To appear in *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*. Cambridge, MA: Lawrence Erlbaum.

Computational Models of Category Learning

Dorrit Billman (Billman@pravda.gatech.edu)
School of Psychology, Georgia Institute of Technology, Atlanta, Georgia 30332

Douglas Fisher (DFisher@vuse.vanderbilt.edu)
Department of Computer Science, Vanderbilt University, Nashville, TN 37235

Mark Gluck (Gluck@psych.stanford.edu)
Department of Psychology, Stanford University, Stanford, CA 94305

Pat Langley (Langley@ptolemy.arc.nasa.gov)
AI Research Branch, MS 244-17, Nasa Ames Research Center, Moffett Field, CA 94035

Michael Pazzani (Pazzani@ics.uci.edu)
Department of Information and Computer Science, University of California, Irvine, CA 92717

1 Introduction

Computational models have played an important role in cognitive science, giving researchers a precise, unambiguous language in which to express theories of cognition. Early work in cognitive simulation focused on the collection of verbal protocols for a small number of subjects, the detailed analysis of those protocols, and the creation of programs that simulated their behavior in considerable detail. This approach, exemplified by the work of Newell and Simon (1972), has led to important insights about human cognition, particularly in the area of problem solving.

Another early approach, exemplified by Feigenbaum's (1963) work on Epam, focused on robust empirical generalizations, showing how such phenomena arose as emergent properties of a computational model. Although originally less common that the former approach, in recent years this research paradigm has been gaining ascendancy. In this paper, we consider theories of category learning that have taken this form. First, we summarize several memory, reasoning, and learning phenomena that models of category learning must explain. Next, we give brief overviews of three category learning models and indicate how these models account for some of the empirical findings. Finally, we discuss some open issues and consider promising directions for future research.

2 Empirical Generalizations about Categorization

Theories attempt to explain empirical phenomena, so we begin by reviewing some generalizations that have emerged from the experimental study of human categorization. We have not attempted to be exhaustive, but we believe the statements that follow provide important constraints on theories of category learning.

1. People are able to represent, access, and acquire concepts that involve logical 'rules' (Bourne, 1966), but they can also handle 'fuzzy' categories for which no 'logical' rules exist (Smith & Medin, 1981; Barsalou, 1985). For example, one might define birds as flying, beaked animals, but some birds cannot fly and some have bills instead of beaks.
2. Categories are influenced by the informational structure of the environment (Rosch, Mervis, Johnson, Gray, & Boyes-Braem, 1976), in that different experiences lead to different concepts. However, they are also influenced by the goals of the perceiver (Barsalou, 1983a, 1983b) and by intuitive beliefs and theories of the world (Chapman & Chapman, 1969; for discussion, Murphy & Medin, 1985).
3. People can detect and exploit correlations among features (Medin, Altom, Edelson, & Freko), as well as information about the

frequency of categories (Medin & Edelson, 1988; Gluck & Bower, 1988a). For instance, feathers are positively correlated with beaks, whereas fur tends to co-occur with teeth. Further, correlations are learned more easily when part of a coherent set of relations (Billman, 1989; Billman & Jeong, 1989).

4. People divide categories into subcategories, with some levels being more 'natural' than others. These 'basic' categories tend to occur at intermediate levels of abstraction. For instance, the category 'bird' is more basic than the categories 'animal' or 'robin'. Research with natural and artificial categories indicates that such concepts are learned earlier developmentally and more rapidly in experiments (Corter, Gluck, & Bower, 1988; Rosch et al., 1976). Objects are often identified at the basic level most rapidly, though this interacts with the typicality of the object (Murphy & Brownell, 1985).

5. Some instances are more 'typical' of a category than others (Rosch & Mervis, 1975). These are named more frequently and accessed more rapidly than less typical ones. For instance, robins are more typical birds than are penguins, and pictures of the former are recognized as birds more quickly than the latter. However, typicality varies across individuals and contexts (Barsalou, 1985, 1987, 1989). Moreover, there is dissociation between membership and typicality in some categories (Armstrong, Gleitman, & Gleitman, 1983) but not in others (Fehr & Russell, 1984; Hampton, 1987).

6. Just as category learning is influenced by prior beliefs, recognition of items as category members includes a top-down aspect: entities are categorized faster in expected contexts than in unexpected contexts (Palmer, 1975). For example, a drawing of a loaf of bread is recognized more rapidly when located in a kitchen than in a street scene.

7. People can represent, access, and acquire categories that involve structural and conceptual relations between components (Barsalou, in press; Fodor & Pylyshyn, 1988). For instance, the relative locations of the eyes and nose are essential aspects of the concept 'face'.

8. People use categories to guide inference as well as classification and recognition. Inferences about new instances are guided by category membership. For example, given that a novel entity can be identified as a bird, one can infer that it probably hatched from an egg.

9. Categories are used in inferences about new properties. Generalization of new properties to individual instances and to sets of instance is guided by category membership (Gelman & Markman, 1987; Nisbett, Ross, Jepson, & Kunda). Indeed, under some conditions, people generalize properties of a single exemplar to an entire category (Osherson, Smith, Wilkie, Lopes, & Shafir, in press; Macario, Shipley, & Billman in press).

In sum, theories should eventually account for category learning, for classification, and for use of categories in inference. Both characteristics of the input (e.g., correlations among attributes, feature frequency) and background knowledge influence all these processes. In the following sections, we provide overviews of three computational models of category formation. In each case, we describe the model's representation of conceptual knowledge, the manner in which that knowledge is used, and the learning mechanisms through which it is acquired. We also examine each model's ability to explain some subset of the empirical generalization listed above.

3 Configural-Cue Adaptive Networks

Gluck and Bower (1988a, 1988b) describe an adaptive network model of human learning that extends Rescorla and Wagner's (1972) theory of classical conditioning to human classification learning. The model represents knowledge as a set of one-layer classifiers, one for each category. Each network has a set of input that correspond to features that may occur in an experience, a set of weighted links, and a single output node. Given a new experience, one outputs a classification probability by adding the weights on matched input features. Learning involves changing the weights on links using Widrow and Hoff's (1960) least mean squares method. Briefly, this alters weights so as to decrease the difference between the actual and desired score for each classifier.

This adaptive network model has accurately accounted for human behavior in experiments on probabilistic classification learning with multiple cues, but it can only learn 'linearly separable' categories, that is, ones that can be separated by a hyperplane through the space of instances. However, one can extend the model to non-linearly separable categories by letting conjunctions of elementary stimulus features serve as 'higher-order' features of a stimulus pattern. Thus, given the presentation of an experimental pattern consisting of elementary features BCD, one assumes that this is reflected not only through activation of input nodes for the single elements B, C, and D, but also through activation of the pair-wise conjuncts BC, BD, and CD. In this approach, a domain involving N elementary features would be represented using $N^2 + N$ input features for each category, but the categorization and learning mechanisms would remain unchanged.

Gluck, Bower, and Hee (1989) have shown how this extended model accounts for several aspects of complex category learning by humans. Moreover, Corter, Gluck, and Bower (1988) have applied the configural-cue approach to model basic-level effects in hierarchically organized categories. For instance, when the network model is trained on three levels of an artificial domain involving hierarchical category structure, high levels of activation are reached sooner for the output nodes corresponding to the intermediate (basic) level categories. The predicted learning curves closely resemble the observed curves for humans, and the model also correctly predicts a shift in relative difficulty between the subordinate and superordinate levels in different experimental conditions.

A key property of the configural-cue model is that it embodies, implicitly, an approximate exponential decay relationship between stimulus similarity and psychological distance, a relationship with considerable independent support in studies of stimulus generalization (Shepard, 1958) and categorization (Nosofsky, 1984). This effect can be seen by noting how the number of overlapping active nodes (similarity) changes as a function of the number of overlapping component cues (distance). If two triplet patterns share one feature (ABC, XYC), they will have only one active node in common and five nodes nonoverlapping; if they share two features (ABC, XBC), they will have three active nodes in common (two component cues and one configural-cue node) and three nonoverlapping nodes. In fact, the network model can be viewed as an extension of Shepard's (1987) theory of stimulus generalization to classification learning.

In addition to basic-level effects, the configural-cue model is consistent with many of the phenomena from Section 2. It can certainly acquire nonlogical categories, and its use of pair-wise features lets it capture correlations. The concepts learned by the model are certainly a function of the environment it experiences, though this is represented only in different weights. In summary, combining the adaptive network model with a representation of stimuli that includes pair-wise configurations of features lets one account for a wider range of learning results from both the animal and human learning literatures.

The configural-cue model has several obvious limitations, including the rapid growth of input nodes with increasing pattern size. In addition, the approach can only make predictions about prelabeled classes, and it cannot handle structural representations of knowledge. Nevertheless, the model is theoretically parsimonious, accounts for a wide range of phenomena, and uses assumptions for which independent evidence already exists. Furthermore, its successes are instructive in identifying empirical phenomena that can be explained as emergent from the same elementary, associative processes found in lower species.

4 Representation Change in Concept Formation

Although some concept-learning tasks involve supervision, people also acquire conceptual knowledge without explicit supervision. The research described in this section takes the primary task of category learning, particularly unsupervised learning, to be recovery of the correlational structure of input. This produces coherent categories useful in prediction and inference. Models developed in this framework directly represent correlational structure as probabilistic patterns or rules. Further, these models assume category formation is intimately linked to representation change, and specifically to change in the attributes and features used to represent input. Representation change is important, but little studied from the perspective of concept learning. We distinguish between two types of representation change.

'Weak' representation change involves a change in attention to or significance of a property. In contrast, 'strong' representation change involves the introduction of new properties (a 'limiting case' of increase in attention).

CARI (Billman & Heit, 1988) and descendant models (Chalnick & Billman, 1988) investigated procedures for attention change in unsupervised learning. In particular, increasing attention is directed to attributes that prove predictive in some relation; this facilitates the discovery of other related predictive patterns when input exhibits any of a broad class of coherent structures (psychological motivation from Billman, 1989). More recently, Billman and Martin have explored a context-sensitive form of attentional learning. The idea is derived from the philosopher Goodman's (1955, 1984) concept of overhypotheses (see also Russell, 1986, on well-defined categories). Categories of a particular type (jewels, animals, ethnic groups) are homogeneous – or predictable – with respect to some properties (crystal structure, diet, language) but not others (weight, age, first name). Members of one category share values of the homogeneous attributes and contrast with other categories on these attributes. This information can be used to make inferences about novel categories and attribute values. Seeing one instance of a new kind of jewel, one would generalize its crystal structure but not its weight to other category members. Further, people do seem to make use of this information (Nisbett, Krantz, Jepson, & Kunda, 1983; Shipley 1989).

In related work, Billman and Martin have investigated two types of 'strong' representation change. One concerns the formation of new attributes from previously uncoordinated features; thus, someone may come to see legs, fins, and wings as values of a new attribute 'limbs'. The second involves the formation of new features from conjoining old features or attribute values. Martin and Billman (in press; Martin, 1989) describe CORA, a computational model of this latter process, which we now describe in some detail.

CORA represents conceptual knowledge in terms of probabilistic inference rules. For instance, it might have the rule IF X HAS WINGS AND X HAS FEATHERS, THEN INFER THAT X FLIES WITH PROBABILITY 0.95. In this case, a conjunction of two features predicts the presence of a third, but simpler and more complex rules are possible. CORA does not represent concepts as separate knowledge structures; rather, it organizes knowledge an interconnected network that directly represents conditional probabilities between features and conjunctions of features. The system also contains knowledge about features which are mutually exclusive, that is, which are alternative values of the same attribute.

Given a new experience with the values of some attributes omitted, CORA uses its inference rules to predict the missing values. For each attribute, it applies all rules that infer a value for that attribute and whose conditions match against the new experience. The system then estimates the overall probability for each possible value as the geometric average of the individual predicted probabilities. Finally, CORA predicts the value of the attribute that has the highest overall probability.

Billman and Martin's system begins with simple inference rules in which the condition sides contain only one feature, and in which all rules have the same associated probability. However, CORA augments this knowledge using two learning mechanisms. First, when the system is given a new experience, it iterates through each observed value, updating the conditional probability of each rule that infers that value. After this, it checks to determine whether the updated rules – in combination – actually predict the observed value over the alternatives. If they do not, this suggests that there are interactions among features which are not being taken into account, and CORA creates a new inference rule whose condition side is the conjunction of two existing rules. The component rules are those which have been most frequently associated with errors in the past. In this way, the system begins with simple inference rules and incrementally constructs more complex ones in which higher-order features are present.

The basic approach shares several characteristics with Gluck's configural-cue model. Both systems represent conjunctions of features, associate weights with these higher-order terms, and combine the relevant 'rules' to make predictions. However, the configural-cue model assumes that all pairwise conjunctions are present from the outset, whereas CORA begins with single features and constructs higher-order terms as necessary. Thus, Gluck's learning method searches only the space of weights on prespecified links, whereas Martin and Billman's system carries out a general-to-specific search through the space of feature combinations. Another difference is that CORA is unsupervised, in that it requires no labeled training instances and predicts the value of any missing attribute.

The CORA model is consistent with a number of the phenomena given in Section 2. The system can handle both logical 'rules' and more 'fuzzy' representations of knowledge, and it constructs different inference rules depending on the informational structure of the environment. CORA predicts that some experiences are more typical than others, but it makes no assumptions about reaction times. The model can certainly exploit correlations among features, and it emphasizes the use of conceptual knowledge for prediction. However, it provides no account of basic-level effects, and it fails to address the structural nature of some knowledge. Nevertheless, the approach provides a promising account of the representation, use, and acquisition of conceptual knowledge, and future work may address these issues.

5 Hierarchical Concept Formation

Many models of concept formation construct classification hierarchies over environmental observations. Perhaps the earliest such system was Feigenbaum's (1963) EPAM, which incrementally formed discrimination networks and used them to classify new observations. Lebowitz's (1982) UNIMEM and Kolodner's (1983) CYRUS built more sophisticated, redundant classification structures, but they sifted observations down alternative paths in a hierarchy much like EPAM.

Here we will focus on Fisher's (1987) COBWEB, which was intended to synthesize ideas from these earlier systems and thus will serve briefly to highlight characteristics of the general hierarchical approach.[1] In particular, the system uses a probabilistic representation of concepts (Smith & Medin, 1981). Each concept specifies a set of attributes and their possible values, along with the conditional probability of that value given the concept. COBWEB also stores the overall probability of occurrence for each concept. Moreover, concepts are organized into a concept hierarchy that is partially ordered according to specificity, with more abstract categories at higher levels, more specific ones at lower levels, and specific observations as terminal nodes.

Like its predecessors, COBWEB sorts new observations downward through its hierarchy, selecting the best branch at each level. In making this decision, it finds the 'best match' according to category utility, an evaluation function proposed by Gluck and Corter (1985) to account for basic levels observed in experimental studies of human categorization. They arrived at this function through a rational analysis (Anderson, in press) of the categorization task, in which basic-level concepts are preferred because they facilitate more accurate predictions about their members.

Like Martin and Billman's model, COBWEB acquires concepts incrementally, during the process of categorization. As it sorts experiences down the hierarchy, the system alters the probabilities stored with each concept and its associated values. In some cases, COBWEB also changes the structure of its hierarchy. If an experience differs sufficiently from the children of a concept, the system stores it as a new child at that level. In other cases, COBWEB finds that merging or splitting existing concepts at a given level will improve the match score. The structure of the resulting hierarchy is a function not only of the experiences given to the system, but of the order in which they are presented.

Note that COBWEB differs from CORA in that it does not explicitly compute correlations among features. Category utility sums a function of individual feature probabilities, so that each node in memory effectively represents a category as an independent cue model (Smith & Medin, 1981). Fisher and Langley (in press) have shown that favoring the creation of categories that maximize this function of individual features will *tend* to reward categories that capture correlations. However, leaves of the concept hierarchy correspond to specific observations (cases, exemplars), giving representational power equivalent to that of exemplar (Smith & Medin, 1981) or relational cue (Medin, 1983) representations. In addition, hierarchical systems offer a natural representation of default and exceptional properties: one assumes the most likely attribute value for a given category, unless and the observed features warrant deeper classification.

[1] We direct readers Gennari, Langley, and Fisher (1989) for a review of hierarchical approaches to concept formation.

Cobweb and most other hierarchical systems assume that observations are represented as nominal attribute-value pairs. However, recent work has extended the basic framework to handle real-valued attributes (Gennari, Langley, & Fisher, 1989) and structured descriptions (Thompson & Langley, 1989), in which acquired concepts are defined in terms of other acquired concepts and relations among them. Moreover, some researchers have adapted such relational representations for use in concept formation over traces of problem-solving experiences (e.g., Yang, Yoo, & Fisher, in press). In this approach, previous experience is organized in memory and accessed for use in guiding behavior on novel problems. Such models may explain the origin functional categories, as well as predicting psychological phenomena in problem-solving domains that are analogous to those found in simple categorization tasks.

Despite its initial motivation in terms of computational efficacy, Cobweb also accounts for a variety of psychological phenomena (Fisher & Langley, in press). Its account of basic-level effects follows from its use of category utility, but with assumptions about retrieval rates, accounts of typicality effects and fan effects (Anderson, 1976) also emerge. Like the other models we have examined, Cobweb can represent and acquire nonlogical categories, but it can handle logical rules as a special case. The system takes advantages of correlations among features, though it represents them only indirectly, and Fisher (1987) has extensively tested its predictive ability. Extensions to the framework show promise for structural knowledge, and the model even makes certain predictions about interactions between typicality and basic levels (Fisher, 1988). Thus, hierarchical methods for concept formation provide another promising framework for explaining the nature of human categorization.

6 Conclusions

We have presented three models that account for some empirical findings in category formation. No one model claims to account for all of the complexity of category formation. In some cases, techniques used in one model can address shortcomings of other models. For example, Cora suggests a way of going beyond pairwise conjunctions in the configural cue model while avoiding the enumeration of all possibilities. In some cases, the models provide contrasting accounts for the same phenomenon. For example, both Cobweb and the configural-cue model propose different mechanisms to account for basic-levels effects.

Some phenomena are not addressed by any of the models. In particular, none provide an explicit account for the priming effect (e.g., Palmer, 1975) of expected contexts, though the use of conditional probabilities in Cora and Cobweb suggest possible extensions. All of the current models emphasize the role of the informational structure of the environment, ignoring the role played by the learner's goals and theories of the world in category formation.

We have focused on demonstrating that empirical generalizations of human learning such as basic-level and typicality effects can arise as emergent properties of computational models. It is equally important that these computational models make predictions and be used to formulate hypotheses about human behavior that can be tested empirically. Furthermore, the precision demanded by computational models often raises issues that might have otherwise been overlooked. Thus, we feel that experimentation and computational modeling play essential complementary roles in understanding cognition.

Acknowledgements

We would like to thank Larry Barsalou for his advice on the complex topic of categorization and Joel Martin for comments on an earlier draft.

References

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. R. (in press). The place of cognitive architectures in a rational analysis. In K. Van Lehn (Ed.), *Architectures for intelligence.* Hillsdale, NJ: Lawrence Erlbaum.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). On what some concepts might not be. *Cognition, 13*, 263–308.

Barsalou, L. W. (1983a). Ad hoc categories. *Memory & Cognition, 11*, 211–227.

Barsalou, L. W. (1983b). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10*, 82–93.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 629–654.

Barsalou, L. W. (in press). Frames, concepts, and conceptual fields. In A. Lehrer & E. Kittay (Eds.), *Semantic fields and other alternatives*.

Billman, D. (1989). Systems of correlations in rule and category learning: Use of structured input in learning syntactic categories. *Language and Cognitive Processes, 4*, 127–155.

Billman, D. O., & Heit, E. (1988). Observational learning from internal feedback: A simulation of an adaptive learning method. *Cognitive Science, 12*, 587–626.

Billman, D., & Jeong, A. (1989). Systematic correlations facilitate learning component rules in spontaneous category formation. Paper presented at the 30th Annual Meeting of the Psychonomic Society, November, 1989, Atlanta, GA.

Bourne, L. E. (1966). *Human conceptual behavior*. Boston: Allyn and Bacon.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodynamic signs. *Journal of Abnormal Psychology, 74*, 272–280.

Corter, J. E., Gluck, M. A., & Bower, G. H. (1988). Basic levels in hierarchically structured categories. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*. Montreal, Canada: Lawrence Erlbaum.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113*, 464–486.

Feigenbaum, E. A. (1963). The simulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought*. New York: McGraw-Hill.

Fisher, D. H. (1987). Knowledge acquisition via incremental conceptual clustering. *Machine Learning, 2*, 139–172.

Fisher, D. (1988). A computational account of basic level and typicality effects. *Proceedings of the Seventh National Conference on Artificial Intelligence* (pp. 233–238). St. Paul, MN: Morgan Kaufmann.

Fisher, D. & Langley, P. (in press). The structure and formation of natural categories. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26). New York: Academic Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–71.

Gennari, J., Langley, P., & Fisher, D. (1989). Models of incremental concept formation. *Artificial Intelligence, 40*, 11–62.

Gluck, M. A., & Bower, G. H. (1988a). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, 27, 166–195.

Gluck, M. A., & Bower, G. H. (1988b). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117*, 225–244.

Gluck, M. A., & Corter, J. E. (1985). Information, uncertainty, and the utility of categories. *Proceedings of the Seventh Annual Conference of the Cognitive Science Society* (pp. 283–287). Irvine, CA: Lawrence Erlbaum.

Gluck, M. A., Bower, G. H., & Hee, M. R. (1989). A configural-cue network model of animal and human associative learning. *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*. Ann Arbor, MI: Lawrence Erlbaum.

Goodman, N. (1955/1984). *Fact, fiction, and forecast*. New York: Bobbs-Merrill.

Hampton, J. A. (1987). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 12–32.

Kolodner, J. L. (1983). Reconstructive memory: A computer model. *Cognitive Science*, 7, 281–328.

Lebowitz, M. (1982). Correcting erroneous generalizations. *Cognition and Brain Theory, 5*, 367–381.

Macario, J., Shipley, E., & Billman, D. (in press). Induction from a single instance: Formation of a novel category. *Journal of Experimental Child Psychology*.

Martin, J. D. (1989). Focusing attention for observational learning: The importance of context. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*. Detroit, MI: Morgan Kaufmann.

Medin, D. (1983). Structural principles of categorization. In T. Tighe & B. Shepp (Eds.), *Perception, cognition, and development*. Hillsdale, NJ: Lawrence Erlbaum.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 37–50.

Murphy, G. L., & Brownell, H. H (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 70–84.

Medin, D. L., & Edelson, S. M. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General*, *117*, 68–85.

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

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday deductive reasoning. *Psychological Review*, *90*, 339–363.

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

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (in press). Category based induction. *Psychological Review*.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*, 519–526.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts.

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

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

Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.

Russell, S. J. (1986). Preliminary steps toward the automatization of induction. *Proceedings of the Fifth National Conference on Artificial Intelligence*. Philadelphia, PA: Morgan Kaufman.

Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, *65*, 242–256.

Shepard, R. N. (1987). Towards a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.

Shipley, E. F. (1990). *Hierarchies and induction*. Unpublished manuscript, Department of Psychology, University of Pennsylvania, Philadelphia, PA.

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

Thompson, K., & Langley, P. (1989). Incremental concept formation with composite objects. *Proceedings of the Sixth International Workshop on Machine Learning* (pp. 371–374). Ithaca, NY: Morgan Kaufmann.

Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record*, *4*, 96–194.

Yoo, H., Yang, H., & Fisher, D. (in press). Concept formation over problem-solving experience: Issues of induction and utility. In D. Fisher & M. Pazzani (Eds.), *Computational approaches to concept formation*. San Mateo, CA: Morgan Kaufmann.