Psychological Review
1995, Vol. 102, No. 1, 185–191

0033-295X/95/$3.00

THEORETICAL NOTES

Base-Rate Neglect in ALCOVE: A Critical Reevaluation

Stephan Lewandowsky
University of Oklahoma

A recent hybrid model of categorization (Attention Learning Covering Map [ALCOVE]; J. K. Kruschke, 1992) has combined the most desirable properties of exemplar models with a connectionist architecture and learning rule. A critically important property of ALCOVE is its apparent ability to account for base-rate neglect, a phenomenon beyond the purview of previous exemplar models. This article reexamines ALCOVE's base-rate neglect predictions and shows that they are confined to a very limited set of circumstances. In most cases, ALCOVE is unable to produce base-rate neglect.

The ability to abstract categories from the presentation of exemplars is at the center of much human cognition and underlies many real-life tasks, ranging from concept formation during language acquisition to the classification of tornadoes by skilled observers (Federal Emergency Management Agency, 1983). Accordingly, empirical and theoretical investigations of categorization performance have received much attention in recent years. In the theoretical arena, focus has often been on comparisons between prototype and exemplar models, the two primary approaches to theories of categorization (Nosofsky, 1992). Exemplar-based models (e.g., Estes, 1986; Nosofsky, 1984, 1986) postulate that training instances are uniquely and distinctly represented in memory, with subsequent categorization judgments based on a test item's similarity to all memorized exemplars. Prototype models, in contrast, assume an abstract central representation of the category that summarizes information across exemplars without representing them individually (e.g., Homa, Sterling, & Trepel, 1981). A related development involves connectionist networks (e.g., Gluck & Bower, 1988; Shanks, 1991), in which weighted connections are adjusted during learning to capture the statistical relationship between exemplars and categories. Many networks are formally equivalent to prototype models because test stimuli are classified on the basis of their similarity to the central tendency of each category (Nosofsky, 1992, p. 150).

This article focuses on a recent theory of categorization (Attention Learning Covering Map [ALCOVE]; Kruschke, 1992) that combined an exemplar-based architecture with a connectionist learning rule. Much has been made of ALCOVE's ability to account for base-rate neglect, a phenomenon originally predicted by connectionist networks and previously thought to be incompatible with exemplar models. This article reevaluates the apparent success of ALCOVE by analyzing its ability to predict base-rate neglect. The examination reveals that the prediction is limited to very special circumstances and that, in most situations, ALCOVE is unable to show base-rate neglect, much like other exemplar models.

Exemplar Versus Network Models

Over the years, a distinct set of empirical strengths and weaknesses has been associated with exemplar models, on one hand, and prototype and network models, on the other. In general, when classification is observed after completion of training, the data favor exemplar models over adaptive networks or prototype models. For example, exemplar models correctly predict that similarity to specific training instances plays an important role in classification performance (e.g., Medin & Schaffer, 1978; Nosofsky, 1991). Exemplar models can also accommodate the finding that certain linearly separable categories are learned more slowly than certain nonlinearly separable ones (Medin & Schwanenflugel, 1981), contrary to the predictions of some connectionist networks (cf. Gluck, Bower, & Hee, 1989).

Networks, on the other hand, have the edge over exemplar models in at least two ways. First, networks can account for performance during all stages of category learning, including trial-by-trial performance on a probabilistic classification task (e.g., Estes, Campbell, Hatsopoulos, & Hurwitz, 1989); exemplar models can do so only with additional assumptions and parameters (Nosofsky, Kruschke, & McKinley, 1992). Second, predictions derived from the error-driven learning rule embodied in most networks have led to the discovery of a variety of phenomena, among them blocking (Shanks, 1991) and—of particular interest here—base-rate neglect (e.g., Estes et al., 1989; Gluck & Bower, 1988). In the standard base-rate neglect setting, participants learn to classify a hypothetical patient as suffering from one of two possible diseases on the basis of symptoms that

I wish to thank John K. Kruschke for making available the source code for ALCOVE, for providing an independent replication of some of the present results, and for numerous stimulating discussions. Helpful comments on an earlier version of this article were provided by J. Vivien Bainbridge, Joshua Hurwitz, and Bennet Murdock. The revision process was aided by reviews provided by Mark Gluck, John Kruschke, Robert Nosofsky, and David Shanks.

Correspondence concerning this article should be addressed to Stephan Lewandowsky, Department of Psychology, University of Oklahoma, Norman, Oklahoma 73019-0535. Electronic mail may be sent to lewan@uoknor.edu.

are probabilistically related to both alternatives. On each learning trial, a set of symptoms is presented, and feedback is provided after the participant "diagnoses" the "patient." One of the symptoms (f_1) has a moderately high conditional probability (.6) of being present given one of the diseases (rare disease, or R) but a low probability (.2) of occurring with the other disease (common disease, or C). However, because disease C is three times more common overall (.75 of all learning trials) than disease R (.25), the presence of f_1 alone, according to Bayes's theorem, is not predictive of either disease. Nonetheless, participants consistently fall short of making full use of base rates and overpredict the rare disease (R) from the presence of f_1 alone, exactly as predicted by adaptive networks and quite contrary to the predictions of exemplar models. This base-rate neglect phenomenon constitutes a highly diagnostic empirical benchmark for models of categorization.

Hybrid Models of Categorization

The seemingly incompatible strengths and weaknesses of the exemplar and network approaches have been reconciled by recent hybrid models of categorization. These models combine exemplar-based representational assumptions with connectionist learning rules (Hurwitz, 1990; Kruschke, 1990, 1992). For example, Kruschke's (1990, 1992) ALCOVE rivals exemplar models in its ability to accommodate an impressive variety of data previously thought to present difficulties for networks. For example, ALCOVE learns nonlinearly separable categories faster than linearly separable ones; it can accommodate categories involving correlated dimensions; it does not suffer from the unrealistically rapid forgetting that affects other networks; and it acquires exceptions to rules in ways that are consistent with children's language acquisition performance (Kruschke, 1992). At the same time, unlike conventional exemplar models and much like simple connectionist networks, ALCOVE accounts for base-rate neglect (Kruschke, 1992). It thus appears that ALCOVE has combined the most desirable properties of exemplar models and adaptive networks, a claim that has been supported by limited empirical comparisons (Nosofsky et al., 1992).

ALCOVE: Summary of the Theory

ALCOVE consists of three layers of connectionist units: a layer of input units to represent the stimulus dimensions, a layer of hidden units corresponding to all training exemplars, and a layer of output units representing the set of applicable categories. In contrast to most other networks, the hidden units in ALCOVE have fixed locations that correspond to particular conjunctions of input values. Consider a binary two-dimensional input space with four possible exemplars: {0, 0}, {0, 1}, {1, 0}, and {1, 1}. Each exemplar j would map into the location of a hidden unit whose activation is computed as follows:

$$a_j^{hid} = \exp[-c \sum_i^{n_{in}} \frac{1}{n_{in}} \mid h_{ji} - a_i^{in} \mid], \tag{1}$$

where c is a freely estimated parameter called the specificity of the unit and n_{in} refers to the number of input dimensions.[1] Hence, a hidden unit is activated to the extent that its location $\{h_{j1}, h_{j2}\}$ is close, in a city-block metric with an exponential similarity function, to the currently presented exemplar. Presentation of, for example, the exemplar {1, 1} would fully activate the hidden unit at that location ($a^{hid} = 1.0$), whereas those located at {1, 0} and {0, 1} would receive less activation ($a^{hid} = e^{-0.5c}$) and that at {0, 0} less still ($a^{hid} = e^{-c}$). The partial and similarity-graded overlap between receptive fields of hidden units is a critical feature of the theory and is at the heart of ALCOVE's capability to generalize to novel instances and to account for a large body of categorization data, including base-rate neglect (Kruschke, 1992).

In a base-rate neglect situation, learning is confined to the weights connecting the hidden units to the output layer (w_{kj}s) with the standard error-driven learning rule:

$$\Delta w_{kj} = \lambda(t_k - a_k^{out})a_j^{hid}, \tag{2}$$

where λ is a learning parameter, t_k is the desired target activation of output unit k, and a_k is the current activation. Target activations are computed as follows:

$$t_k = \begin{cases} \max(+1, a_k^{out}) \text{ if the stimulus is in category } k \\ \min(-1, a_k^{out}) \text{ otherwise.} \end{cases} \tag{3}$$

Current activations of output units are given by the sum of the weighted inputs received from all (n_{hid}) hidden units:

$$a_k^{out} = \sum_j^{n_{hid}} w_{kj} a_j^{hid}. \tag{4}$$

Equations 1 through 4 illustrate ALCOVE's close relationship to both classes of categorization models. The exemplar-based hidden units and the use of an exponential similarity gradient (Equation 1) are akin to Nosofsky's (e.g., 1986) generalized context model (GCM). However, unlike GCM, with its constant-increment rule, ALCOVE uses connectionist learning principles to adjust the weights between each exemplar and the output units (Equations 2 and 3), thus also rendering it akin to connectionist networks (e.g., Gluck & Bower, 1988). Kruschke (1993) provided further details about ALCOVE's relationship to other theories of categorization.

ALCOVE: Base-Rate Neglect Predictions

The discovery of base-rate neglect (Gluck & Bower, 1988) has been followed by much commentary and numerous follow-up studies (e.g., Estes et al., 1989; Gluck & Bower, 1990; Markman, 1989; Shanks, 1990), all primarily concerned with the proper theoretical interpretation of the finding and all supportive of its empirical robustness and replicability. Table 1 shows the probability structure of the particular sequence of 240 training trials used by Estes et al. (1989), by Nosofsky et al. (1992), and in all simulations of ALCOVE to date. Attention should focus on the first symptom (f_1), which occurs equally often with each disease and is therefore without any objective diagnostic value when

[1] The present discussion is limited to those parameters and features of the theory that enter into the base-rate neglect predictions. Numerous other important features not relevant here are omitted and can be found in Kruschke (1990, 1992).

Table 1
Conditional Probabilities and Frequencies of Symptoms and Diseases in Simulation 1

	Disease			
	Rare		Common	
Symptom	Probability	Frequency	Probability	Frequency
f_1	.6	36	.2	36
f_2	.4	24	.3	54
f_3	.3	18	.4	72
f_4	.2	12	.6	108

presented on its own. However, as shown in Table 2, during training f_1 occurs more frequently by itself (i.e., Pattern 9) with the rare disease (a frequency of 16) than with the common disease (a frequency of 4).[2] The feedback associated with a response on those learning trials thus preferentially reinforces the rare disease; this, in turn, leads to base-rate neglect at a later test when f_1 is presented by itself and the participant must estimate the probability of the rare disease being present (for further details, see Gluck & Bower, 1988, or Kruschke, 1992, p. 31).

Consider the first two rows in Table 3, which show, respectively, the objective diagnostic value of each feature on its own (computed from Bayes's theorem; note the .50 for f_1) and, for comparison, representative probabilities with which participants choose the rare disease when given single symptoms (Nosofsky et al., 1992, Experiment 1). The difference between participants' $P(R \mid f_1)$ of (.75) and the objective value (.50) represents base-rate neglect. The next two rows of Table 3 show the base-rate neglect predictions obtained from a standard adaptive network (Estes et al.'s [1989] predictions for Experiment 2) and from ALCOVE (Kruschke, 1992, Figure 8). It is apparent that the models deviate from the objective values in a way analogous to the behavior of human participants.

Table 2
Symptom Combinations and Frequencies for All Training Patterns Used in Simulation 1

	Symptom				Frequency	
Pattern	f_1	f_2	f_3	f_4	Rare disease	Common disease
1	–	–	–	–	4	24
2	–	–	–	+	2	40
3	–	–	+	–	5	16
4	–	–	+	+	2	23
5	–	+	–	–	7	11
6	–	+	–	+	1	15
7	–	+	+	–	3	8
8	–	+	+	+	0	7
9	+	–	–	–	16	4
10	+	–	–	+	4	5
11	+	–	+	–	3	4
12	+	–	+	+	0	10
13	+	+	–	–	5	4
14	+	+	–	+	3	5
15	+	+	+	–	5	1
16	+	+	+	+	0	3

Note. Presence of symptom is denoted by +; absence is denoted by –.

ALCOVE: Simulations of Base-Rate Neglect

One peculiarity of base-rate neglect research to date has been the use of a single specific sequence of 240 learning trials in many behavioral experiments (e.g., Estes et al., 1989; Nosofsky et al., 1992) and in all simulations of ALCOVE (Kruschke, 1990, 1992; Nosofsky et al., 1992). The two simulations presented next show that ALCOVE's base-rate neglect predictions are tied to use of that specific sequence of learning trials.

Simulation 1: Sequence Sensitivity

Parameters were set to the best-fitting estimates provided by Kruschke (1992, p. 32), with $\lambda = .0393$, $c = 2.553$, and $\phi = 1.056$. The last parameter is a scaling factor required to convert the activations of output units to response probabilities (see Kruschke, 1992, Equation 3), and λ and c refer, respectively, to the learning rate and the specificity of the receptive fields of the hidden units. The input layer consisted of 4 units set to 0 or 1 to indicate absence or presence, respectively, of the corresponding symptom. At the hidden layer, 16 units represented the set of training patterns (shown in Table 2), and 2 output units corresponded to the two diseases. Predictions were obtained after completion of training by presenting test patterns with a single symptom (analogous to test trials in typical behavioral experiments; e.g., Estes et al., 1989). For these test trials, activations of hidden units were computed by collapsing across absent input dimensions.

The results of Simulation 1 are shown in Table 3. The simulation used the same sequence of 240 training trials used by Kruschke (1992), except that a minor modification was introduced for the second-to-last trial (row labeled 239) or the last 2 trials (239 + 240). For those trials, the correct outcome was reversed, such that the common disease served as the target response and was used for feedback during weight update. This minor deviation from the exact conditions used by Kruschke (1992) completely eliminated base-rate neglect in ALCOVE, even though the overall probabilistic structure of the training sequence was virtually unaffected. Indeed, because the last 2 trials involved Patterns 5 and 7 (see Table 2), the reversal of the correct outcome removed two instances in which the rate dis-

[2] The present discussion assumes that symptoms are either present or absent. In several experiments (e.g., Nosofsky, Kruschke, & McKinley, 1992), substitutive features have been used instead, in which presence versus absence is replaced by two opposing symptoms (e.g., high vs. low blood pressure). Substitutive features reduce the potential confusion between learning trials in which only one symptom is present (and the absence of the others has informational value; see Pattern 9 in Table 2) and test trials in which a judgment must be made on the basis of a single feature (and in which the absence of the others has no informational value). This is relevant in behavioral experiments and in determining the input representation for the simple adaptive network (cf. Gluck & Bower, 1988, p. 239), but it has little bearing on the ALCOVE simulations reported here.

Table 3
Observed and Predicted Choice Probabilities for Rare Disease Given a Single Symptom

	Symptom			
Source	f_1	f_2	f_3	f_4
Objective	.500	.310	.200	.100
Nosofsky et al. (1992)	.750	.465	.346	.215
Network (Estes et al., 1989)	.620	.180	.300	.050
ALCOVE (Kruschke, 1992)	.628	.379	.258	.119
Simulation 1				
239	.490	.230	.170	.070
239 + 240	.340	.130	.090	.040
Simulation 2				
Lewandowsky (1993)	.830	.330	.220	.060
ALCOVE: fixed parameters	.360	.200	.130	.060
ALCOVE: best fitting	.460	.310	.230	.140

ease was reinforced without f_1 being present. If anything, this may be expected to enhance base-rate neglect because two additional learning trials exist that associate the common disease with the absence of f_1 and, by plausible inference, the rare disease with its presence.

Simulation 2: Long-Run Predictions

The first simulation showed that, at least with the parameter values used here, base-rate neglect in ALCOVE is tied to the particular sequence of training trials used by Estes et al. (1989) and Nosofsky et al. (1992) and in all simulations by Kruschke (1990, 1992). Although this apparent sensitivity presents cause for concern, the possibility remains that the same parameter values can, in the long run (across numerous random sequences of properly structured training trials), give rise to base-rate neglect. Alternatively, other sets of parameter values may exist that would allow the model to predict base-rate neglect under a wider variety of circumstances.

This should be reasonably expected of ALCOVE. Table 3 shows the results of an unpublished experiment conducted by Lewandowsky (1993) in which each of 18 participants received a different random sequence of 120 training trials. After training, participants were presented with single symptoms and had to choose which disease was present. The data confirmed that participants reliably exhibited base-rate neglect with any sequence that implemented the requisite overall probability structure shown earlier in Table 1. Similarly, the original network predictions of Gluck and Bower (1988) were derived analytically from the asymptotic probability structure of the training trials and, therefore, also did not depend on a particular sequence of trials.

Table 3 also shows the predictions of ALCOVE for the parameter values used in Simulation 1; however, the values are averaged across 200 randomly generated sequences of training trials, each with the proper probability structure. There was no trace of base-rate neglect, and the results therefore rule out the possibility that ALCOVE, with the present set of parameter values, can predict base-rate neglect in the long run. Finally, the bottom row of Table 3 shows the results for another 200 randomly generated training sequences but with a new set of best-fitting parameter values. Parameters were estimated by minimizing the root mean squared deviations (RMSD) between the choice probabilities observed by Lewandowsky (1993) and those predicted by ALCOVE averaged across the 200 training sequences. The best-fitting estimates were $\phi = 0.356$, $\lambda = .324$, and $c = 5.0$. It is clear from the table (and the RMSD of .19) that base-rate neglect was again absent.[3]

Taken together, the first two simulations suggest that ALCOVE, unlike the adaptive network and unlike human participants, can produce base-rate neglect only under a narrow set of circumstances, namely, when parameters are estimated with one particular sequence in which the rare disease is the correct choice on the last two trials. These results run counter to earlier explanations offered for ALCOVE's capabilities (e.g., Kruschke, 1992) and thus demand further analysis and explanation.

Analysis of Weight Changes

Recall that ALCOVE, like other connectionist networks, learns by gradient descent (i.e., by adjusting the weights such that the "error," or discrepancy, between the network's current output and the desired target is reduced). In the simulations reported here, learning was restricted to the connections between hidden and output layers. It follows that the presence, as well as the absence, of base-rate neglect must result from the behavior of those weights during learning, as formalized in Equation 2. Equation 2 also points to further analysis of the present results. Disregarding the learning parameter λ, the right-hand side of Equation 2 corresponds to the derivative of the weights with respect to the error (McClelland & Rumelhart, 1988, p. 130).

The vector of all such weight-error derivatives (WEDs) points in the direction of steepest descent down the error surface, with

[3] Training sequences were composed of 120 trials each to conform to the experimental details of Lewandowsky (1993). Parameters were also estimated with longer sequences (240 trials), and the predictions were found to be virtually identical. In addition, ALCOVE was fitted individually to five different random sequences of 240 training trials. Base-rate neglect was observed only in the one case in which, by chance, the rare disease was the correct response for the last trial.

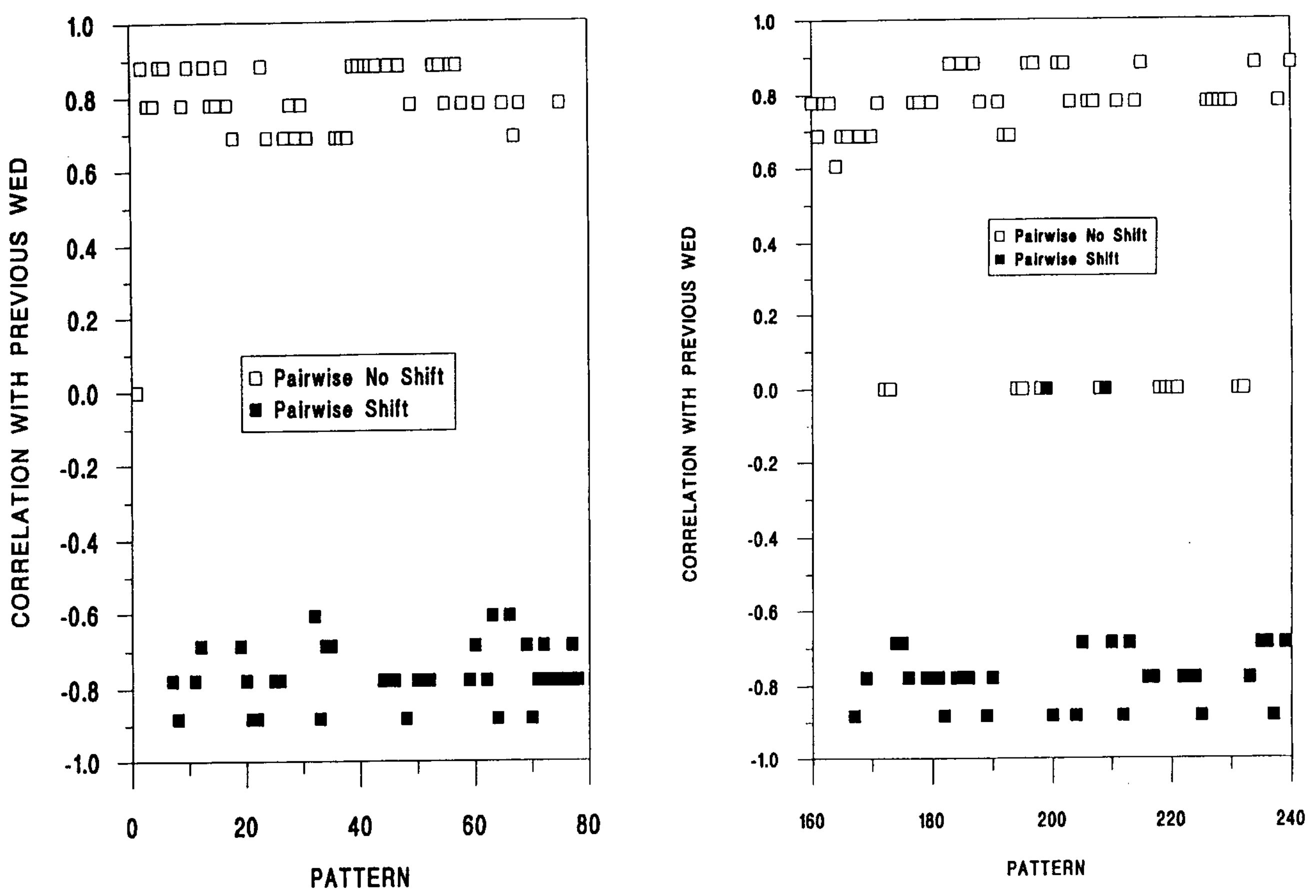

Figure 1. Observed vector correlations (cosines) between successive weight-error-derivative (WED) vectors in Simulation 1 early (left panel) and late (right panel) in training.

the length of the vector indicating the step size. Hence, analysis of successive WED vectors during learning describes the path taken by the network when minimizing the error between output and target (e.g., Hetherington, 1991). In particular, comparison of the direction of each WED vector with its predecessor from the previous trial using vector correlation (cosine of the angle between the two vectors) reveals whether the descent is smooth and relatively straight (large positive correlations) or involves backtracking (negative correlations). Figure 1 shows the pairwise correlations between successive WED vectors in ALCOVE across the sequence of base-rate neglect learning trials in Simulation 1. The left panel focuses on early learning trials, and the right panel refers to the end of learning.

The WED correlations are graphically differentiated according to whether or not the same disease was the correct response for both successive trials. It is clear that the correlations fall into one of two clearly demarcated clusters: large positive correlations for trials involving the same disease (no shift) and large negative correlations when disease is shifted between trials. Thus, gradient descent continues in a fairly uniform direction if the target response remains the same but changes direction when the target is changed. By comparison, the effect of the different input patterns is relatively minor, as shown by the homogeneity of each cluster.

Two qualifications apply to interpretation of the WEDs. First, the pattern of correlations alone cannot be used to infer the presence or absence of base-rate neglect. Second, some change in direction of weight update on shift trials is a necessary consequence of the learning rule (Equation 2) and the way in which target activations are computed (Equation 3). If target diseases are altered between trials, the sign of the weight updates is reversed, and, by implication, a negative correlation with the previous WED vector must result. However, the large absolute magnitudes of those correlations are not dictated by the learning rule, and they reveal that weights are pulled in virtually the opposite direction whenever target responses are changed. Thus, even though ALCOVE may predict base-rate neglect when the most recent weight update favored the rare disease, this ability is lost rapidly with reinforcement of the common disease.

The WED analysis can be extended by plotting the values of the weights themselves across learning trials.[4] Figure 2 shows the values of the 16 weights connecting the hidden layer to the output unit for the rare disease during the last 20 trials of the

[4] I wish to thank John Kruschke for convincing me of the utility of the weight analysis.

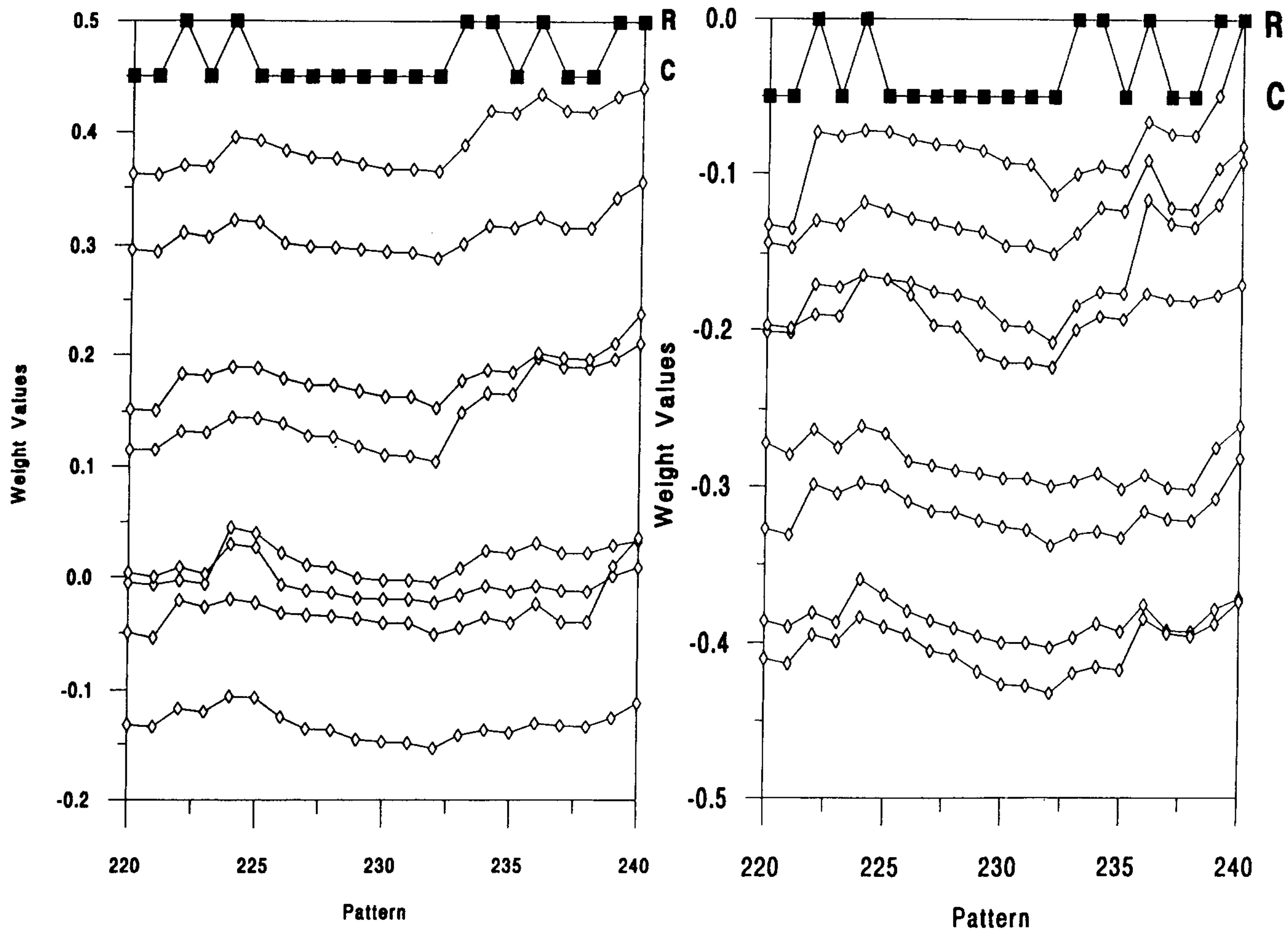

Figure 2. Values of the 16 weights connecting the hidden layer to the output unit for the rare disease during the final 20 learning trials of Kruschke's (1992) sequence. Weights were arbitrarily assigned to panels on the basis of their absolute value. The correct response (rare [R] or common [C]) for each training pattern is shown at the tops of the panels.

training sequence used by Kruschke (1992). The figure also indicates the target disease (*R* or *C*) for each pattern.

In confirmation of the earlier WED analysis, the weights in Figure 2 collectively respond to each change of target disease: All weights increase in the presence of the rare disease, and they all decrease in the presence of the common disease. Perhaps more noteworthy is the fact that the absolute magnitude of each weight change is fairly small. The largest individual weight change on a single trial is approximately .05, and most weights move by an even smaller amount. Moreover, the total weight change across the 20 trials is negligible, suggesting that ALCOVE has reached a relatively stable state. Nonetheless, under exactly the same circumstances, reversal of the correct disease for the last 2 trials eliminates base-rate neglect. It follows that small weight changes in ALCOVE can have large unanticipated consequences.

The simulation results, in conjunction with the WED and weight analyses, form a fairly unequivocal picture that contradicts previous verbal explanations of why ALCOVE produces base-rate neglect (e.g., Kruschke, 1990, 1992; Nosofsky et al., 1992). Those explanations were based on the fact that particularly strong association weights are established between the rare disease and certain exemplar nodes (those containing f_1) that have close neighbors favoring the common disease. In consequence, when f_1 is presented on its own, the rare disease is preferred (see Kruschke, 1992, p. 32, for additional details). The WED analysis clarified that the weights were less robust than assumed by Kruschke and colleagues, and inspection of the weight changes revealed that small changes that verbal explorations of a theory would neither detect nor deem significant can have large and unanticipated ramifications. In consequence, Kruschke's explanation is of little generality.

Conclusions

This article has focused on a purported crucial property of the ALCOVE model of categorization that differentiates it from, and potentially raises it above, conventional exemplar models: its ability to predict base-rate neglect. The article has demonstrated that this prediction is confined to some very special circumstances and that, in general, the theory is unable to predict base-rate neglect. This does not negate the impressive explanatory power of ALCOVE in other situations, but it suggests that additional data-based comparisons must be sought to differentiate the theory from conventional exemplar models and to justify its additional computational complexity.

References

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology, 18,* 500–549.

Estes, W. K., Campbell, J. A., Hatsopoulos, N., & Hurwitz, J. B. (1989). Base-rate effects in category learning: A comparison of parallel network and memory storage-retrieval models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 556–571.

Federal Emergency Management Agency. (1983). *Tornado resource workbook.* Washington, DC: U.S. Government Printing Office.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117,* 227–247.

Gluck, M. A., & Bower, G. H. (1990). Component and pattern information in adaptive networks. *Journal of Experimental Psychology: General, 119,* 105–109.

Gluck, M. A., Bower, G. H., & Hee, M. R. (1989). A configural-cue network model of animal and human associative learning. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society* (pp. 323-332). Hillsdale, NJ: Erlbaum.

Hetherington, P. A. (1991). *The sequential learning problem in connectionist networks.* Unpublished master's thesis, McGill University, Montreal, Quebec, Canada.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 418–439.

Hurwitz, J. B. (1990). *A hidden-pattern unit network model of category learning.* Unpublished doctoral dissertation, Harvard University, Cambridge, MA.

Kruschke, J. K. (1990). *ALCOVE: A connectionist model of category learning* (Research Report 19). Bloomington: Cognitive Science Program, Indiana University.

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

Kruschke, J. K. (1993). Human category learning: Implications for backpropagation models. *Connection Science, 5,* 3–36.

Lewandowsky, S. (1993). *Learning and forgetting of base-rate neglect.* Manuscript in preparation.

Markman, A. B. (1989). LMS rules and the phenomenon of base-rate neglect. *Journal of Experimental Psychology: General, 118,* 417–421.

McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing.* Cambridge, MA: MIT Press.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Journal of Experimental Psychology: General, 117,* 68–85.

Medin, D. L., & Schwanenflugel, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 355–368.

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

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

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 3–27.

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

Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 211–233.

Shanks, D. R. (1990). Connectionism and human learning: Critique of Gluck and Bower (1988). *Journal of Experimental Psychology: General, 119,* 101–104.

Shanks, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 433–443.

Received November 29, 1993
Revision received April 28, 1994
Accepted April 29, 1994 ■

New Editors Appointed, 1996–2001

The Publications and Communications Board of the American Psychological Association announces the appointment of two new editors for 6-year terms beginning in 1996. As of January 1, 1995, manuscripts should be directed as follows:

- For *Behavioral Neuroscience,* submit manuscripts to Michela Gallagher, PhD, Department of Psychology, Davie Hall, CB# 3270, University of North Carolina, Chapel Hill, NC 27599.
- For the *Journal of Experimental Psychology: General,* submit manuscripts to Nora S. Newcombe, PhD, Department of Psychology, Temple University, 565 Weiss Hall, Philadelphia, PA 19122.

Manuscript submission patterns make the precise date of completion of 1995 volumes uncertain. The current editors, Larry R. Squire, PhD, and Earl Hunt, PhD, respectively, will receive and consider manuscripts until December 31, 1994. Should either volume be completed before that date, manuscripts will be redirected to the new editors for consideration in 1996 volumes.