiliar stimuli. Stimuli that are artificially created by the experimenter are, by design, unfamiliar to the observer. In contrast, when focusing on semantic concepts, stimuli can differ in the extent to which an observer has prior, real-world knowledge about them. For instance, about *apple*, a Western observer knows where they grow, what they taste like, what they look like, and so on, but there are other natural world items, such as exotic tropical fruits, that can be highly unfamiliar to the observer. Previous research by Smits, Storms, Rosseel, and De Boeck (2002) has shown that the relative success of exemplar and prototype models may depend on the degree to which the stimuli are familiar. They showed that the advantage of the GCM over the GPM was larger for familiar stimuli than for unfamiliar stimuli.

Outline

In this article, we provide a systematic evaluation of geometric and featural representations based on the ability of categorization models to account for categorization decisions involving familiar and unfamiliar stimuli from semantic concepts. First, we provide a formal description of the four different models to be considered, two of which are existing, geometric categorization models, and two are their newly constructed featural counterparts. Next, we describe two previously published data sets involving the categorization of familiar and unfamiliar stimuli from semantic concepts, involving two semantic domains, food and animals. These data sets are then used to assess the relative merits of the geometric and featural categorization models separately for the familiar and the unfamiliar stimuli. The model comparisons suggest that familiar stimuli can be better represented using a featural approach than using a geometric approach, and that visually presented, unfamiliar stimuli can be better captured with a geometric representation, using a limited number of underlying dimensions. Finally, we relate our findings to previous findings concerning geometric and featural representations and discuss future work.

MODELS

Two similarity-based views on category learning have been widely studied: the exemplar view and the prototype view (see Nosofsky, 1992; Vanpaemel & Storms, 2010, for overviews of empirical comparisons). The exemplar view assumes that a category is represented by all of the previously encountered category members, the exemplars. The prototype view assumes that a category is represented by an abstracted summary, reflecting the central tendency of the category. We first focus on existing MDS-based exemplar and prototype models and then modify these to featural models.

Geometric and Featural Exemplar Models

Probably the most successful exemplar model is the GCM (Nosofsky, 1984, 1986). It assumes that a categorization decision about a stimulus is driven by the similarity of the stimulus to the relevant categories. In particular, in the case of two categories, A and B, the GCM assumes

that the probability of stimulus x being classified in Category A is given by:

$$P(\mathrm{A} \mid x) = \frac{\beta s(x, \mathrm{A})}{\beta s(x, \mathrm{A}) + (1-\beta) s(x, \mathrm{B})}, \quad (1)$$

where $0 < \beta < 1$ serves as a response bias parameter toward Category A, and $s(x, \mathrm{A})$ denotes the similarity of stimulus x toward Category A. The more similar a stimulus is considered to Category A, the more likely it is to be endorsed in that category. Being an exemplar model, the similarity of stimulus x toward Category A is defined by the sum of the similarity of stimulus x toward all Category A exemplars:

$$s(x, \mathrm{A}) = \sum_{j \in \mathrm{A}} s(x, j). \quad (2)$$

In the GCM, the similarity between two stimuli x and j, $s(x, j)$, is computed from a geometric stimulus representation,[1] in which x and j are represented as points in a D-dimensional space with coordinates $x = (x_1, \ldots, x_D)$ and $j = (j_1, \ldots, j_D)$. The similarity between two stimuli is assumed to be an exponentially decreasing function of the distance between the stimuli, indicating that similar stimuli lie close together and dissimilar stimuli lie far apart:

$$s(x, j) = \exp[-c\, d(x, j)], \quad (3)$$

where $0 < c < \infty$ is a scaling parameter and $d(x, j)$ is the Minkowski distance between x and j, given by:

$$d(x, j) = \left(\sum_k w_k \mid x_k - j_k \mid^r \right)^{1/r}, \quad (4)$$

where w_k denotes the weight of the kth dimension, and all D weights are restricted to sum to 1. The distance between two stimuli increases with the number of dimensions where there is a mismatch, as well as with the size of the mismatch. Further, these mismatches are weighted by the importance of a particular dimension in a given categorization task. The metric r is usually given a value of either 1 or 2, corresponding to a city-block or a Euclidean distance, respectively.

Rather than relying on a geometric representation to compute the similarities in Equation 2, as the GCM assumes, a featural exemplar model (FEM) computes these similarities starting from a featural stimulus representation, in which x and j are represented as vectors of m binary variables $x = (x_1, \ldots, x_m)$ and $j = (j_1, \ldots, j_m)$, where $x_k = 1$ if x possesses the kth feature, and $x_k = 0$ otherwise. Following Lee and Navarro (2002), we assume that similarities can be computed from featural distances, implying that Equation 3 is sensible also in the context of featural representations. Further, rather than on the geometric distance of Equation 4, the FEM relies on the following featural distance:

$$d(x, j) = \sum_k w_k x_k \left(1 - j_k\right) + \sum_k w_k \left(1 - x_k\right) j_k. \quad (5)$$

Like the geometric distance used in the GCM, the featural distance used in the FEM is sensitive to mismatches only—that is, the number of features on which stimuli match does not affect the distance.

Geometric and Featural Prototype Models

The GPM (Minda & Smith, 2001; Nosofsky, 1987; Reed, 1972; Vanpaemel & Storms, 2008) is to a large extent similar to the GCM. Like the GCM, it assumes that categorization is driven by a similarity comparison of the stimulus to the category (Equation 1) and assumes a geometric representation (Equations 3 and 4). The crucial difference with the GCM is how a category, and hence the similarity of a stimulus to a category, is defined. Rather than summing the similarity to all individual exemplars (Equation 2), the similarity of stimulus x toward Category A is defined by the similarity of x toward one representative item, the Category A prototype π_{A}:

$$s\left(x, \mathrm{A}\right) = s\left(x, \pi_{\mathrm{A}}\right). \quad (6)$$

Although different proposals have been made to define a prototype, the most common definition of π_{A} is the average of all n Category A members:

$$\pi_{\mathrm{A}k} = \frac{1}{n} \sum_{j \in \mathrm{A}} j_k, \text{ for } k = 1, \ldots, D. \quad (7)$$

Given a featural representation of the stimuli, a featural prototype model (FPM) can be constructed by Equations 1, 3, 5, and 6, where the prototype π_{A} of Equation 6 is defined by the modal values of the features of all Category A members.[2]

DATA

We considered two previously published data sets involving food items (fruits and vegetables; Smits et al., 2002) and animal items (carnivores and herbivores; Verbeemen, Vanpaemel, Pattyn, Storms, & Verguts, 2007). We briefly summarize the stimuli and data collection procedures, but refer the interested readers to these publications for a description in full detail.

Smits et al. (2002) used a stimulus set consisting of pictures of 109 fruits and vegetables. Of these 109 stimuli, 35 fruit stimuli and 44 vegetable stimuli were obtained using an exemplar generation task and can therefore be considered as highly familiar, well-known stimuli, such as *carrot*. The remaining 30 stimuli were mostly exotic items, such as *rambutan*, purchased in specialty shops that import food from Central African and Southeast Asian countries. A group of 30 participants classified the familiar stimuli as belonging to either fruits or vegetables, whereas a different group of 20 participants did the same for the unfamiliar stimuli.

Similarity ratings were computed using the following procedure.[3] A first group of participants generated features for the categories fruits and vegetables. The 10 most frequently generated features for both categories were selected. However, 3 features occurred in both lists. Consequently, a set of 17 different features was used for a feature applicability judgment task, completed by a second group of participants. From these feature applicability vectors, a matrix containing all pairwise similarities was obtained by correlating these vectors for all 109 stimuli, after summing over participants.

Verbeemen et al. (2007) used a stimulus set consisting of pictures of 101 carnivores and herbivores. Of these 101 stimuli, 35 carnivores and 38 herbivores were again obtained from an exemplar generation task for each of the categories, implying that these were familiar, well-known category members, such as *cow*. The remaining 28 stimuli were exotic, unknown animals, such as *indri*. A group of 32 participants classified all stimuli as being carnivores or herbivores. As in Smits et al. (2002), another group of participants was asked to generate features, but contrary to the procedure followed by Smits et al. (2002), the features were generated at the level of the exemplars, rather than at the category level. A final group of participants rated the applicability for each of the 39 most frequently generated features to each of the 101 animals. A similarity matrix was then obtained by correlating the summed feature applicability vectors for all 101 stimuli.

EVALUATING THE GEOMETRIC AND FEATURAL MODELS OF CATEGORIZATION

All four categorization models (GCM, FEM, GPM, and FPM) were evaluated on both data sets. To assess the appropriateness of the two approaches to similarity representation as a function of stimulus familiarity, categorization responses to well-known, familiar stimuli were analyzed separately from categorization responses to unknown, unfamiliar stimuli. Of interest is the comparison of the relative performance of the GCM versus the FEM and of the GPM versus the FPM.

Generating the Stimulus Representations

Smits et al. (2002) and Verbeemen et al. (2007) used the (computed) pairwise similarities to derive geometric representations of all 109 items of the food domain and all 101 items of the animal domain, respectively. Based on the stress profile, a three-dimensional MDS solution was chosen for both the food domain and the animal domain. These two previously generated geometric representations were employed in the present article.

Rather than using the features generated by the participants as the featural representations, featural representations were generated using the same computed pairwise similarities that were used to generate the geometric representations, in order to compare the geometric- and feature-based approaches as accurately as possible. The featural representation is thus not restricted to have as many features as or fewer features than the original feature vectors. The procedure to derive the featural representation (and the geometric representation as well) is seen as a data reduction technique, implying that the features (or dimensions) that make up the representation do not need to correspond to the features used to compute the pairwise similarities.

To derive a featural stimulus representation, for each domain, cluster solutions with 12 to 26 clusters were computed using the additive clustering algorithm described in Lee (2002). For all representations, the universal cluster (the feature that applies to all stimuli) was excluded since it cannot add to any distances between stimuli and is redundant in the categorization model, leaving a set of cluster solutions with 11 to 25 effective clusters. Note that additive clustering focuses on common features only, whereas the featural distance employed in the featural categorization models (Equation 5) is based on distinctive features. Although contradictory at face value, these choices were inspired by Gati and Tversky's (1984) proposal that in categorization decisions, dissimilarities are weighted more heavily, whereas in similarity decisions, common features tend to be weighted more heavily (see Lee & Navarro, 2002, for an in-depth discussion).

Evaluating the Geometric and Featural Categorization Models

Each geometric model (GCM and GPM) had four free parameters for each data set: the bias parameter β, the sensitivity parameter c, and two free dimension weights, because dimension weights are restricted to add to 1. The metric parameter r was not treated as a free parameter but was fixed at $r = 2$, because the dimensions identified by the MDS procedure are most likely integral rather than separable. Each featural model (FEM and FPM) relying on q effective clusters had $q+1$ free parameters: β, c, and $q-1$ free feature weights, because the weights are restricted to sum up to 1.

All models were evaluated using the Bayesian information criterion (BIC; Schwarz, 1978), which is defined by $\mathrm{BIC} = -2 \ln \mathrm{L} + k \ln n$, where ln L is the maximized log likelihood, k is the number of free parameters, and n is the number of data points. The BIC increases whenever either the lack of fit ($-2 \ln \mathrm{L}$) increases or the parametric complexity (k) increases. Because the goal is to select a model that fits the data well and at the same time is low in complexity, the model with the smallest BIC is to be preferred. In this sense, BIC balances goodness-of-fit and model complexity. An additional advantage of BIC is that there exists a useful framework within which to evaluate the significance of magnitude differences in BIC values across models. For two models A and B, the difference in BIC, $\Delta\mathrm{BIC}_{\mathrm{BA}} = \mathrm{BIC}_{\mathrm{B}} - \mathrm{BIC}_{\mathrm{A}}$, approximates twice the logarithm of the Bayes factor ($2 \ln \mathrm{BF}_{\mathrm{AB}}$) for comparing A and B, for which Kass and Raftery (1995, p. 777) have suggested an interpretative scale of evidence. In particular, the evidence in favor of Model A (or, equivalently, against Model B) ranges from *not worth more than a bare mention* (for $2 \ln \mathrm{BF}_{\mathrm{AB}}$ in the range 0–2), to *positive* (2–6), to *strong* (6–10), to *very strong* (>10).

Figure 1 depicts $\Delta\mathrm{BIC} = \mathrm{BIC}(\text{featural}) - \mathrm{BIC}(\text{geometric})$ for the geometric and featural models, across all 15 cluster solutions considered. Positive values indicate evidence for the geometric models (GCM and GPM, respectively), whereas negative values indicate evidence in favor of the featural models (FEM and FPM, respectively). The crosses ($\times$) refer to the data concerning the familiar stimuli, whereas the asterisks ($*$) refer to the data concerning the unfamiliar stimuli. Panels A and C depict the results for the food domain and panels B and D depict the results

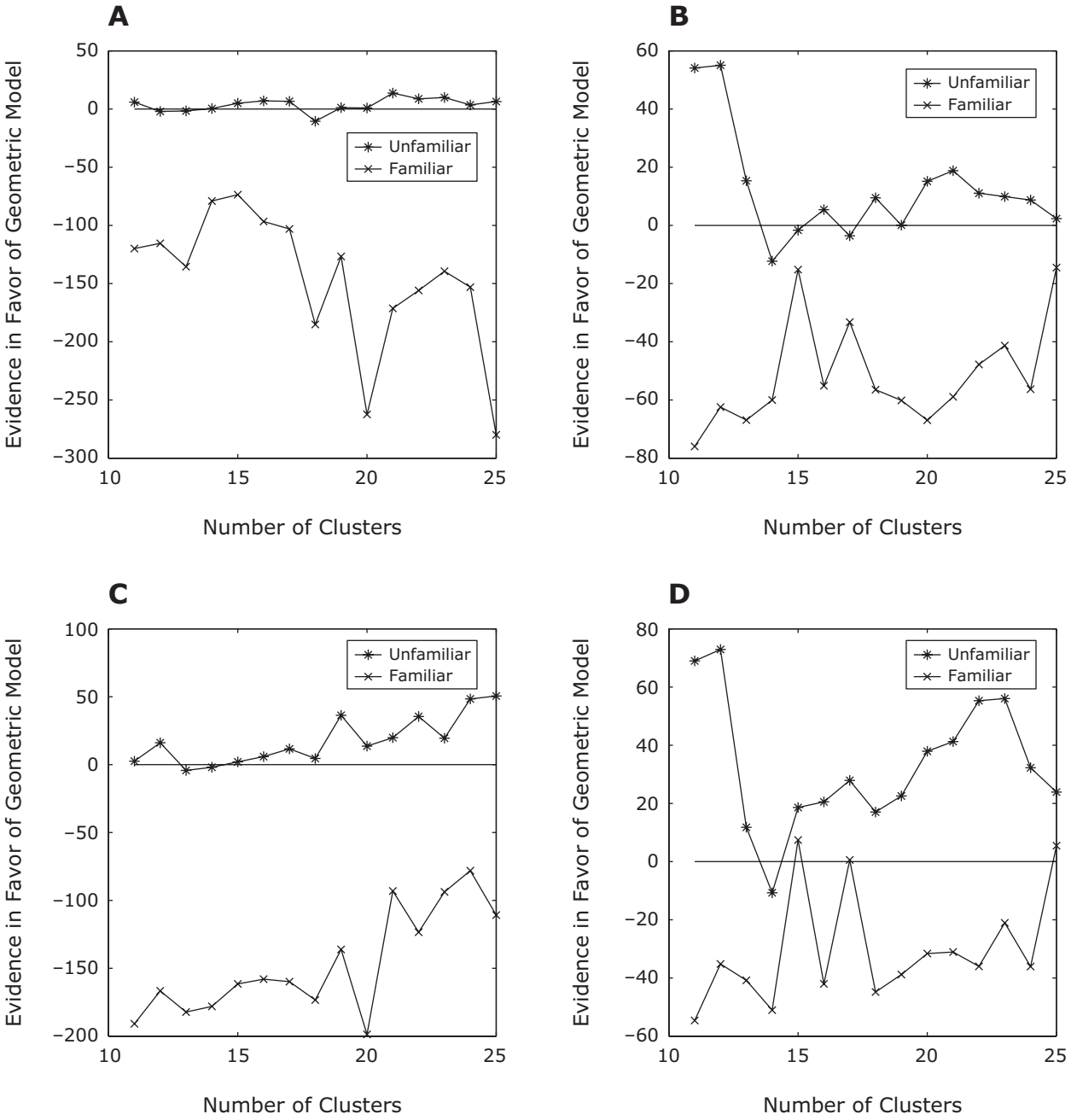


Figure 1. Empirical comparison of geometric and featural categorization models across 15 cluster solutions. Panel A compares geometric and featural exemplar models in the food domain, panel B compares geometric and featural exemplar models in the animal domain, panel C compares geometric and featural prototype models in the food domain, and panel D compares geometric and featural prototype models in the animal domain. The crosses (×) refer to the data concerning the familiar stimuli, whereas the asterisks (∗) refer to the data concerning the unfamiliar stimuli. Positive values indicate evidence for the geometric models; negative values indicate evidence for the featural models.

for the animal domain, for the exemplar models (GCM and FEM) and the prototype models (GPM and FPM), respectively.

Figure 1 shows a surprisingly consistent pattern. Focusing on the exemplar models (first row), there is overwhelming evidence for a featural representation when familiar stimuli are considered for both the food (panel A) and the animal (panel B) domains, across all 15 cluster solutions. As far as the unfamiliar stimuli are concerned, the evidence is somewhat less strong, but for all but 3 of the 15 cluster solutions, the geometric representation is best supported by the data, again for both domains. The same general picture emerges for the prototype models (second row). Again, the featural representations are

overwhelmingly supported for all (in the food domain, panel C) or nearly all (in the animal domain, panel D) cluster solutions when the familiar stimuli are considered. Model preference changes dramatically for the unfamiliar stimuli: two (in the food domain) or one (in the animal domain) solutions aside, the geometric models are very strongly supported.

DISCUSSION

The goal of the present study was to assess the appropriateness of geometric versus featural representations as a function of stimulus familiarity. Rather than evaluating different representations directly using similarity data, we evaluated different categorization models, relying on different representations, using categorization data involving semantic concepts (fruits and vegetables; carnivores and herbivores). Four categorization models were evaluated on their ability to account for the categorization decisions relating to either unfamiliar (e.g., *guaga* or *genetta*) or familiar (e.g., *broccoli* or *squirrel*) items. The results indicated that categorization of unfamiliar stimuli was best accounted for by models assuming a geometric stimulus representation, whereas the categorization of familiar stimuli was best captured by models that assume a featural stimulus representation. Encouragingly, these results not only hold for a large number of cluster solutions, but also across two different semantic domains and across the prototype versus exemplar distinction.

Relation to Earlier Findings

As noted in the introduction, previous work focusing on how geometric and featural representations account for similarity data has highlighted that featural representations are most appropriate for conceptual domains, whereas geometric representations are better suited for perceptual domains (Pruzansky et al., 1982; Tversky & Hutchinson, 1986). Our findings can be understood in the light of these earlier findings by considering the notion of familiarity.

Although the distinction between familiar and unfamiliar stimuli and the exact nature of what determines familiarity in this context is most likely very subtle and needs to be studied more thoroughly, it seems reasonable to assume that people have elaborated conceptual knowledge of familiar stimuli (like where the animals live, what they look like, and how they behave, or where the fruits grow, how they taste, and the way they are eaten or prepared) and that such conceptual knowledge is lacking for unfamiliar ones. Consequently, the very least one could say is that familiar and unfamiliar stimuli vary strongly in terms of the amount and type of information available to base categorization decisions on. When trying to categorize an unknown, novel item (e.g., a piece of fruit seen in a marketplace in a foreign tropical country), all the categorizer can do is combine a few perceptual cues such as shape, texture, color, smell, and taste. (Note that in pictured stimuli, such as the stimuli used in this article, the latter two cues are not available.) When trying to categorize a well-known, familiar item (such as a piece of fruit seen in the local produce store), the categorizer has a vast knowledge base to rely on when categorizing. This connection between unfamiliarity and perceptual information, on the one hand, and familiarity and conceptual information, on the other hand, ties together the present findings with the previous findings based on similarity analyses.

In sum, our finding fits nicely with the idea of a gradual increase in knowledge: An observer who knows little about an unfamiliar semantic stimulus can only judge it using surface perceptual information, a process that is best captured by geometric representations. As the observer becomes increasingly acquainted with a stimulus, knowledge about the semantic stimulus grows. Categorization then becomes increasingly based on conceptual rather than perceptual information, which is best captured by featural representations.

Future Work

Despite the clarity and interpretability of the observed pattern in the model comparison results, people most likely use a mixture of both conceptual and perceptual information for categorizing both unfamiliar and familiar semantic stimuli. Interestingly, recent work on stimulus representation has started to explore hybrid representations, incorporating both a geometric and a featural representation (Lee & Navarro, 2005; Navarro & Lee, 2003). In such a hybrid representation, a stimulus is allowed to take continuous values on a number of dimensions and to have a number of discrete features at the same time. Assessing the performance of categorization models based on such hybrid representations is an important goal for future research.

AUTHOR NOTE

This project was in part sponsored by Grant FWO G.0513.08 from the Fund for Scientific Research, Flanders, Belgium, to G.S. We thank Rob Nosofsky, Dan Navarro, Michael Lee, and several anonymous reviewers for their comments on previous versions of this article. Address correspondence to W. Vanpaemel, Department of Psychology, University of Leuven, Tiensestraat 102, B3000 Leuven, Belgium (e-mail: wolf.vanpaemel@psy.kuleuven.be).

REFERENCES

Borg, I., & Groenen, P. J. F. (2005). *Modern multidimensional scaling: Theory and applications* (2nd ed.). New York: Springer.

Carroll, J. D., & Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology*, **31**, 607-649.

Dry, M. J., & Storms, G. (2009). Similar but not the same: A comparison of the utility of directly rated and feature-based similarity measures for generating spatial models of conceptual data. *Behavior Research Methods*, **41**, 889-900.

Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology*, **16**, 341-370.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, **90**, 773-795.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44.

Lee, M. D. (2001). Determining the dimensionality of multidimensional scaling models for cognitive modeling. *Journal of Mathematical Psychology*, **45**, 149-166.

Lee, M. D. (2002). Generating additive clustering models with limited stochastic complexity. *Journal of Classification*, **19**, 69-85.

Lee, M. D., & Navarro, D. J. (2002). Extending the ALCOVE model of

category learning to featural stimulus domains. *Psychonomic Bulletin & Review*, **9**, 43-58.

LEE, M. D., & NAVARRO, D. J. (2005). Minimum description length and psychological clustering models. In P. D. Grünwald, I. J. Myung, & M. A. Pitt (Eds.), *Advances in minimum description length: Theory and applications* (pp. 355-384). Cambridge, MA: MIT Press.

LOVE, B. C., MEDIN, D. L., & GURECKIS, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, **111**, 309-332.

MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.

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, & Cognition*, **27**, 775-799.

NAVARRO, D. J., & GRIFFITHS, T. L. (2008). Latent features in similarity judgments: A nonparametric Bayesian approach. *Neural Computation*, **20**, 2597-2628.

NAVARRO, D. J., & LEE, M. D. (2003). Combining dimensions and features in similarity-based representations. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in neural information processing systems* (pp. 67-74). Cambridge, MA: MIT Press.

NOSOFSKY, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 104-114.

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, & Cognition*, **13**, 87-108.

NOSOFSKY, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *Essays in honor of William K. Estes, Vol. 1: From learning theory to connectionist theory* (pp. 149-167). Hillsdale, NJ: Erlbaum.

PRUZANSKY, S., TVERSKY, A., & CARROLL, J. D. (1982). Spatial versus tree representations of similarity data. *Psychometrika*, **47**, 3-24.

REED, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, **3**, 382-407.

SCHWARZ, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.

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

SHEPARD, R. N., & ARABIE, P. (1979). Additive clustering: Representation of similarities as combinations of discrete overlapping properties. *Psychological Review*, **86**, 87-123.

SMITS, T., STORMS, G., ROSSEEL, Y., & DE BOECK, P. (2002). Fruits and vegetables categorized: An application of the generalized context model. *Psychonomic Bulletin & Review*, **9**, 836-844.

TENENBAUM, J. B. (1996). Learning the structure of similarity. In D. S. Toretzky, M. C. Mozer, & M. Hasselmo (Eds.), *Advances in neural information processing systems* (pp. 3-9). Cambridge, MA: MIT Press.

TVERSKY, A. (1977). Features of similarity. *Psychological Review*, **84**, 327-352.

TVERSKY, A., & HUTCHINSON, W. (1986). Nearest neighbor analysis of semantic spaces. *Psychological Review*, **93**, 3-22.

VAN MECHELEN, I., & STORMS, G. (1995). Analysis of similarity data and Tversky's contrast model. *Psychologica Belgica*, **35**, 85-102.

VANPAEMEL, W., & STORMS, G. (2008). In search of abstraction: The varying abstraction model of categorization. *Psychonomic Bulletin & Review*, **15**, 732-749.

VANPAEMEL, W., & STORMS, G. (2010). Abstraction and model evaluation in category learning. *Behavior Research Methods*, **42**, 421-437.

VERBEEMEN, T., VANPAEMEL, W., PATTYN, S., STORMS, G., & VERGUTS, T. (2007). Beyond exemplars and prototypes as memory representations of natural concepts: A clustering approach. *Journal of Memory & Language*, **56**, 537-554.

NOTES

1. The model has been termed *generalized context model* because it generalizes Medin and Schaffer's (1978) successful context model. Alternative names that have been used occasionally include *MDS-based context model* and *MDS-based exemplar model*. For the present purpose, it can be regarded as the *geometric exemplar model*.

2. The prototype is defined by the mode rather than by the average, as in Equation 7, since taking the average over binary features can result in a prototype that is not featural itself, in the sense that it takes nonbinary values. In our analyses, using an average prototype in the FPM leads to results quantitatively very similar and qualitatively identical to those from using a modal prototype.

3. This procedure differs from the commonly used procedure of obtaining pairwise similarity ratings, which involves directly asking the participants to rate the similarity between any pair of stimuli in the domain. Taking this approach would result in 5,886 or 5,050 pairs to be judged, for the food and animal domains, respectively, a number too large to handle in any reasonable time. The procedure followed in this article, which starts from an exemplar by feature matrix, is a more feasible one, given the large number of stimuli involved. Additionally, the procedure has been shown to be superior to the more standard direct rating procedure in the context of semantic concepts. In particular, Dry and Storms (2009) showed that similarities derived from feature applicabilities were noticeably superior to directly rated similarities, both in terms of goodness of fit and in terms of predictions of empirical measures of graded category structure such as typicality ratings, exemplar generation frequencies, and response times in a speeded categorization task.

(Manuscript received March 30, 2009;
revision accepted for publication March 9, 2010.)