In Vivo or In Vitro: Cognitive Architectures and Task-Specific Models

Bradley C. Love
The University of Texas at Austin

The AMBR project has a worthy and lofty goal—to create systems that can successfully simulate human behavior. In service of this goal, AMBR teams created category learning systems that were embedded within a cognitive architecture. To test and more fully develop these architectural proposals, human performance data were collected in a behavioral study that paired structures utilized in Shepard, Hovland, and Jenkins's (1961) classic category learning studies with secondary tasks that evoked air traffic control (ATC) scenarios.

More narrowly defined models developed in the category learning literature have successfully accounted for human performance on Shepard, Hovland, and Jenkins's (1961) six problem types (e.g., Kruschke, 1992; Love & Medin, 1998; Nosofsky, Palmeri, & McKinley, 1994). One interesting question is how these task-specific models compare to the architectural solutions proposed by the AMBR teams. After a brief discussion of the surface commonalities and differences between task-specific and architectural approaches, this chapter considers how models from each approach are developed, utilized, and evaluated. The manner in which models from each class integrate diverse behavioral findings is also discussed. Finally, I suggest when one approach is preferable to the other and argue that task-specific and architectural models are destined to converge.

COMMONALITIES AND DIFFERENCES

Salient commonalities and differences exist between these two classes of models. On the convergence front, many of the AMBR models successfully

LEA—THE TYPE HOUSE—MODELING HUMAN BEHAVIOR WITH INTEGRATED COGNITIVE ARCHITECTURES (GLUCK/PEW)

borrow from the category learning literature. For example, the ACT-R team posits an exemplar-based representation of category information, which is the most popular modeling framework in the category learning literature (e.g., Kruschke, 1992; Medin & Schaffer, 1978; Nosofsky, 1986). The ACT-R model also adopts an exponentially shaped function for relating psychological distance to perceived similarity. This choice, motivated by the work of Roger Shepard (Shepard, 1964, 1987), is often employed in category learning models (e.g., Kruschke, 1992; Love & Medin, 1998; Nosofsky, 1986).

Convergence is not limited to the ACT-R proposal. DCOG, like some current models in the category learning literature (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Erickson & Kruschke, 1998), holds that multiple memory and learning systems underlie category learning performance. Perhaps the best example of transfer from psychological models to AMBR is the EASE team's choice of the RULEX model (Nosofsky et al., 1994) as one of its category learning modules.

The category learning experiment directed by BBN is a force for convergence and divergence. In terms of convergence, BBN's results for the difficulty ordering of the Shepard et al. (1961) Problems I, III, and VI replicated findings used to developed category learning models. This is a testament to the robustness of the original findings given the differences in the BBN experiment (e.g., stimuli, secondary tasks, etc.). Unfortunately, time and resource constraints prevented BBN from including all six problem types from the original Shepard et al. study. This is somewhat problematic because the Type II problem, a highly nonlinear category structure (i.e., XOR with an irrelevant stimulus dimension), is easily learned by humans (more easily than the Type III problem), yet many models predict that Type II should be more difficult than Type III. Type II is unusual because its relative difficulty ordering is not predicted by similarity-based generalization models that do not include rule or attentionally driven components. Corroborating this conclusion, Rhesus monkeys who cannot entertain verbal rules have more difficulty learning Type II than all other problems except Type VI (Smith, Minda, & Washburn, 2004). The omission of Type II denied the AMBR teams the opportunity to directly confront one of the most challenging data points from the original Shepard et al. study.

Of course the AMBR models faced other significant challenges because the BBN data involved tasks other than category learning. Although the psychology literature in general has considered dual-task manipulations that create working memory loads (see Pashler, 1994, for a review), the category learning literature has not exploited such methods (though see Love & Markman, 2003, for an exception). Furthermore, current category learning models do not have a component that directly corresponds to working memory capacity. Perhaps consideration of dual-task data would create an additional point of convergence. For instance, the COGNET team chose to augment their system with a working memory component in light of the dual-task manipulation. Although the BBN data offer a well-thought-out method for incorporating secondary tasks with category learning tasks, unfortunately the secondary task loads were not sufficient to disrupt learning performance, and thus the actual data do not provide strong constraints for extending existing category learning models.

The differences discussed earlier between category learning and AMBR models are symptomatic of a basic difference in approach. The AMBR models propose a complete cognitive architecture. These models are intended to be applicable to all relevant aspects of performance from perception to action. Category learning is simply one application for these architectures. In contrast, category learning models are task specific and focus on explaining human category learning performance. This chapter brings the relationship between these two approaches into focus.

MODEL OBJECTIVES, DEVELOPMENT, AND EVALUATION

In this section, we consider the function of models in the category learning literature and the AMBR program. Model evaluation is also discussed. A later section considers how category learning and AMBR models differ in architectural commitments and in the sense in which they are integrative.

What Are Models For?

Psychologists, including modelers, strive to develop theories that will lead to a mechanistic understanding of cognition. Category learning models can serve multiple functions on the path to this ultimate goal. One function is to offer an instantiation of abstract theories to enable rigorous evaluation. Often the full import of a theory is not understood until it is actualized as a model. For example, until the context model (Medin & Schaffer, 1978), psychologists did not appreciate that an exemplar-based model could account for abstraction results (e.g., Posner & Keele, 1968) that on the surface appear to uniquely support a prototype account of category representation. Once implemented, it became clear that activating numerous studied exemplars stored in memory could lead to strong endorsement of a category prototype (the central tendency of the studied exemplars) even in the absence of studying the prototypical stimulus.

Beyond prediction, category learning models serve as conceptual tools that help researchers understand the determinants of behavior. As such models that are easily understood often prove the most useful. Transpicuous models assist in conducting thought experiments that can yield ideas for actual experiments. The Shepard et al. original studies were partially motivated by similarity-based generalization models. When the Shepard et al. results proved inconsistent with these models, a renewed cycle of model construction and testing began. As can be seen by this example, models in psychology play a critical role in theory development and revision.

In contrast, the goal of AMBR is to develop software systems that can successfully simulate human behavior. Some AMBR teams do not view their systems as psychological theories. For example, the COGNET teams state that their model is an engineering tool for creating practical applications rather than a platform for generating and testing psychological theories. Despite these striking surface differences in the goals of category learning models and the AMBR program, both communities are destined to converge if they remain true to their goals. As the psychological theories underlying category learning models are extended and refined, they will improve in their ability to predict behavior, and related models will provide more complete and accurate simulations of human behavior. Conversely, as simulation test beds become more accurate in making a priori behavioral predictions, they will likely converge to the mechanisms underlying human behavior. In this sense, the two communities are taking two paths to the same destination.

In terms of evaluating models, the different paths these two communities have chosen are reflected in how they evaluate models. Because category learning researchers are primarily concerned with theory development, there is a greater focus on which aspects of the model drive overall model behavior. Although this attribution process has the desirable outcome of building a firm foundation for future progress, it can have the negative consequence of stifling bold proposals and leading to theoretical quagmires. As models converge, a greater number of experiments become necessary to tease apart the alternatives, and the return of investment in terms of predicting behavior diminish rapidly. For example, the ongoing debate between exemplar and prototype theorists focuses on a deep, but narrow, set of findings (from the viewpoint of those outside the field). The debate is difficult to resolve through human experimentation and model simulation because of the subtleness of the predictions the models make (e.g., Nosofsky & Zaki, 2002; Smith, 2002). Although such debates make us keenly aware of the assumptions and capabilities of different formalisms, they are not associated with rapid advances in our ability to simulate human behavior under general conditions.

Of course the downside of the approach adopted by the AMBR participants is that it is impossible to know with certainty which aspects of the model are driving performance and if all assumptions are warranted. The data simply underconstrain the models. Aspects and specific predictions (which can be ascribed to certain mechanisms in the models) are not adequately tested through controlled and systematic human experimentation. In this sense, the AMBR models exist as indivisible units, which can make it difficult to know with confidence what specific lessons are to be gleamed from a simulation.

One proposed solution to this problem is to evaluate cognitive architectures across a wide range of tasks and criteria (Newell, 1980, 1990). By addressing multiple criteria simultaneously, as opposed to specializing in one task, architectural proposals can be viewed as the "decathletes" of cognitive modeling (Vere, 1992). Unfortunately, the criteria proposed for evaluating architectures (e.g., Anderson & Lebiere, 2003; Newell, 1980, 1990) are vague and their application subjective. For example, Anderson and Lebiere (2003), revisiting Newell's earlier work, listed "Integrate diverse knowledge" and "Be realizable within the brain" as 2 of 12 criteria for evaluating architectures. Evaluating such criteria is far from trivial. Furthermore, when the primary research goal is to understand human category learning in particular, knowing that an architecture addresses a number of other tasks in general is only satisfying to the extent that these other successes make one confident that the architecture's account of category learning is sound. Unfortunately, such confluence is likely unwarranted given that numerous solutions to a particular task (such as category learning) can be built within a single architecture. An example of this scenario is discussed later in this chapter.

Model Development and Revision

The differences in approach between task-specific and AMBR models can also be seen in how models are developed. Category learning models tend to undergo a competitive and incremental evolution, whereas AMBR models tend to take bold steps and reuse significant components from existing models. For example, DCOG represents a radically new proposal that has many novel components that are not fully evaluated. Other models, like those from the EASE group, recycle large components from previous efforts or incorporate existing models as submodules. Both of these developmental paths are consistent with the idea that systems in the AMBR community are evaluated at a coarse level (i.e., the architecture or its modules are the unit of evaluation) in comparison with category learning models.

This is not to say that lessons learned over time are not incorporated into an architecture. Architectures undergo evolution and borrow elements from one another (see Anderson & Lebiere, 2003, for a brief history of ACT-R's development). However, simulations of a particular task (such as category learning) do not follow this pattern of development, perhaps reflecting that proponents of cognitive architectures are more concerned with addressing a wide array of tasks than focusing on a particular task.

In contrast, category learning models tend to develop in a more evolutionary manner. For example, the context model is an early example of an exemplar model (Medin & Schaffer, 1978). Nosofsky developed the generalized context model (GCM) by combining the context model with Roger Shepard's work on stimulus generalization (Nosofsky, 1986; Shepard, 1964, 1987). Subsequently, Kruschke combined connectionist learning rules (Rumelhart, Hinton, & Williams, 1986) with the work of Nosofsky to create ALCOVE (Kruschke, 1992). None of these steps in the evolution of the exemplar model involved combining or recycling modules. Each step represented a new model that integrated successful principles from existing models into a coherent whole.

Like evolution, modeling proposals often branch in divergent directions. For example, the exemplar-based random walk (EBRW) model combines GCM with instance-based models of automaticity (Logan, 1988) and work in diffusion decision models (Ratcliff, 1978) to create a new exemplar-based model that can account for reaction time distributions. Branches in category learning research can also crisscross. For example, the SUSTAIN (a clustering model) model's development was influenced by work in exemplar, prototype, clustering, and rule-based approaches (Love & Medin, 1998; Love, Medin, & Gureckis, 2004). The trajectory of model development reflects the theoretically driven nature of work in psychology. Just as there is usually an element of truth in yesterday's dominant theory, there is also some truth in past leading models that is reflected in successive generations of models.

Evaluating Models Through Group and Individual Data

The AMBR project's goal of a priori parameter-free prediction is also embodied to a certain extent in category learning modeling. In many ways, it is as important to know what a model does not predict as it is to know what it does predict. A model that can predict every possible pattern of results through different settings of its parameters is not explanatory. Although psychologists have often erred in the direction of favoring more complex and flexible models, there is a growing awareness of the importance of a priori prediction and considering model complexity. One approach has been to reuse parameters across a number of different simulations (e.g., Love et al., 2004). Another approach has been to consider whether a model can predict a pattern of results not observed (e.g., Love et al., 2004; Markman & Maddox, 2003). A computationally intensive approach is to integrate over the entire parameter space to determine the qualitative predictions of a model (e.g., Johansen & Palmeri, 2002). Psychologists have also developed and applied sophisticated model selection statistics that take into account both the fit and functional complexity of a model (e.g., Pitt, Myung, & Zhang, 2002). Of course the process of deriving predictions from a model and then testing them through experimentation is an exercise in a priori prediction.

As the AMBR project unfolded, AMBR teams became aware of the prevalence and importance of individual differences. Comprehensive simulations of human behavior require accounting for individual differences. Likewise comprehensive theories of categorization must address individual differences. Although the majority of psychological research has relied on group averages, there is a growing appreciation of the importance of characterizing individual differences. Exploring individual differences offers another avenue for discriminating between theories. For example, a theory that holds that people stochastically discover rules and one that holds that people incrementally strengthen connections both predict smooth learning curves for the averaged data, but make different predictions at the level of individual subjects (cf. Estes, 2002).

In the AMBR project, as is often the case in category learning research, individual differences were explored only after their magnitude became apparent. Ideally, the goal of understanding individual differences would be reflected in the basic experimental design. One excellent example of this methodology is the work of Johansen and Palmeri (2002). Johansen and Palmeri trained subjects on a category structure that could be learned by applying an imperfect rule or generalizing across exemplars. They predicted that people's behavior would be rule governed early in learning and exemplar driven late in learning. To evaluate this possibility, they introduced transfer trials without feedback after every few blocks of training. These transfer trials required subjects to apply their category knowledge to new examples. The patterns of transfer performance could easily be characterized as rule or exemplar based or following several other patterns. Subjects were binned by strategy; as predicted, the distribution of transfer pattern shifted toward the exemplar-based pattern from rule-based patterns over the course of training. Rather than relying on average performance, Johansen and Palmeri presented their data as a distribution of outcomes over subjects. Models that account for the average data may not account for the distribution of individual subject performance. For example, a model which predicts that subjects are using a blend of rules and exemplars will not account for a transfer distribution that indicates that subjects are basing their responses on either rules or exemplars.

Another approach to individual differences in the category learning literature has been to study a small number of subjects over many (i.e., hundreds of) trials. This tradition traces its roots back to work in psychophysics, where individual subject analysis (over many trials) is stressed. With so much data per subject, parameters can be fit to individual subjects.

Those using this methodology often employ models in a descriptive manner. For example, General Recognition Theory (GRT), which is a generalization of signal detection theory, can be used to estimate parameters for single subjects (Ashby & Townsend, 1986). These parameters become the data representation for the subject. In other words, it is not the model's job to simulate or fit human data. Rather the model's job is to redescribe data in a more understandable form so that the theoretical significance of the results can be appreciated. In such analyses, different patterns of parameter values are expected for subjects in different learning conditions.

ARCHITECTURE AND INTEGRATION

In this section, we consider architectural differences between category learning and AMBR models, as well as different senses in which models from these two communities are integrative. Many of the differences are reflected in the disparate functions models serve in these two communities (as discussed in the previous section).

Different Senses of Integration

Category learning models trace their roots back to work in mathematical psychology and animal learning (e.g., Spence, 1936). Animal learning studies tend to be fairly simple. After all there are only so many paradigms an animal can be trained in. For example, animals cannot be trained in a simulated ATC task as the BBN subjects were. A lot of early (and recent) work in human category learning, including the Shepard et al. studies, have been influenced by previous work in animal learning. The combined influence of mathematical psychology, with its emphasis on rigor, elegance, and transparency, along with experiments that provide limited ways for subjects to interact with the stimuli, have driven the field toward simple models suited to addressing a narrow range of tasks. The scientific motivations for pursuing such a path are discussed in the previous section.

The end result is models that are simple and built exclusively for category learning (as defined by laboratory paradigms). Category learning models are not architectural proposals like AMBR models are. They tend to test a core set of ideas. For instance, many category learning models are test beds for determining the nature of our category representations. Models have been useful in determining whether categories are represented by exemplars, clusters, prototypes, rules, or combinations thereof. The models tend not to integrate many other mental functions except for those directly relevant to category learning, such as attentional mechanisms. Perception and action are usually removed from the mix by modeling tasks in which these processes are not pivotal in determining category learning performance. By simplifying the data and models, core issues in category learning and representation can be brought into focus.

AMBR models and other proposals for the cognitive architecture have grander ambitions. These models attempt to provide a complete account of cognition. For example, the EASE team's model performs eye movements across the ATC display to actively gather information. These models also have a notion of goal and a control structure to determine what to do next. The DCOG model goes as far as to suggest how personality factors govern strategy shifts. In these ways, AMBR models are much more general and integrative than category learning models.

Still there is more than one way to be integrative. Category learning models tend to be fairly integrative in terms of relating findings from different studies to one another. Shepard et al. is just one study amid hundreds. Category learning models serve to integrate findings into common theoretical frameworks. One way to view the difference in focus between AMBR and category learning models is a choice between depth and breadth.

Integration also occurs at the level of relating different induction tasks to one another. There are more ways to learn about categories than the supervised classification learning procedure used in Shepard et al. and the BBN experiment. For example, people often learn through inference-based learning. In inference learning, the category membership of a stimulus is known, but the value of an unknown stimulus dimension must be predicted (Yamauchi, Love, & Markman, 2002; Yamauchi & Markman, 1998). An example inference trial is, "This is a bird. Is it warm-blooded? Does it eat worms?" Other category learning studies have explored how people learn categories in the absence of supervision and how such unsupervised learning compares with supervised learning (Love, 2002). The results from these studies are not reviewed here, but one important thread that runs through all the results is that human performance is determined by the interaction of induction task (e.g., inference or classification) and learning problem (e.g., Type II or VI) so that studying a single induction task (e.g., classification learning) leads to an incomplete account of human category learning. Category learning models have successfully accounted for these interactions (Love, Markman, & Yamauchi, 2000; Love et al., 2004).

In addition to addressing a wider range of real-world learning situations, these alternative learning tasks have also been useful in model selection. For instance, a number of category learning models with different theoretical commitments can successfully account for human performance on Shepard, Hovland, and Jenkins's (1961) six problem types (e.g., Kruschke, 1992; Love & Medin, 1998; Nosofsky et al., 1994). Consideration of other learning tasks and other measures, such as item recognition following category learning, has proved useful in selecting candidate models (Love et al., 2004; Sakamoto & Love, in press). One would be mistaken to think that the simulation of one data set verifies a model's account of category learning given the varied and challenging nature of the data.

Other modeling efforts have explored integrating learning from examples with prior knowledge (Heit, 2001). Understanding the influence of prior knowledge on learning is critical because prior knowledge can reverse the difficulty ordering of acquiring different category structures (Pazzani, 1991; Wattenmaker, Dewey, Murphy, & Medin, 1986) and can even alter the representations of stimulus items (Wisniewski & Medin, 1994).

Another area of integration for psychological models has been the juncture of brain and behavior. In recent years, there has been a flourish of activity directed at understanding which neural circuits support category learning. Some researchers have proposed models that are closely tied to learning systems in the brain (Ashby et al., 1998). Architectural proposals have also begun to consider cognitive neuroscience data (Anderson et al., 2003).

In summary, integration is occurring in both cognitive architecture models and psychological models of category learning. The different senses of integration often reflect the different goals of these two communities. The fact that category learning researchers, like those working within a cognitive architecture, prize integration might reflect that category learning is a fairly encompassing task with many manifestations.

Who Needs a Cognitive Architecture?

Clearly, an integrative cognitive architecture is required for applications that aim to simulate human behavior from perception to action. Developing cognitive architecture models is a necessity for the AMBR teams. A more interesting question is what the category learning community potentially loses from not working within an architecture. For instance, is working within a cognitive architecture necessary for developing a complete model of category learning? Does working within a cognitive architecture provide useful theoretical constraints on developing a model of category learning?

At this point, the answer to both questions appears to be no. The AMBR teams' results suggest that cognitive architecture plays little role in specifying models of category learning. For instance, the EASE team has constructed two different learning models within their system. It is likely that other category learning models could also be implemented within their system with equivalent results. Similarly, the ACT-R team proposes an exemplar-based category learning system within ACT-R. Also working within the ACT-R architecture, Anderson and Betz (2001) forward a different exemplar-based category learning system, a rule-based category learning system, and a hybrid system. Anderson and Betz's rule-based category learning system is identical to the rule-based category learning system constructed within EASE. In summary, multiple and different category learning systems are proposed within a single architecture, and identical learning systems are proposed in different architectures. At this stage, architecture provides little constraint on the development of category learning models.

Although architecture per se does not constrain the development of category learning models at this juncture, future category learning research may benefit from work in cognitive architecture. Architectures, like ACT-R and EASE, specify how much time cognitive, perceptual, and motor operations require. These parameter values can be used within category learning models to make response time predictions. Also future experimental results may increase the importance of cognitive architecture in developing models of category learning. For instance, dual-task learning data demonstrating interactions between category structure (e.g., Types I, III, VI) and secondary task load may be most readily addressed by working within an existing architecture. As data of this type become available, work in category learning and cognitive architecture will likely merge.

Until such interactions emerge, the best strategy for architecture proponents is to implement existing category learning models within a chosen architecture. In the future, as data become available demonstrating shared resources between category learning systems and other cognitive systems, architectural approaches will serve as more than programming languages with timing constraints. In light of such findings, the divide and conquer approach to understanding category learning and other aspects of cognition and perception might not be as appealing. Those working outside an architecture would need to address concerns that the category learning systems they construct will be inoperable with other cognitive systems. Even at this future juncture, one potentially fruitful approach could be to "grow" existing category learning models to encompass other facilities and cognitive bottlenecks such as a working memory. These facilities could be adjusted to reflect the impact of concurrent tasks on category learning performance. As these interactions with other systems sharing resources become more pervasive, the best approach will likely be to build category learning models within an existing architecture.

CONCLUSIONS

Architectural models of cognition that seek to successfully simulate human behavior and task-specific models that are tools for developing theories of cognition adopt two different paths that will both hopefully lead to successful a priori predictions about human behavior under a variety of conditions. The sense in which these two classes of models are integrative differs greatly. Architectural proposals are integrative by definition. They must simulate behavior from perception to action. Task-specific models, such as models of category learning, are integrative by virtue of placing disparate findings from numerous experiments into a common theoretical framework, considering multiple induction tasks, subsequent item recognition, the effects of prior knowledge on learning, and the relationship between brain and behavior.

Despite adopting different paths, work in one community is relevant to the other. As can be seen in the AMBR proposals, formalisms from the category learning literature are utilized in AMBR models. Similarly, the challenges faced by AMBR models highlight future challenges for category learning modelers. Although cognitive architecture does not strongly constrain theories of category learning at this juncture, it is likely to in the near future. As appropriate dual-task data become available, psychological models of category learning are likely to merge with cognitive architecture proposals.

Finally, applications of cognitive architecture models to realistic behavioral simulations provide valuable information to all of psychology. Theories and experiments created in the laboratory can fail to address variables that contribute to significant variance in real-world settings. To the extent that architectural proposals borrow from theories developed in the laboratory, theorists working in the laboratory gain a valuable check on their efforts. Without such checks, models developed in the laboratory face the danger of not generalizing. Clearly, progress in one community benefits the other.

REFERENCES

Anderson, J. R., & Betz, J. (2001). A hybrid model of categorization. *Psychonomic Bulletin & Review, 8,* 629–647.

Anderson, J. R., & Lebiere, C. (2003). The Newell Test for a theory of cognition. *Behavioral and Brain Sciences, 26,* 587–640.

Anderson, J. R., Qin, Y., Sohn, M.-H., Stenger, V. A., & Carter, C. S. (2003). An information-processing model of the BOLD response in symbol manipulation tasks. *Psychonomic Bulletin and Review, 10,* 241–261.

Ashby, F., Alfonso-Reese, L., Turken, A., & Waldron, E. (1998). A neuropsychological theory of multiple-systems in category learning. *Psychological Review, 105,* 442–481.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154–179.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107–140.

Estes, W. (2002). Traps in the route to models of memory and decision. *Psychonomic Bulletin & Review, 9,* 3–25.

Heit, E. (2001). Background knowledge and models of categorization. In U. Hahn & M. Ramscar (Eds.), *Similarity and categorization* (pp. 155–178). New York: Oxford University Press.

Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology, 45*, 482–553.

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

Logan, G. D. (1988). Towards an instance theory of automatization. *Psychological Review, 95*, 492–527.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review, 9*, 829–835.

Love, B. C., & Markman, A. B. (2003). The non-independence of stimulus properties in human category learning. *Memory & Cognition, 31*, 790–799.

Love, B. C., Markman, A. B., & Yamauchi, T. (2000). Modeling classification and inference learning. *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, 136–141.

Love, B. C., & Medin, D. L. (1998). SUSTAIN: A model of human category learning. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence* (pp. 671–676). Cambridge, MA: MIT Press.

Love, B. C., Medin, D. L., & Gureckis, T. (2004). SUSTAIN: A network model of human category learning. *Psychological Review, 111*, 309–332.

Markman, A. B., & Maddox, W. T. (2003). Classification of exemplars with single and multiple feature manifestations: The effects of relevant dimension variation and category structure. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 29*, 107–117.

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

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4*, 135–183.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

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

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101*(1), 53–79.

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

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244.

Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 416–432.

Pitt, M., Myung, I., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review, 109*, 472–491.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 241–248.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by backpropagating errors. *Nature, 323*, 533–536.

Sakamoto, Y., & Love, B. C. (in press). Schematic influences on category learning and recognition memory. *Journal of Experimental Psychology: General.*

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

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

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs, 75*(13, Whole No. 517).

Smith, J. (2002). Exemplar theory's predicted typicality gradient can be tested and disconfirmed. *Psychological Science, 13*, 437–442.

Smith, J. D., Minda, J. P., & Washburn, D. A. (2004). Category learning in rhesus monkeys: A study of the Shepard, Hovland, and Jenkins tasks. *Journal of Experimental Psychology: General, 133*, 398–414.

Spence, K. W. (1936). the nature of discrimination learning in animals. *Psychological Review, 43*, 427–449.

Vere, S. A. (1992). A cognitive process shell. *Behavioral and Brain Sciences, 15*, 460–461.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology, 18*, 158–194.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science, 18*, 221–281.

Yamauchi, T., Love, B. C., & Markman, A. B. (2002). Learning nonlinearly separable categories by inference and classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 585–593.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language, 39*, 124–149.