Psychonomic Bulletin & Review
2002, 9 (4), 836-844

Fruits and vegetables categorized: An application of the generalized context model

TIM SMITS, GERT STORMS, YVES ROSSEEL, and PAUL DE BOECK
University of Leuven, Leuven, Belgium

In the study reported in this paper, we investigated the categorization of well-known and novel food items in the categories *fruits* and *vegetables*. Predictions based on Nosofsky's (1984, 1986) generalized context model (GCM), on a multiplicative-similarity prototype model, and on an instantiation model as applied in Storms, De Boeck, and Ruts (2001) were compared. Despite suggestions in the literature that prototype models predict categorization from large categories better than exemplar models do, our results showed that the exemplar-based GCM yielded clearly better predictions than did a (multiplicative-similarity) prototype model.

Review articles about psychological research in the area of concepts and categories (e.g., Komatsu, 1992; Medin, 1989; Medin & Smith, 1984; E. E. Smith & Medin, 1981) usually describe two large domains of studies that are rather far apart. Issues typically addressed in the first domain are how existing lexicalized concepts (e.g., *birds*, *sports*, *furniture*, etc.) are represented in semantic memory and which mental processes operate on this information (see, e.g., Hampton, 1979, 1993; Murphy & Medin, 1985; Rosch, 1975, 1978; Rosch & Mervis, 1975). It has almost invariably been assumed that natural language concepts are represented through attributes associated with the concepts and possessed by (some of) the elements for which the concept is used—that is, by some of its extension (e.g., prototype theory; see Hampton, 1993, and Rosch, 1978).

In the second domain, issues such as how people learn new categories and how they decide when a category applies are addressed. In a typical category-learning experiment, participants learn to discriminate between two new artificial categories through the presentation of exemplars of both categories. After this learning phase, they are presented with new items in a transfer phase, and representational assumptions and process information are derived from their categorization decisions. Research in this domain has resulted in very good predictions from exemplar models (e.g., Medin & Schaffer, 1978; Nosofsky, 1992). These exemplar models reject the idea of abstract, summary representations of concepts and, instead, assume that people encode information about specifically encountered exemplars. Although the nature of these artificial categories is rather distinct from the nature of natural language concepts, proponents of exemplar models have often advocated the relevance of their findings for natural language concepts.

For several reasons, it is not obvious how the findings from artificial category learning should be generalized to learning and using natural language concepts. For example, for most natural language concepts, it is not clear which features are important in categorization decisions, whereas only a few well-defined features are manipulated in most artificial categories. Another problem is that, in the context of natural language concepts, it is not clear what an exemplar representation is (Medin, 1986). Komatsu (1992) described different possible views on the notion of exemplars in the context of natural language concepts. At one extreme, an exemplar representation might be a family-resemblance representation that abstracts across different specific instances. In this view, the concept *fish* may consist of the set of representations of trout, goldfish, shark, and so on, which are then considered exemplars of the category *fish*, but which are themselves abstractions. At the other extreme, exemplar representations may involve no abstraction at all. In this view, concepts are assumed to be stored in memory as a set of representations, each of which is a specific memory trace of a particular previously encountered instance. The dubious meaning of "exemplars" in the context of natural language concepts makes it difficult to apply the exemplar models developed in the category learning literature to natural language categories. In the context of the present paper, exemplars of the superordinate-level concepts *fruits* and *vegetables*, such as *tomato*, *apple*, *banana*, and so forth, are instantiations of these categories.

One remarkable exception is the instantiation principle presented by Heit and Barsalou (1996). The instantiation model essentially assumes that people generate instantiations of a category on which to base category-related decisions. More specifically, Heit and Barsalou predicted the typicality of lower level concepts (e.g., *mammals*) within a

This research project was supported by Grant G.0266.02 from the Belgian National Science Foundation (Fundamental Human Sciences) to G.S. We thank Sarah Casaer for assistance in purchasing the stimulus items and Marcel Lenaerts for making the photographs. All of the data described in this manuscript can be obtained from G.S. (in Excel files) upon request. Correspondence concerning this article may be addressed to G. Storms, Department of Psychology, Tiensestraat 102, B-3000 Leuven, Belgium (e-mail: gert.storms@psy.kuleuven.ac.be).

higher level concept (e.g., *animal*) from the typicality of instances of the lower level concept (e.g., *dog*, *horse*, *cat*) within the higher level concept (*animal*). Storms and his colleagues (De Wilde, Vanoverberghe, Storms, & De Boeck, in press; Storms, De Boeck, & Ruts, 2000; Verbeemen, Vanoverberghe, Storms, & Ruts, 2001) further applied the instantiation model to account for typicality ratings and response times in natural language concepts. Storms, De Boeck, and Ruts (2001) also used instantiation to predict categorization decisions for novel tropical food items as *fruits* or *vegetables.* To our knowledge, other influential exemplar models, such as the context model (Medin & Schaffer, 1978) or the generalized version of Nosofsky (the generalized context model, or GCM; Nosofsky, 1984, 1986), have never been applied to the categorization of natural language concepts.

The purpose of the present paper is twofold. First, we describe a study in which the GCM was applied to predict categorization choices of novel stimuli into well-known natural language concepts. The fruit-and-vegetable experiment described by Storms et al. (2001) was replicated and, for the first time, the GCM was applied to natural language concepts. Predictions of the GCM are compared with those of a (multiplicative-similarity) prototype model and with those based on the instantiation principle. Second, it has recently been suggested (Minda & Smith, 2001) that a prototype model with a multiplicative-similarity function (see also Nosofsky, 1992) yields better predictions for categorization decisions than exemplar models do, as the categories become larger. Studying *fruits* and *vegetables* allowed us to test this conjecture by comparing prototype and exemplar predictions for category sizes that are considerably larger than what is suitable in experiments in which participants have to learn new categories in a limited period of time.

Given that we aim to apply the GCM to data similar to those from Storms et al. (2001), in the following, the GCM and the experiment of Storms et al. (2001) will be described.

THE GENERALIZED CONTEXT MODEL

Nosofsky's (1984, 1986) GCM was derived from Medin and Schaffer's (1978) context model. Like the context model, the GCM assumes that subjects' classification of a new stimulus is based on its similarity to stored category exemplars. However, whereas stimuli in the context model vary along binary-valued dimensions, the GCM situates stimuli along continuous dimensions, usually after a multidimensional scaling (MDS) procedure is applied (Borg & Groenen, 1997). The model is based on Shepard's (1958) similarity choice model. Formally, for the case of two categories A and B, the probability that a given stimulus X is classified in category A is given by

$$P(A \mid X) = \frac{\beta_A \eta_{XA}^{\alpha}}{\beta_A \eta_{XA}^{\alpha} + (1 - \beta_A)\eta_{XB}^{\alpha}},$$

where β_A is a response bias toward category A and η_{XA} and η_{XB} are similarity measures of stimulus X toward all stored exemplars of categories A and B, respectively. The parameter α, first introduced into the GCM by Ashby and Maddox (1993), represents a response-scaling parameter. When $\alpha = 1.0$, observers respond by *probability matching* to the relative summed similarities. When α grows larger than 1.0, observers respond more deterministically with the category that yields the largest summed similarity (McKinley & Nosofsky, 1995; Nosofsky & Johansen, 2000). If α is less than 1.0, observers respond less deterministically than they do through probability matching. The similarity measures are summed similarities of the stimulus X toward all stored exemplars. Formally,

$$\eta_{XA} = \sum_{j \in A} \exp\left\{ -\left[c \left(\sum_{k=1}^{D} w_k \left| y_{xk} - y_{jk} \right|^r \right)^{1/r} \right]^q \right\},$$

where c is an overall scaling parameter, y_{xk} and y_{jk} are the coordinates of stimulus X and of the jth stored exemplar on dimension k, respectively, and w_k is the weight of dimension k. The weights of the different dimensions are restricted to sum to 1.0. The Minkowski r metric usually takes values between 1 and 2, wherein $r = 1$ results in the city-block metric and $r = 2$ results in Euclidean distances. Finally, the parameter q determines the shape of the similarity function, $q = 1$ resulting in a similarity function with exponential decay and $q = 2$ resulting in a Gaussian decay.

THE INSTANTIATION PREDICTOR

Storms et al. (2001) compared the predictions of an instantiation-based measure and a prototype-based measure in a categorization task using novel food items. The participants in their experiment were presented with actual specimens of tropical food items and were asked to categorize them as either a *fruit* or a *vegetable.* The instantiation prediction was based on the summed ratings, averaged over 10 raters, of the similarity of the presented stimuli to the seven most frequently mentioned fruits and to the seven most frequently mentioned vegetables in an exemplar-generation task. The participants in the generation task and those in the rating task were different. Still other participants judged whether or not characteristic features of fruits and vegetables applied to the stimuli. The prototype measure was based on the summed applicability frequencies of these features. The prediction obtained with the instantiation predictors for *fruits* and *vegetables* was somewhat better than the corresponding prediction based on the prototype predictors.

A final group of participants wrote down what the presented stimuli most reminded them of. The answers were considered to be "nearest neighbors" of the presented stimuli. Storms et al. (2001) showed that the instantiation predictor could be improved by adding rated similarities to these nearest neighbors (i.e., to fruits or vegetables that were given in answer to the question of what the presented stimuli most reminded the participants of). The resulting *extended instantiation measure* was based on the (unweighted) summed similarity ratings to the seven most frequently generated exemplars of each category (mentioned above) and to six or seven additional exemplars of each category taken from the *reminds-me-of* task. In the next section, a replication experiment will be described, which was implemented to compare the predictions of an extended-instantiation model with those of Nosofsky's (1984, 1986) GCM and those of a multiplicative-similarity prototype model.

THE EXPERIMENT

As was stated above, the participants in the Storms et al. (2001) study categorized actual specimens of tropical food items as *fruits* or as *vegetables.* A disadvantage of this procedure is that all tasks needed for the study had to be completed in 3 days because of the rotting process, which changed the color and form of some of the stimuli. Therefore, pictures of the specimens were used instead of the tropical fruits and vegetables themselves.

Method

Participants. Fifty students of the University of Leuven participated in the experiment voluntarily. Ten of them participated in a matrix-filling task, 30 participated in the categorization task, and 10 participated in the similarity-rating task.

Materials. To select the *well-known fruits and vegetables*, data from the exemplar-generati on task of Storms, De Boeck, Van Mechelen, and Ruts (1996) were used. In that study, 25 participants each generated 10 exemplars of the categories *fruits* and *vegetables*. This resulted in a list of 34 well-known fruits and 45 well-known vegetables. The 30 *novel* stimuli were exotic (mostly tropical) fruits and vegetables, purchased at specialty shops, that were presumed to be unknown to the participants. All of these novel stimuli were also used in Storms et al. (2001). A complete list of the 109 novel and well-known fruits and vegetables is given in the Appendix. Each of the 109 fruits and vegetables was photographed on a white background with an indication of its size. The photographs were printed in color (15 × 8 cm). The features of the concepts *vegetables* and *fruits* were also taken from the feature-generation task of Storms et al. (1996).

Procedure. The participants in the *matrix-filling task* were asked to judge whether or not each feature of a list of 17 features (of the concepts *fruits* and *vegetables*, taken from Storms et al., 1996) applied to the 79 well-known fruits and vegetables and to the 30 novel stimuli, by marking (for every item) each feature with a "1" for "yes" or a "0" for "no." They were instructed to make their best guess if they were not sure (which was sometimes the case for the novel stimuli). The pictures were presented in a randomized order. (For a complete list of the features used in the matrix-filling task, see Storms et al., 2001, Appendix B).

In the *categorization task*, the participants were asked to assign a category label, "fruit" or "vegetable," to each of the 109 pictured items (i.e., 30 novel food items and 79 well-known fruits and vegetables). The participants were also asked if they knew the item and to name it if they did. The pictures were again presented in a randomized order.

In the *similarity-rating task*, the participants were presented with pictures of the 30 novel fruits and vegetables only and rated, on a 10-point rating scale, how similar these 30 items were to 14 fruits and 13 vegetables. The latter 27 fruits and vegetables included the 7 most frequently generated exemplars of both categories, plus other fruits and vegetables that most resembled the 30 novel food items according to the results of the reminds-me-of task described in Storms et al. (2001).[1] The list of these 27 fruits and vegetables consisted of: apple, banana, pear, orange, kiwi, peach, grape, cherry, lemon, litchi, prune, coconut, berry, nut, lettuce, carrot, tomato, red cabbage, cauliflower, leek, pea, endive, cucumber, corncob, onion, bean, and sugar beet.

Results

Multidimensional scaling. The procedure that is most often used in the categorization literature to obtain an MDS solution of a stimulus set is to gather pairwise similarity judgments, which are then used as input for a scaling program. However, collecting similarity judgments from among 109 stimuli would require far too many comparisons. Furthermore, with complex stimuli such as the foods used in this study, there are likely to be highly idiosyncratic bases for making such judgments, which would most likely result in very noisy data. Therefore, the following procedure was used. First, the matrices of the 10 participants from the matrix-filling task were summed. Then a 109 × 109 stimulus-similarity matrix was derived by correlating the feature vectors of all pairs of stimuli. Next, the similarity matrix was analyzed using the MDS program ALSCAL (Takane, Young, & De Leeuw, 1977). Solutions in two to five dimensions resulted in stress values of .16, .09, .07, and .05, respectively. Because the improvement in fit for solutions in four and in five dimensions was minor, the three-dimensional solution was chosen for further analysis. The solution was rotated to a principal-axis orientation.[2] Figure 1 shows the projection of the stimuli on Dimensions 1 and 2 and on Dimensions 1 and 3. All novel stimuli (represented as diamonds) are labeled as indicated in the Appendix. The well-known fruits and vegetables (represented as squares and as triangles, respectively) are labeled only if they are outliers or if they are referred to elsewhere in the paper.

Categorization choices. As in Storms et al. (2001), the few "yes" responses to the question of whether the participants knew the presented novel food items were verified. If the naming of the item (in the case that they said they knew the item) was correct (which was very rarely the case), then the classification decision for the corresponding item was discarded from the data set. The proportion of *fruit* classifications (which is, of course, the complement of the proportion of *vegetable* classifications) functioned as the dependent variable in this study. For the well-known fruits and vegetables, these proportions were usually close to 1 or to 0, with a few exceptions (such as "avocado" and "tomato"). However, the proportion differed considerably over the 30 novel items that were presented. Only 2 novel items were unanimously classified by all of the participants in the classification task. Furthermore, for 9 novel stimuli, at most two thirds of the participants gave the majority classification answer. These results show that the participants differed in their classification choices and, thus, that the task was not trivial. The reliability of the classification decisions of the novel items was estimated by applying the Spearman–Brown formula to the split-half correlation after the participant group was randomly divided into two groups of equal size. The reliability estimate of the classification decisions was .95.

Fitting the GCM and the prototype model. All of the models were fitted to the data using a maximum-

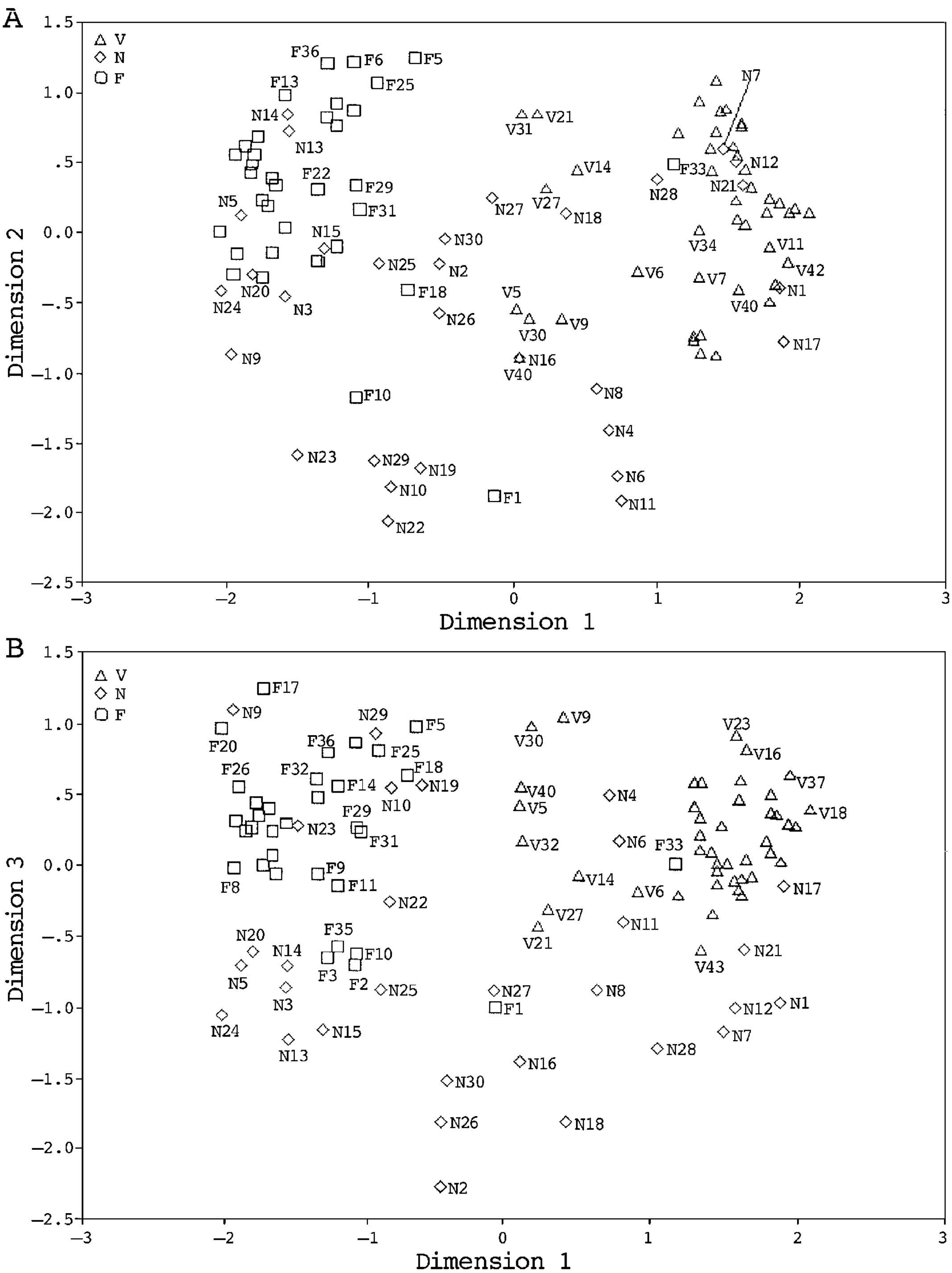

Figure 1. (A) Plot of Dimension 1 versus Dimension 2 of the MDS representation. (B) Plot of Dimension 1 versus Dimension 3 of the MDS representation.

likelihood criterion. The fits of the city-block metric and Euclidean distances ($r = 1$ and $r = 2$, respectively) were compared in several versions of the GCM. Only the results of the Euclidean distance function will be reported, because it resulted in clearly better fits, both in terms of the models' log-likelihood and in terms of the correlation between the observed and the predicted classification proportions. This finding indicates that the underlying dimensions are integral dimensions (Nosofsky, 1986; Shepard, 1964). Likewise, only the results for the exponential-

Table 1
Log-Likelihood Value (Log L), AIC, Sensitivity (c), α, Bias (β), for the Category *Fruit*, and Two Dimension Weights in the Exemplar-Based GCM and in the Prototype Model

Model	Log L	AIC	c	α	β for Fruit	w_1	w_2
Analysis of all 109 Stimuli							
Exemplar	−419.88	849.8	9.45	.41	.288	.27	.06
Prototype	−502.47	1,012.9	1.198	–	.374	.80	.16
Analysis of the 30 Novel Stimuli Only							
Exemplar	−71.03	152.06	9.75	.39	.238	.28	.03
Prototype	−83.96	175.91	0.81	–	.341	.98	.00
Analysis of the 79 Well-Known Stimuli Only							
Exemplar	−345.69	701.38	9.24	.45	.310	.25	.05
Prototype	−405.33	818.65	1.21	–	.392	.96	.04

Note—AIC, Akaike's (1974) information criterion.

decay similarity function ($q = 1$) will be reported, because it resulted in clearly better fits than did the Gaussian decay function ($q = 2$).

A version of the GCM with five free parameters was fit to the data. These five parameters were α, c, a bias parameter β, and two dimension weights (because the weights of the three dimensions are restricted to add up to 1.0). Next, a multiplicative-similarity prototype model (Nosofsky, 1987) was fit, which is identical to the GCM in terms of its assumptions about similarity, selective attention, and the response-ratio rule. However, unlike the GCM, the prototype model used only two stored exemplars—that is, the prototype of *fruits* and the prototype of *vegetables*. The coordinates of these two prototypes were calculated by averaging the coordinates for each dimension over all well-known exemplars within each category. This prototype model used only four free parameters instead of five, given that the values of α and of c cannot be estimated separately in this model. Note that, in the special case in which the present model is applied to stimuli that vary along binary-valued, separable dimensions (which is the paradigm used by Minda & Smith, 2001, and J. D. Smith & Minda, 1998), the model reduces to the multiplicative-similarity model used by J. D. Smith and Minda.

When the categorization proportions for the whole set of (well-known and novel) stimuli are being predicted, the correlations between the observed and the predicted classification proportions were .93 and .90 for the GCM and for the prototype model, respectively. The fits of the models were compared using Akaike's (1974) information criterion (AIC), a likelihood-based statistic that punishes for the number of free parameters used. The model with the lower AIC value is favored by the test; the test clearly favored the GCM over the prototype model. Table 1 presents the log-likelihood value, the AIC value, and the best-fitting parameters of all the stimuli. (Note that only two dimension weights are presented. The weight of the third dimension equals 1.0 minus the weights of the other two dimensions.)

In the analyses described above, the exemplar and prototype models were fit to all 109 stimuli, including both the well-known and the novel items. The advantage of the GCM over the prototype model is due to the latter's inability to predict the near-perfect classification of most of the well-known foods that are outliers in the MDS solution. If a model succeeds in classifying these outliers correctly, but fails for the less extreme classifications, its overall fit may still be fairly good. To examine this possibility, in a second analysis, we compared the predictions of the two models, but now only for the novel items. This amounts to excluding all stimuli that were used in the supervised category-learning phase. The correlations between the observed and the predicted classification proportions were .92 and .88 for the GCM and for the prototype model, respectively. The scatterplot of the correlation between the observed categorization proportions and the corresponding predictions of the GCM, shown in Figure 2, illustrates that the categorization proportions of many of the novel items are well distributed between 0 and 1. The likelihood of the models was again compared using AIC, and the GCM was favored over the prototype model. The middle part of Table 1 presents the log-likelihood value, the AIC value, and the best-fitting parameters of the novel stimuli.

Finally, in a third analysis, we compared the predictions of the two models, but now only for the well-known items. The correlation between the observed and the predicted classification proportions was .98 and .90 for the GCM and for the prototype model, respectively. As before, AIC favored the GCM over the prototype model. The lower part of Table 1 presents the log-likelihood value, the AIC value, and the best-fitting parameters for the well-known stimuli.

The GCM predictions were better than the prototype predictions in all three analyses, but the superiority of the GCM was clearest in the analysis of the well-known fruits and vegetables. A scatterplot of the observed and the predicted classification proportions showed that a few items with a relatively large distance from the prototype of the category to which they belong (tomato, corncob, cactus fruit, and medlar; see Figure 1) resulted in rather poor predictions of the prototype model. The GCM succeeded in predicting the corresponding categorization proportions much better, because the GCM uses exemplar information from all well-known exemplars, including these four items.

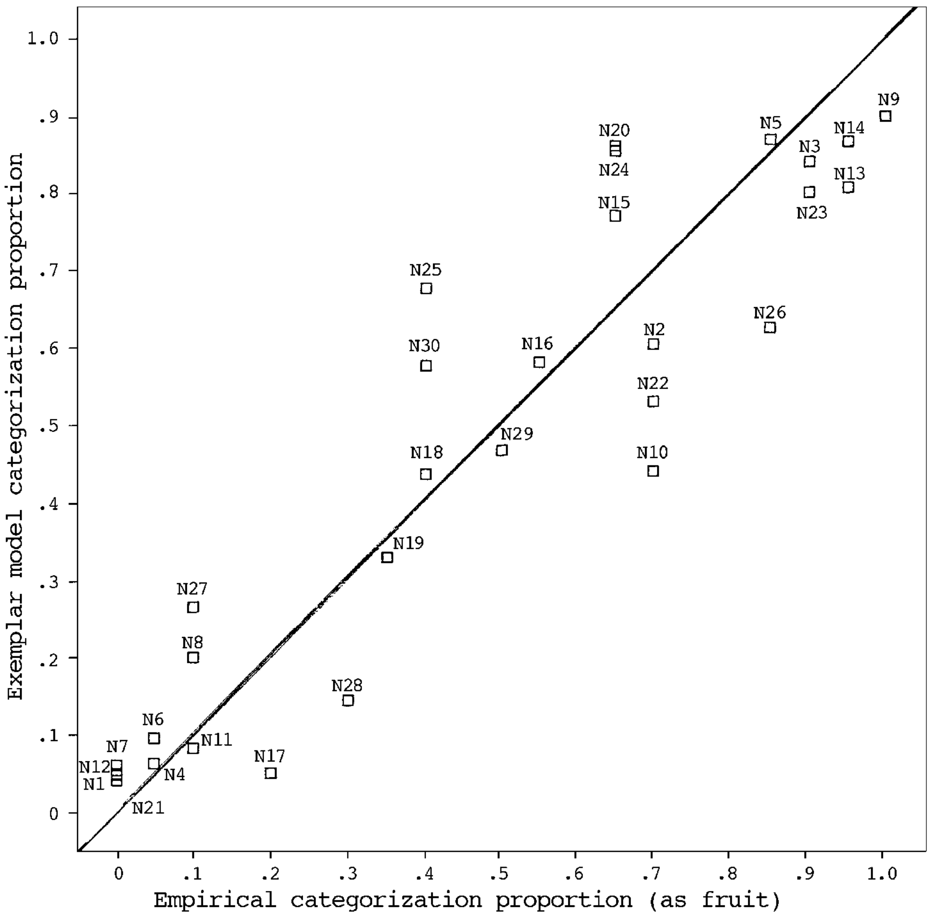

Figure 2. Scatterplot of the observed and predicted fruit classifications for the GCM.

Comparing predictions from the GCM with those of the (extended) instantiation model. The predictions of the GCM and those of the (extended) instantiation model can be directly compared for the categorizations of the novel stimuli, since (apart from the presentation mode, i.e., actual food items versus pictures of these items) the stimulus set used in Storms et al. (2001) and that used in the present study are identical. An instantiation-based prediction was calculated in the following way. First, for every stimulus that had to be categorized, the ratings of similarity to the 14 *fruit* and the 13 *vegetable* exemplars used in the Storms et al. (2001) instantiation predictor were summed. Next, the probability of a *fruit* classification for an item was set equal to the summed similarity to the 14 *fruits*, divided by the summed similarity to the 27 *fruits* and *vegetables*, after a response bias parameter β and a scaling parameter α analogous to those used in the GCM were added. The optimal estimated scaling and bias parameters were .75 and .45, respectively. The correlation between the observed and the predicted classification proportions was .87. This value is lower than the corresponding correlation for the GCM (which was .94, as noted above) and about the same as the correlation for the prototype model (which was .88). The AIC value for the extended intantiation model was 155.69, which was slightly higher than the corresponding value for the GCM but lower than the AIC value for the prototype model. Thus, the instantiation model fitted the categorizations of the novel stimuli better than the prototype model did, but slightly worse than the GCM did.

Discussion

To our knowledge, the experiment described in this paper is the first straight application of the GCM to categorization in natural language categories. The model provided good predictions of the probability with which observers classified both familiar and novel fruits and vegetables as members of these categories.

The predictions of the GCM were also compared with those of a (multiplicative-similarity) prototype model. Both models were compared for the whole stimulus set (i.e., the well-known fruits and vegetables as well as the novel stimuli) and for the novel stimuli and the well-known stimuli separately. In all three analyses, the GCM clearly outperformed the prototype model, both in terms of the correlation between observed and predicted categorization proportions and in terms of the statistical tests based on the likelihood of the data.

A detailed inspection of the observed and the predicted classification proportions showed that especially the well-known items with a relatively large distance from the prototype of the category to which they belonged resulted in rather poor predictions of the prototype model. The GCM succeeded in predicting the categorization proportions of these outliers much better, because it includes self-similarity in its predictions. These results suggest that a model that assumes an all-or-none memory trace for all of the well-known exemplars, but in which generalization is based solely on the category prototype (Minda & Smith, 2001), might predict the results better than a pure prototype model, such as the one discussed above. (Note, however, that such a model most likely still yields a worse fit than the GCM does, since the fit of the GCM was better than that of the prototype model even when both models were limited to the novel stimuli only.)

The predictions of the GCM for the novel stimuli were also compared with the corresponding extended-instantiation predictions based on Storms et al. (2001), after that model was granted a response bias and a scaling parameter analogous to those of the GCM. The predictions of the instantiation model were derived from direct similarity ratings of the novel stimuli to the most frequently generated exemplars of the categories *fruits* and *vegetables* and to the well-known fruits and/or vegetables that they were thought to resemble most. The GCM predictions were somewhat better than those derived from the instantiation-based predictors, although the difference in fit was small. The slight superiority of the GCM over the instantiation model is possibly caused by the similarity-scaling transformation of the GCM.

The present application of the GCM does differ from previously published applications of the model not only in that natural language categories were used. Earlier applications of the GCM used either a limited set of stimuli, constructed from a small number of dichotomously valued features (e.g., Nosofsky, 1992) or large sets of stimuli that varied along two continuous dimensions (e.g., McKinley & Nosofsky, 1995). Most likely, the latter types of stimuli were not easily recognized by the participants as clearly distinguishable and identifiable stimuli. The application described in this paper is unique in that a very large set of 79 well-learned and clearly distinguishable stimuli are used to predict a considerably large set of 30 transfer stimuli.

Minda and Smith (2001) recently compared exemplar-based predictions with those of a multiplicative-similarity prototype model in conditions with varying category sizes (ranging from 4 to 15 stimuli per category). They found that the prototype model yielded better predictions in all conditions, but that its advantage over the exemplar-based model increased with category size. Their conclusion that (multiplicative-similarity) prototype models are particularly suited to predict categorization in large categories is in clear contradiction with our findings. There may be several reasons why the results of our experiment favored the exemplar-based GCM over the prototype model. First, the category size of (well-known) fruits and vegetables is considerably larger than the sizes used in Minda and Smith's (2001) experiments. Second, and arguably most important, all the well-known exemplars of the very familiar categories *fruits* and *vegetables* were probably much better learned (over a period of about 20 years of experience with food) by our subjects than were the abstract stimuli used in the experiments of Minda and Smith. This argument is compatible with other recent findings in the literature that the advantage of exemplar models becomes stronger in later stages of learning, when exemplars are very well learned (see, e.g., Johansen & Palmeri, in press; J. D. Smith & Minda, 1998).

In conclusion, the results of the present study show that the GCM predicts categorization proportions of novel and well-known fruits and vegetables better than both a (multiplicative-similarity) prototype model and an instantiation-based predictor. These results illustrate the GCM's capacity to fit data from natural language concepts with a large set of well-known stored exemplars.

REFERENCES

Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, **37**, 372-400.

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

De Wilde, E., Vanoverberghe, V., Storms, G., & De Boeck, P. (in press). The instantiation principle re-evaluated. *Memory*.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning & 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* (pp. 67-95). London: Academic Press.

Heit, E., & Barsalou, L. W. (1996). The instantiation principle in natural language categories. *Memory*, **4**, 413-451.

Johansen, M. K., & Palmeri, T. J. (in press). Representational shifts in category learning. *Cognitive Psychology*.

Komatsu, L. K. (1992). Recent views of conceptual structure. *Psychological Bulletin*, **112**, 500-526.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling* (Sage University Paper Series on Quantitative Applications in the Social Sciences, No. 07-011). Beverly Hills, CA: Sage.

McKinley, S. C., & Nosofsky, R. M. (1995). Investigation of exemplar and decision-bound models in large-size, ill-defined category structures. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 128-148.

Medin, D. L. (1986). Comment on "Memory storage and retrieval processes in category learning." *Journal of Experimental Psychology: General*, **115**, 373-381.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, **44**, 1469-1481.

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

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology*, **35**, 113-138.

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.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual comprehension. *Psychological Review*, **92**, 289-316.

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.), *From learning theory to connectionist theory: Essays in honour of William K. Estes* (Vol. 1, pp. 149-167). Hillsdale, NJ: Erlbaum.

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. *Psychonomic Bulletin & Review*, **7**, 375-402.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, **7**, 532-547.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.

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

Shepard, R. N. (1958). 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. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, **1**, 54-87.

Smith, E. E., & Medin, D. M. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1411-1436.

Storms, G., De Boeck, P., & Ruts, W. (2000). Prototype and exemplar-based information in natural language categories. *Journal of Memory & Language*, **42**, 51-73.

Storms, G., De Boeck, P., & Ruts, W. (2001). Categorization of novel stimuli in well-known natural concepts: A case study. *Psychonomic Bulletin & Review*, **8**, 377-384.

Storms, G., De Boeck, P., Van Mechelen, I., & Ruts, W. (1996). The dominance effect in concept conjunctions: Generality and interaction aspects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1-15.

Takane, Y., Young, F. W., & De Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika*, **42**, 7-67.

Verbeemen, T., Vanoverberghe, V., Storms, G., & Ruts, W. (2001). The role of contrast categories in natural language concepts. *Journal of Memory & Language*, **44**, 618-643.

NOTES

1. The fruits and vegetables that were selected resembled at least one of the novel stimuli more than any of the seven most frequently generated fruits or vegetables did.

2. Since MDS was used as a data-reduction technique here, it is not easy to label the dimensions. In an attempt to determine the type of information reflected in the MDS solution, regression analyses were done to predict the feature-applicability frequencies of each of the individual features that were used to obtain the input correlation matrix of the MDS from the coordinates on each of the three dimensions in the solution. (See Kruskal & Wish, 1978, for more details on this procedure.) High R^2 values of above .80 were obtained for the features "grows under or just above the ground" and "is green." Because both features are rather independent ($r = .19$), they (or similar correlated features) seem to determine the solution to a large extent.

(*Continued on next page*)

APPENDIX
List of the Stimuli

Well-known Fruits	Well-known Vegetables	Novel Foods*
Medlar (F1)	Lentil (V1)	Lemon grass (N1)
Blackberry (F2)	Chick peas (Variety 1) (V2)	Mangosteen (N2)
Raspberry (F3)	French (green) bean (V3)	Kiwano (N3)
Peanut (F5)	Sprouts (V4)	Bitter melon (N4)
Chestnut (F6)	Chili (hot pepper) (V5)	Tomarillo (N5)
Nectarine (F7)	Eggplant (V6)	Okra (N6)
Pomegranate (F8)	Artichoke (V7)	Turmeric (N7)
Fig (F9)	Chicory (V8)	Thai eggplant (N8)
Cactus fruit (F10)	Cucumber (V9)	Guava (N9)
Papaya (F11)	Red cabbage (V10)	Bergamot (N10)
Lytchee (F12)	Broccoli (V11)	Lilac (N11)
Pineapple (F13)	Cauliflower (V12)	Jerusalem artichoke (N12)
Coconut (F14)	Celery (V13)	Rambutan (N13)
Kiwi (F15)	Winter radish (V14)	Edos (N14)
Mango (F16)	White cabbage (V15)	Kumquat (N15)
Apple (F17)	Parsley (V16)	Patisson (N16)
Avocado (F18)	Spring onion (V17)	Kan-toon (N17)
Star fruit (F19)	Leek (V18)	Cola nut (N18)
Pear (F20)	Carrot (V19)	Big gourd (N19)
Orange (F21)	Garlic (V20)	Pitahaya (N20)
Lemon (F22)	Tomato (V21)	Sweet potato (N21)
Melon (F23)	Turnip (V22)	Tiny korella (N22)
Passion fruit (F24)	Lettuce (V23)	Jujube (N23)
Walnut (F25)	Radish (V24)	Cherimoya (N24)
Grapes (F26)	Celeriac (V25)	Ripe tamarind (N25)
Mandarin (F27)	Pumpkin (V26)	Young pepper (N26)
Prune (F28)	Fennel (V27)	Banana blossom (N27)
Grapefruit (F29)	Courgette (V28)	Taro (N28)
Banana (F30)	Pepper (V29)	Chayote (N29)
Date (F31)	Corncob (V30)	Safon (N30)
Lime (F32)	White bean (V31)	
Rhubarb (F33)	Bean (V32)	
Cherry (F34)	Onion (V33)	
Strawberry (F35)	Potato (V34)	
Monkey nut (F36)	Asparagus (Variety 1) (V35)	
	Chick Peas (Variety 2) (V36)	
	Gherkin (V37)	
	Endive (V38)	
	Chervil (V39)	
	Seaweed (V40)	
	Salsify (Variety 1) (V41)	
	Asparagus (Variety 2) (V42)	
	Salsify (Variety 2) (V43)	
	Spring onions (V44)	

*The names of the novel items are in different languages. It turned out to be impossible to find English names for all these items.

(Manuscript received July 31, 2001;
revision accepted for publication December 15, 2001.)