Title

Comparisons of prototype- and exemplar-based neural network models of categorization using the GECLE framework

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 26(26)

ISSN
1069-7977

Author
Matsuka, Toshihiko

Publication Date
2004

Peer reviewed

Comparisons of prototype- and exemplar-based neural network models of categorization using the GECLE framework

Toshihiko Matsuka (matsuka@psychology.rutgsers.edu)
RUMBA, Rutgers University – Newark
101 Warren St., Smith Hall 327, Newark, NJ 07102 USA

Abstract

In the present study, GECLE (Matsuka, 2003) was used as a general modeling framework to systematically compare the plausibility of two prominent assumptions about internal representations of neural network (NN) models of human category learning. In particular, exemplar-model friendly Medin and Schaffer's 5/4 stimulus set (1978) was used for comparing prototype- and exemplar-based NN models. The results indicate that some prototype-based models performed as good as or better than an exemplar-based model in replicating the empirical classification profile. In addition, a phenomenon called A2 advantage (i.e., people tend to categorize the less "prototypical" stimulus A2 more accurately than more "prototypical" stimulus A1) reported in empirical studies (e.g., Medin & Schaffer 1978) was also successfully reproduced by these prototype-based NN models.

Introduction

There have been an increasing number of studies debating how stimuli are internally represented in human cognition during the last few decades (e.g., Minda & Smith 2002; Nosofsky & Zaki 2002). Most of these debates have been based on quantitative models of categorization, and only a few have considered representational aspects of adaptive, network, or learning models of categorization. Several studies (Matsuka, 2002; Matsuka, Corter, & Markman, 2003) have compared exemplar-based (EB) and prototype-based (PB) adaptive network models of categorization, but there has been no systematic comparison of specific assumptions in EB and PB modeling. Although these comparative studies provided information on the models' capabilities for reproducing human-like categorization learning, they did not necessarily provide information that can lead to specific understanding of the nature of human category learning. That is because model-to-model comparisons are not informative for testing the plausibility of each specific assumption, rather such model comparisons are essentially omnibus tests collectively comparing all variations in assumptions at once. In other words, it has been difficult to use the results of these previous comparative studies to understand which specific assumptions are supported by the empirical data. Therefore, it seems desirable to make systematic comparisons between competing model assumptions using a general modeling framework that allows us to manipulate and test one or a limited number of model assumptions at a time.

In the present study, a generalized exploratory modeling approach for human category learning is introduced. Then, using this general framework two assumptions about how categories are internally represented, namely prototypes and exemplars, are compared in a systematic fashion.

GECLE

GECLE (for Generalized Exploratory models of Category LEarning) is a general and flexible exploratory approach for modeling human category learning, that is capable of modeling human category learning with many variants using different model assumptions (Matsuka, 2003). This general modeling framework allows model assumptions to be manipulated separately and independently. For example, one can manipulate assumptions about how stimuli are internally represented (e.g. exemplars vs. prototypes), or about how people selectively pay attention to input feature dimensions (e.g., paying attention to dimensions independently or not).

The GECLE model uses the Mahalanobis distances (in the quadratic form) between the internally represented reference points (RP: corresponding to either exemplars or prototypes) and the input stimuli as the measure of similarity between them. Thus, unlike other neural network models of category learning, GECLE does not necessarily assume that attention is allocated independently dimension-by-dimension. Rather, it assumes that humans in some cases might pay attention to correlations among feature dimensions. This allows GECLE to model processes interpretable as dimensionality reduction or mental rotation in the perception and learning of stimuli. Such processes may increase the interpretability of stimuli in categorization tasks. Another motivation for using the Mahalanobis distance is that the capability for paying attention to correlations among feature dimensions may be necessary for classification tasks defined on integral stimuli.

In the GECLE framework, the attention parameters (which are the diagonal and off-diagonal elements of the covariance matrices) can be considered as *shape* and *orientation* parameters for receptive fields or attention coverage areas of the reference points. It should be noted, however, that one can constrain GECLE to incorporate the "dimensional attention processes" assumption (i.e., attention is allocated independently on a dimension-by-dimension basis) by forcing the off-diagonal entries in the covariance matrices to be equal to zero.

Another unique feature of GECLE's attention mechanism is that it allows each reference point to have uniquely shaped and oriented attention coverage area, which is referred to as "local attention coverage structure" (Matsuka 2003). Again, one can impose a restriction on the model's attention mechanism by fixing all covariance matrices to be the same, which may be called "global attention coverage structure". Many NN models of category learning, ALCOVE (Kruschke, 1992) for example, incorporate the global attention coverage structure.

The local attention coverage structure model is complex, but may plausibly model attention processes in human category learning. For example, it allows models to be sensitive to one particular feature dimension when the input stimulus is compared with a particular reference point that is highly associated with category X, while the same feature dimension receives little or no attention when compared with another reference point associated with category Y. Thus the local attention coverage structure causes models to learn and be sensitive to within-cluster or within-category feature configurations, while the global attention coverage structure essentially stretches or shrinks input feature dimensions in a consistent manner for all RP receptive fields and all categories.

Another way of interpreting GECLE's capabilities for paying attention to correlations among feature dimensions and having local attention coverage structures is that the model learns to define what the feature dimensions are for each RP and to allocate attention to those dimensions independently. In contrast, for almost all previous adaptive models of category learning, the definition of the feature dimensions is static and supplied by individuals who use the models.

Some studies showed that humans learn much better in "filtration" tasks, in which information from only one dimension is required for (perfect) categorization, than in "condensation" tasks, in which information from two dimensions is required (e.g., Gottwald & Garner, 1975). This finding has been used as evidence that people pay attention to each dimension independently, rather than dependently (i.e., paying attention to correlations). Thus, a model paying attention to correlations or having diagonal attention coverage, as GECLE does, may not replicate filtration advantage. However, Matsuka (2003, 2004) successfully replicated the filtration advantage using a prototype based correlation-attentive GECLE with local attention coverage structure. He suggested that for a prototype based GECLE, the condensation stimuli require a stricter correspondence or synchronization between prototype search (i.e., shifting centroids of prototypes) and psychological scaling of the two feature dimensions (i.e., attention processes) as compared with the filtration stimuli. This is because the "correct" prototypes and "correct" scaling are defined by two dimensions in the condensation stimuli as compared to one dimension in the filtration stimuli.

In its natural form, the GECLE may be considered as a model using prototype internal representation, because it tries to learn to locate its reference points at the centers of each category cluster. However, with proper user-defined parameter settings, it can behave like a model with an exemplar-based internal representation.

Quantitative Descriptions (Algorithm)

The feedforward and learning algorithms of the GECLE are typical for implementation of the Generalized Radial Basis Function (Haykin, 1999; Poggio & Girosi, 1989, 1990). GECLE uses the following function to calculate the distances or similarity between internally represented reference points and input stimuli:

$$D_j^n\left(x^n, r_j\right) = \left(x^n - r_j\right)^T \Sigma_j^{-1}\left(x^n - r_j\right) \tag{1}$$

where x^n is an I-ruple vector representing an input stimulus consisted of I feature dimensions presented at time n, r_j, also an I-ruple vector, that corresponds to the centroids of reference point j, expressing its characteristics, and Σ_j^{-1} is the inverse of the covariance matrix, which defines the shape and orientation of the attention coverage area of reference point j. For a model with global attention coverage structure, there is only one global Σ^{-1} for all reference points.

The psychological similarity measures $D_j(x,r)$ cause some activations in internal "hidden" units or reference points (i.e., exemplars or prototypes). The activation of "hidden" basis unit j, or h_j, is obtained by any differentiable nonlinear activation transfer function (ATF), or

$$h_j = G\left(D_j(x,r)\right) \tag{2}$$

given that its first derivative $G'(\cdot)$ exists. An exponential function, $\exp(-cD_j(x,r))$, is an example of an ATF. The ATF must be a differentiable function, because GECLE uses a gradient method for learning, where the partial derivatives are used for updating the learnable parameters. However, it is possible to eliminate this restriction by incorporating a form of derivative-free learning algorithm such as stochastic learning (Matsuka & Corter 2004).

The activations of hidden units are then fed forward to output nodes. The activation of the kth output node, O_k, is calculated by summing the weighted activations of all hidden units connected to the output node, or

$$O_k = \sum_{j=1}^{J} w_{kj} h_j \tag{3}$$

where w_{kj} is the association weight between output node k and reference point j. The probability that a particular stimulus is classified as category C_k, denoted as P(C), is assumed equal to the activity of category k relative to the summed activations of all categories, where the activations are first transformed by the exponential function (Kruschke, 1992)

$$P(C) = \frac{\exp(\phi O_c)}{\sum_k \exp(\phi O_k)}. \tag{4}$$

ϕ is a real-value mapping constant that controls the "decisiveness" of classification responses.

GECLE uses the gradient method to update parameters. The error function is defined as the sum of squared differences between targeted and predicted output values (i.e., L_2 norm), or

$$E(w,r,\Sigma^{-1}) = \frac{1}{2}\sum_{k=1}^{K} e_k^2 = \frac{1}{2}\sum_{k=1}^{K}\left(d_k - O_k\right)^2 \tag{5}$$

Then the following functions are used to update parameters.

$$\Delta w_{jk} = \frac{\partial E}{\partial w_{jk}} = -\eta^w e_k h_j \tag{6}$$

where η^w is the learning rate for the association weights.

$$\Delta r_j = \frac{\partial E}{\partial r_j} = -\eta^r \sum_{k=1}^{K} e_k w_{jk} G'\left(D_j(x,r)\right)\Sigma_j^{-1}\left(x - r_j\right) \tag{7}$$

where $G'(\cdot)$ is a derivative of $G(\cdot)$. Equation 7 can be considered as a function that locates or defines prototypes of stimuli. For the exemplar-based modeling η^r must be set to zero to maintain the static nature of reference points.

$$\Delta\Sigma_j^{-1} = \frac{\partial E}{\partial \Sigma_j^{-1}} = \eta^{\Sigma}\sum_{k=1}^{K} e_k w_{jk} G'\left(D_j(x,r)\right)\left(x-r_j\right)\left(x-r_j\right)^T \quad (8)$$

For models with global attention coverage structure, Equation 8 should be summed over both k and j.

Hierarchy of Constraints on Attention Parameters

There is a hierarchy of constraints that one can impose on the attention parameters Σ^1 to manipulate GECLE's attention mechanisms. There are two levels of uniqueness of Σ^1 (global and local attention coverage structure), in each of which there are three levels of constraints on entries in Σ. The following is a list of six possible levels of restriction. Note that regardless of the types of restriction, the entries (s_{im}) in Σ_j are assumed and constrained to satisfy the following conditions: $s_{ii} \geq 0$ & $|s_{im}| \leq$ MIN(s_{ii}, s_{mm}).

Global Attention Coverage Structures

A. Global Pure Radial (GPR): Constraints on Σ_j: $s_{ii} = s$, for all i: $s_{im} = 0$, for all $i \neq m$; $\Sigma_j = \Sigma$ for all reference points j.
B. Global Uncorrelated Non-radial (GUN): Constraints on Σ_j: $s_{im} = 0$, for all $i \neq m$; $\Sigma_j = \Sigma$ for all reference points j.
C. Global Correlated Non-radial (GCN): Constraints on Σ_j: $\Sigma_j = \Sigma$ for all reference points j.

Local Attention Coverage Structures

D. Local Pure Radial (LPR): Constraints on Σ_j: $s_{ii} = s$, for all i; $s_{im} = 0$, for all $i \neq m$.
E. Local Uncorrelated Non-radial (LUN): Constraints on Σ_j: $s_{im} = 0$, for all $i \neq m$.
F. Local Correlated Non-radial (LCN): Constraints on Σ_j: none.

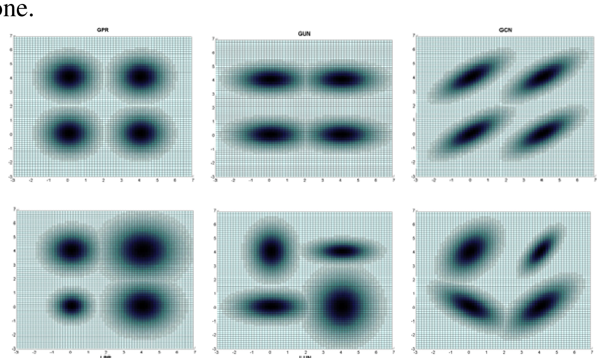

Figure 1. Six types of attention structures in the GECLE framework. Clockwise from top left. GRP, GUN, GCN, LCN, LUN, and LRP.

Simulations

In this section, three simulation studies were conducted to compare adaptive network models of category learning utilizing prototypes or exemplar internal representations using the GECLE framework. Here, a classical category learning study (Medin & Schaffer 1978) was replicated with several variants of GECLE. Simulation 1 reports the predictions by several GECLE models based on "optimal" parameter values. In Simulation 2, the general tendencies in some key aspects associated with the stimulus set were investigated with the same GECLE models used in Simulation 1. The plausibility of prototype models was further investigated using two variants of prototype-based GECLE in Simulation 3.

Simulation 1

In Simulation 1, I simulated category learning using the well-known Medin and Schaffer's 5/4 stimulus set (1978). Table 1 shows the schematic representation of the stimulus set. Eight different GECLE-based models were involved in the present simulation study. Among them there were seven prototype-based models (PB) with 2,3,4,5,6,7, or 8 prototypes and one exemplar-based model (EB) with all 9 unique exemplars. The global attention structure with dimensional attentional processes (i.e., GUN) was used for all eight models. They were run in a simulated training procedure to learn the correct classification responses for the training set. The models were run for 100 blocks of training, where each block consisted of a complete set of the training instances. The final parameter values used for each model were chosen by a simulated annealing method to minimize the objective function (i.e., sum of squared error: SSE) in reproducing the classification profile reported in the original Medin & Schaffer's work (1978). There are a total of 50 simulated subjects in each condition.

The following one-parameter exponential activation transfer function was used for the models:
$h_j = \exp\left(-c \cdot D_j(x,r)\right)$

One of the main interests of the present simulation study was how well the eight models could reproduce observed classification profile reported in Medin & Schaffer (1978). The other related interest was how well each model performs on stimuli A1 and A2 (see Table 1). These two stimuli have been considered to be very important and diagnostic, because PB and EB tend to give different predictions for these particular stimuli (e.g., Nosofsky & Zaki, 2002). Specifically, EB models are used to explain empirical results that show that humans are better able to categorize less "prototypical" A2 than more "prototypical" A1 (e.g., Medin & Schaffer 1978). Moreover, simulation studies (e.g., Nosofsky & Zaki 2002) indicate that EB gives a better fit for differential performance on these particular stimuli.

Table 1. Stimulus set used in Simulation 1

	Training Set					Transfer Set			
	Cat	D1	D2	D3	D4	D1	D2	D3	D4
A1	A	1	1	1	0	1	0	0	1
A2	A	1	0	1	0	1	1	1	1
A3	A	1	0	1	1	0	1	0	1
A4	A	1	1	0	1	0	0	1	1
A5	A	0	1	1	1	1	0	0	0
B1	B	1	1	0	0	0	0	1	0
B2	B	0	1	1	0	0	1	0	0
B3	B	0	0	0	1				
B4	B	0	0	0	0				

Results: Table 2 shows two fit indices for the eight models, namely SSE as an absolute fit index, and SSE multiplied by the number of learnable parameters (NLP) as a (crude) relative fit index that may account for the model complexity. A pure prototype model (here a pure prototype is defined as a model that has as many RPs as the number of categories) performed worst before and after controlling for the model complexities. In addition, it failed to show the A2 advantage. Rather as in many previous studies, it predicted that A1 was easier than A2. However, other PB models performed well; PB8 resulted in the best absolute fit, and PB5 resulted in the best relative fit.

When the PB models are compared with the EB model, some PBs fit the observed profile better than EB, particularly after controlling for the model complexities. More interestingly, as the EB model, almost all PBs were able to predict the A2 advantage (Table 2, last column).

Although, this Medin and Schaffer 5/4 stimulus set has been used as evidence supporting exemplar-based models and undermining prototype-based models, the results of the present simulation study appear to show no competitive advantage of the exemplar-based model. Instead, some PB models were able to reproduce the observed classification profile and the A2 advantage equally successfully with smaller numbers of learnable parameters.

Table 2. Results of simulation 1

Model	NLP	NRP	SSE	SSE x NLP	A2-A1
PB2	16	2	0.1438	2.301	-5.633
PB3	22	3	0.0694	1.527	3.643
PB4	28	4	0.0361	1.011	5.444
PB5	34	5	0.0250	0.850	9.046
PB6	40	6	0.0215	0.860	2.663
PB7	46	7	0.0193	0.888	4.314
PB8	52	8	0.0182	0.946	3.273
EB9	58*	9	0.0201	1.166	8.011

NLP: Number of Learnable Parameters
NRP: Number of Reference Points (e.g. prototype or exemplar)
* Location parameters for exemplar were static & not subject for learning, but assumed that optimized locations were learned when the exemplars were created.

Discussion of Simulation 1: All GECLE models that were capable of learning to locate the reference points were interpreted as prototype-based models in the present simulation study. However, it might not have been a sensible interpretation for some of those models, particularly for models with larger numbers of prototypes (e.g., PB5 ~ PB8). That is, it does not seem logical to create eight prototypes from only nine unique stimuli. Rather, there may be better interpretations for these models. Two possible alternative interpretations are discussed below.

First, it might be more sensible to interpret PB GECLE with larger numbers of prototype as models utilizing "fuzzy" or modular prototypes (or simply modules) as the reference points (RP) in a combinatorial fashion: it tries to create and memorize modules (defined by or being prototypes of subsets of stimuli belonging to a particular category) that summarize characteristics of particular feature dimensions more correctly than the other feature dimensions for a particular category, and uses combinations of the module activations triggered by similarities between the modules and input stimuli for categorizing. This combinatorial coding seems to be a very efficient use of limited mental resources for categorizing virtually unlimited number of unique instances.

Alternatively, those models that were interpreted as prototype-based GECLE with many prototypes might have been utilizing RPs that were more sensible to be interpreted as probabilistic, partial, or erroneous exemplars, instead. That is, although the models might have tried to store correct exemplars in their memory, the process was not fully completed because of the limited mental resources, resulting in imprecise exemplars memorization, in which a particular feature of a particular exemplar was more correctly memorized than other features. Then, these imprecise exemplars were utilized for categorizing the stimuli.

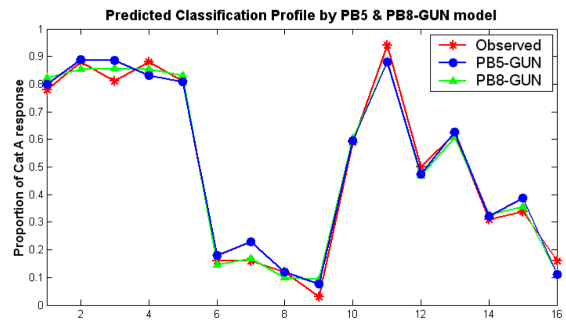

Figure 2. Predicted classification profiles by two best prototype based GECLE models (i.e., PB8-GUN: lowest absolute fit; PB5-GUN: lowest relative fit).

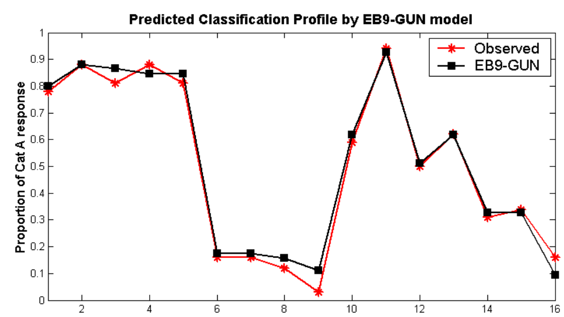

Figure 3. Predicted classification profiles by exemplar based GECLE model (i.e., EB9-GUN).

Simulation 2

Simulation 2 is a replication of Simulation 1 with 10,000 randomly chosen parameter configurations to investigate general tendencies in the A2 advantage by the same eight models used in Simulation 1. Here, the 10,000 simulated subjects with randomly assigned parameter values were trained to classify the 5/4 stimulus set. The ranges of parameters were [0.1 10] for c and ϕ, [0.001 1] for the three learning rates.

Results & discussion: Table 4 summarizes the results of Stimulation 2. In short, the A2 advantage was observed in almost all PB and EB models, indicating that the results of Simulation 1 are reasonably generalizable in this regard.

More interestingly, the EB model showed lesser magnitude of the A2 advantage than several PBs. This was mainly because EB9 learned to produce network output activations correctly with many parameter configurations (i.e., minimizing the error defined as Equation 5 perfectly) since the model was supplied the correct locations of all unique stimulus exemplars from the beginning of the training. This in turn, resulted in very small differences in classification responses for Stimuli A1 and A2, because the activations triggered by Stimulus A1 and A2 for the output nodes were almost identical (i.e., L_2 was minimized). This implies that any EB-based GECLE or any EB-based model such as ALCOVE would find this learning task (here, learning task does not correspond to categorization, but L_2 minimization, i.e., Eq. 5) easy because it can satisfactorily complete the task with virtually any parameter settings inasmuch as the locations of exemplars were well defined. Although this may be true if the condition of correctly memorizing exemplars is met, there is no guarantee for satisfying the condition in real human cognition. But, more likely, the condition would not be tenable for some people (i.e., some memorize exemplars more correctly and/or faster than other individuals). This difference in memorization ability may be one of the factors creating individual differences in category learning. This aspect of exemplar type modeling alone does not invalidate the assumption of exemplar-type internal representation, but it does suggest that exemplar-based (computational) models of categorization could be benefited from integrating an algorithm or quantitative explanation of how people learn and memorize exemplars.

On the contrary, exemplar theorists may argue that the upper limits of the randomly selected learning rate parameters (or the number of training epochs) were set unrealistically high. Although this argument is likely valid and thus the interpretation of the results may require some caution, it is still true that exemplar model may need to have learning algorithm for exemplar initialization, maintenance, and memorization.

Table 4. Results of Simulation study 2: Differences in classification accuracies for A2 and A1. (numbers of observed cases shown in parentheses).

Model	Overall	Classification Accuracy (CA) in training	
		100≥ CA >90%	90≥ CA >80%
PB2	1.011	-8.725(117)	-8.178(162)
PB3	2.184	0.056(250)	0.539(295)
PB4	2.521	0.331(556)	1.261(369)
PB5	3.071	0.885(905)	3.007(342)
PB6	2.711	0.661(1212)	3.816(365)
PB7	2.962	0.342(1690)	4.029(367)
PB8	2.446	0.330(2037)	2.885(393)
EB9	0.050	0.014(7660)	0.087(837)

Note: Observed classification accuracy for the training set is 0.85

Simulation 3

Simulations 1 and 2 showed that the pure prototype model, PB-2, accounted poorly for phenomena associated with the Medin and Schaffer's stimuli. However, these results might have resulted from incorrect assumptions about the prototype modeling. For example, I assumed that the locations of prototypes were continuously updated throughout the training, but in reality, people may quickly identify prototypes which may be less likely to be updated unless absolutely necessary. Another possible explanation is that people may have a uniquely shaped activation area for each prototype and/or pay attention to correlation among feature dimensions. For example, Matsuka (2003 & 2004) showed that there may be an interaction between types of internal mental representation and types of attention mechanism: the prototype-based model performed better when it incorporated unique attention structure with the capability of paying attention to dimensional correlations; whereas the exemplar-based model performed better with global attention structure with independent dimensional attention processes (i.e., no attention to correlations).

In the present simulation study, pure prototype modeling was reinvestigated using two variants of the original PB2 GECLE. The first one, SPB-2 is a static version of PB-2. That is SPB-2 is identical to PB2 appeared in Simulations 1 and 2, but the locations of prototypes were supplied from the beginning of the training and the learning rate for RPs was set to zero. Thus, this model resembles EB-based GECLE (except that RPs were prototypes) in that the locations of RPs were static. The second one, CPB2, is PB2-GECLE with the most complex attention mechanism, namely LCN (see Figure 1, lower right panel), having a unique receptive field for each prototype and the capability of paying attention to correlation.

For SPB2, the prototype for each category was created by averaging the feature values of each dimension of every object in a particular category, thus [0.8 0.6 0.8 0.6] for Category A and [0.25 0.5 0.25 0.25] for Category B. The rest of the procedures of the present simulation study follow those of Simulations 1 and 2.

Table 5a. Simulation 3: Results based on optimal parameters

Model	NLP	NRP	SSE	SSE x NLP	A2-A1
SPB2	16	2	0.1972	3.155	-9.208
CPB2	32	2	0.0377	1.206	11.130

Table 5b. Simulation 3: A2 advantage based on randomly drawn parameters.

Model	Overall	Classification Accuracy (CA) in training	
		100≥ CA >90%	90≥ CA >80%
SPB2	-2.346	-3.740 (2263)	-4.535(1920)
CPB2	2.931	-0.814(2215)	5.964(1505)

Results & discussion: A great decrease in SSE was obtained for CPB2 as compared with the original PB2, and after controlling for the model complexity by the simple linear adjustment (i.e., SSE x NLP) it performed nearly as good as EB9 (1.206 vs.1.166). In addition, unlike PB2, CPB2 was able to replicate the A2 advantage, and it was shown to be generalizable to some extent in the second part of the present simulation study using the randomly drawn parameters (Table 5b). In contrast SPB2 performed worse than PB2 for replicating the observed classification profile. Moreover, SPB2 consistently failed to replicate the A2 advantage in the randomized simulation study.

Discussion on Simulations

Medin and Schaffer's 5/4 stimulus (1978) has been used as a benchmarking stimulus set for computational models of categorization and category learning, usually favoring exemplar models (e.g. Matsuka et al. 2003; Minda & Smith 2002; Nosofsky & Zaki, 2003). However, the results of the present simulation studies showed that several GECLE models with prototype internal representation performed as good as or better than the exemplar-based GELCLE. One type of those successful prototype-based GECLE was the model that created and utilized multiple *modular* prototypes for categorization. The modular prototype is a prototype defined by subsets of stimuli belonging to a particular category that summarize characteristics of particular feature dimensions more correctly than the other feature dimensions for the particular category (however, the modular prototypes may be interpreted as imprecise exemplars). The other type of the successful prototype-based GECLE was the one with uniquely shaped and oriented attention coverage areas and with the capability of paying attention to correlations among feature dimensions.

There are at least few concerns associated with the present simulation studies. First one, as discussed in Simulation 1, is that as the number of GECLE's reference points (RP) increases, it become philosophically difficult within the cognitive science paradigm to interpret what these RP are representing (e.g., modular prototypes vs. imprecise exemplars). The other concern is the way the numbers of learnable parameters were counted for the exemplar-based GECLE (see notes on Table 2). That is, in the present simulation studies, the location parameters of the exemplars were counted as learnable parameters. On one hand, the locations of exemplars may be learnable, because they are initialized at the "optimal" location without error. On the other hand, they may not be learnable, because they reside in static locations.

Conclusions

Generalized Exploratory model of human Category LEarning (GECLE) is a flexible and general framework for modeling human category learning that is capable of manipulating a limited number of assumptions independently and systematically. In the present study, the plausibility of two different assumptions about internal representation was investigated with GECLE using exemplar-model-friendly Medin & Schaffer 5/4 stimulus set (1978). The results of simulations showed no competitive advantage of previously favored exemplar-based modeling. Rather, they appeared to suggest some prototype models performed better than an exemplar model. In addition, the exploratory nature of GECLE yielded new plausible prototype-based adaptive models of category learning with different structures and model assumptions.

Although, several models were examined in some depth in the present research, the results were based only on a simulation of one empirical study. More simulation studies with several other stimulus sets should help identify models or assumptions with descriptive validities more accurately. In addition, measurements of several different cognitive processes associated with category learning, such as, attention allocation should be collected in empirical studies, in order to restrict model parameters and to better differentiate among models.

References

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation (2nd ed.).* Upper Saddle River, NJ: Prentice Hall.

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

Matsuka, T (2002). Attention processes in computational models of categorization. Unpublished Doctoral Dissertation. Columbia University, NY.

Matsuka, T. (2003). General exploratory model of human category learning. Accepted for publication.

Matsuka, T. (2004). Interactions between representation and attention processes in category learning. Poster presented at the 11th Annual Meeting of Cognitive Neuroscience Society. San Francisco.

Matsuka, T. & Corter, J.E (2004). Stochastic learning algorithm for modeling human category learning. Accepted for publication.

Matsuka, T., Corter, J. E. & Markman, A. B. (2003). Allocation of attention in neural network models of categorization. Under revision.

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

Minda, J.P. & Smith, J.D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 275-292.

Nosofsky, R.M. & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 924-940.

Nosofsky, R.M., Gluck, M.A., Palmeri, T.J., McKinley, S.C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins

Poggio, T. & Girosi, F. (1989) A Theory of Networks for Approximation and Learning). *AI Memo 1140/CBIP Paper 31*, Massachusetts Institute of Technology, Cambridge, MA.

Poggio, T. & Girosi, F. (1990). Regularization algorithms for learning that are equivalent to multilayer networks. *Science, 247,* 978-982.