UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Feature- vs. Relation-Defined Categories: Probab(alistic)ly Not the Same

Permalink

https://escholarship.org/uc/item/31m404wn

Journal

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

ISSN

1069-7977

Authors

Kittur, Aniket
Hummel, John E,
Holyoak, Keith J.

Publication Date

2004

Peer reviewed

Feature- vs. Relation-Defined Categories: Probab(alistic)ly Not the Same

Aniket Kittur (nkittur@ucla.edu)
John E. Hummel (jhummel@lifesci.ucla.edu)
Keith J. Holyoak (holyoak@lifesci.ucla.edu)
Department of Psychology, 1285 Franz Hall
University of California, Los Angeles
Los Angeles, CA 90095

Abstract

Relational categories underlie many uniquely human cognitive processes including analogy, problem solving, and scientific discovery. Despite their ubiquity and importance, the field of category learning has focused almost exclusively on categories based on features. Classification of feature-based categories is typically modeled by calculating similarity to stored representations, an approach that successfully models the learning of both probabilistic and deterministic category structures. In contrast, we hypothesize that relational category learning is analogous to schema induction, and relies on finding common relational structures. This hypothesis predicts that relational category acquisition should function well for deterministic categories but suffer catastrophically when faced with probabilistic categories, which contain no constant relations. We report support for this prediction, along with evidence that the schemas induced in the deterministic condition drive categorization of novel and even category-ambiguous exemplars.

Relational and Feature-Based Categorization

Most mathematical models of human category learning start with the assumption that people represent categories as lists of features, and assign instances to categories by comparing the features of an instance to the features stored with the mental representation of the category (either a prototype or stored exemplars; e.g., Bruner, Goodnow, & Austin, 1956; Kruschke, 1992; Kruschke & Johansen, 1999; Nosofsky, 1992; Rosch & Mervis, 1975; Shiffrin & Styvers, 1997). Accordingly, most studies of human category learning in the laboratory investigate how people learn categories with exemplars consisting of well-defined (to the experimenter, at least) features.

In the real world, as some researchers have forcefully pointed out (e.g., Barsalou, 1993; Keil, 1989; Murphy & Medin, 1985; Rips, 1989; Ross & Spalding, 1994) categories are less often defined in terms of lists of features than in terms of *relations* between things: either relations between the features or parts of an exemplar (e.g., the legs need to be in a particular kind of relation to the seat in order for an object to serve as a chair), or relations between the exemplar and the user's goals (e.g., any object that affords sitting can, in some circumstances, be considered a chair), or relations between the exemplar and other objects in the world (e.g., what makes an object a "conduit" is a relation between that object and whatever thing flows through it, whether it be water, light, electricity, information, or karma). In spite of their importance in human cognition, comparatively little is known about how people learn relational categories.

Relational category learning is important because relational concepts (i.e., mental representations of relational categories) play an essential role in virtually all aspects of human thinking, including our ability to make and use analogies, problem solving, scientific discovery, and even aspects of perception (see, e.g., Gentner, 1983; Gentner et al., 1997; Green, 2004; Hesse, 1966; Holyoak & Thagard, 1995; Hummel, 2000). The utility of relational representations is that they permit generalization from a small (often as few as one or two) number of examples to a large (potentially infinite) number of new cases (as in the case of inferences generated through the use of analogies, schemas and rules; Gick & Holyoak, 1983; Pirolli & Anderson, 1985; Ross, 1987).

Relational concepts cannot be adequately represented as lists of features (as assumed by most current models of category learning), but instead must be mentally represented as relational structures such as schemas or theories (Gentner, 1983; Holland, Holyoak, Nisbett, & Thagard, 1986; Hummel & Holyoak, 2003; Keil, 1989; Murphy & Medin, 1985). This observation suggests that the operations governing relational schema induction may also underlie the acquisition of relational categories (see, e.g., Kuehne et al., 2000).

At least one theory of schema induction, Hummel and Holyoak's, 2003, LISA model, predicts that a schema induced from two or more examples retains (roughly) the structured intersection of what the examples have in common. For example, consider two analogous stories about love triangles. In the first, Abe loves Betty, but Betty loves Chad, so Abe is jealous of Chad; in the second Alice loves Bill, but Bill loves Cathy, so Alice is jealous of Cathy. Drawing an analogy between these stories maps Abe to Alice, Betty to Bill, and Chad to Cathy (along with the roles of the *loves* and *jealous-of* relations). The schema LISA induces from this analogy retains what the examples have in common, and de-emphasizes the ways in which they differ. For example, since the analogy maps males to females and vice versa, the resulting schema effectively discards the actors' genders, stating (roughly) "person1 loves person2 but person2 loves person3, so person1 is jealous of person3," where persons1…3 are generic people, rather than being specifically males or females (see Hummel & Holyoak, 2003).

Importantly, this intersection discovery process also takes place at the level of whole propositions. For example, if the second story contained a proposition stating that, as a

result of her jealousy, Alice was mean to Cathy, but the first story had no corresponding proposition, then LISA would simply drop this proposition in its entirety from the resulting schema.

If we assume that relational category learning is a process of relational schema induction, then this property of dropping unmapped propositions (i.e., unmapped relations) from the induced schema (i.e., category representation) leads to a counterintuitive prediction: If a relational category has a probabilistic structure, such that every member of the category shares some relations with every other member of the category, but there is no relation that *all* members share, then category learning should fail catastrophically. The reason is that the process of schema induction will drop any relation that is absent from any exemplar from the emerging schema. If every relation is absent from some exemplar (i.e., no relation is present in every exemplar), then schema induction will eventually drop every relation from the schema. By the end, the induced schema will be the empty set.

To clarify, consider a simple relational category with four exemplars, each with three relations chosen from the set r1, r2, r3 and r4 (for our current purposes it does not matter what r1…r4 are, only that they are relations of some sort). Let exemplar 1 (e1) contain the relations r1, r2 and r3. That is, e1 = [r1, r2, r3]. Similarly, let e2 = [r2, r3, r4]; e3 = [r1, r3, r4]; and e4 = [r1, r2, r4]. Note that mapping, for example, e1 to e2 results in a schema (s1,2) that contains relations r2 and r3 (which e1 and e2 share), but lacks r1 (which e1 possesses but e2 does not) and r4 (which e2 possesses but e1 does not): s1,2 = [r2, r3]. Mapping s1,2 onto, say, e3, produces a schema containing only r3, and mapping that schema onto e4 produces a schema containing no relations. The resulting schema is clearly not a useful basis for classifying exemplars as members of the category.

The point is that relational category learning is predicted to be extremely difficult when the categories have a strictly probabilistic structure (i.e., with no relation shared by all exemplars). By contrast, if there is even a single relation that is shared by all exemplars, then category learning should improve dramatically relative to the purely probabilistic case. Categorization performance should also improve dramatically, even with purely probabilistic categories, if the relational structure is replaced with a feature-based structure. Learning of feature-based categories is well known to be robust to probabilistic category structures, a fact that underlies prototype effects (e.g., Posner & Keele, 1968).

In summary, we predict a sharp dissociation between relational and feature-based category learning with respect to their robustness to probabilistic category structures: Both relational and feature-based categories should be learnable when they have a deterministic structure, even if only a single relation or feature reliably predicts category membership; similarly, feature-based categories should be learnable whether they have a deterministic structure or a probabilistic one. By contrast, relational categories should be extremely hard to learn from examples when those examples are presented in a probabilistic structure.

We tested this hypothesized dissociation between relational and feature-based category learning using a 2x2 design, in which relational vs. feature-based categories were crossed with probabilistic vs. deterministic category structures. In order to control all extraneous sources of potential effects, the same basic stimulus set was used in all four conditions; only the assignment of stimuli to categories varied.

Method

Subjects. 33 UCLA undergraduate students participated for course credit.

Instructions. Participants were read a cover story describing a computer manufacturer trying to determine the function of accidentally unlabelled computer chips. Subjects then engaged in a training phase followed by a transfer phase. During both phases, subjects were instructed to indicate the category to which the onscreen stimulus belonged by pressing one of two keys. The categories were labeled "math" chips and "graphics" chips.

Materials. On each trial, the subject saw an exemplar consisting of an octagon and a square, arranged on a fixed background designed to resemble a computer chip (see Figure 1). Each exemplar had both relational properties (e.g., octagon *bigger than* square) and featural properties (e.g., octagon of size 3).
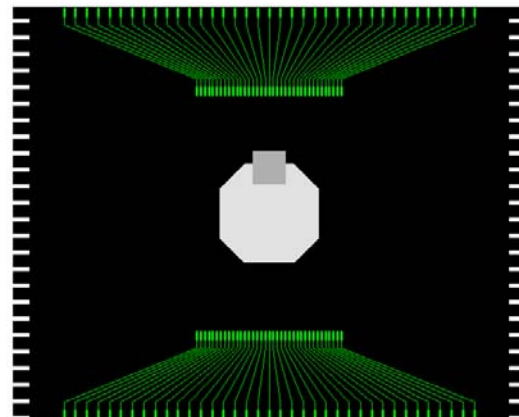
Press f if this chip would go in a graphics computer, j if it would go in a math computer

Figure 1: Example stimulus.

The properties of each exemplar were determined by an identical family resemblance category structure (see Table 1). The prototypes of the two categories were defined as (1,1,1,1) and (0,0,0,0), and distortions were made by chang-

ing the value of one or more dimensions to its opposite[1]. Each column in Table 1 represents an exemplar, and the particular value on each dimension (1 or 0) defines the value of a relation (in the relational condition) or a feature (in the featural condition) for each exemplar. The values for both the relational and featural properties are listed in Table 2. For example, the relational prototype with structure (1,1,1,1) would have an octagon *bigger, darker, above,* and *in front of* a square, while the prototype with structure (0,0,0,0) would be the exact opposite. The properties were set up so that using features could not result in learning to criterion in the relational condition, and using relations in the feature condition would also lead to sub-criterion responding.[2] Stimulus generation and display as well as response collection were done with a program written in Matlab.

Design. The experiment used a 2 (category structure: probabilistic vs. deterministic) X 2 (relevant property: features vs. relations) between-subjects design. The only difference between the conditions in terms of the stimuli used was that, in the deterministic condition, a single distorted exemplar from each category was not presented during training, so that one dimension was constant for all exemplars of a category. The choice of which dimension was held constant was counterbalanced across subjects.

Table 1: Family resemblance category structure. Each column represents an exemplar, and each row a dimension.

Category A				Ambiguous						Category B				
1	1	1	1	0	1	1	1	0	0	0	0	0	0	1	0
1	1	1	0	1	1	0	0	1	0	1	0	0	1	0	0
1	1	0	1	1	0	1	0	1	1	0	0	1	0	0	0
1	0	1	1	1	0	0	1	0	1	1	1	0	0	0	0

Procedure. During the training phase subjects classified only distorted exemplars of each category (depicted in the light gray columns of Table 2). All distortions for each category were shown in random order exactly once per block. Responses were followed by accuracy feedback, during which the exemplar remained on the screen. Subjects pressed the space bar to proceed to the next trial. The training phase continued until the subject responded cor-

rectly on at least seven out of eight trials for two consecutive blocks[3], or until they had finished 75 blocks (600 trials) without reaching this criterion.

Following the training phase, subjects were informed that they would be tested on chips for which feedback could not be given. During this transfer phase subjects classified all 16 possible exemplars, including the prototypes and ambiguous exemplars. Subjects completed five blocks, with each block showing all 16 exemplars in random order exactly once.

After the transfer phase, each subjects completed a questionnaire in which they were asked to write down the criteria they used to categorize the exemplars.

Table 2: Category definitions.

Relational categories

Exemplar	Relation	Exemplar	Relation
1	Bigger	0	Smaller
1	Darker	0	Lighter
1	Above	0	Below
1	In Front	0	Behind

Feature-based categories

Exemplar	Feature	Exemplar	Feature
1	O *size 3*	0	O *size 2*
1	O *shade 4*	0	O *shade 3*
1	S *size 1*	0	S *size 2*
1	S *shade 1*	0	S *shade 2*

Note: Prototype exemplars are shown with their defining properties on each dimension. In the relational condition, each dimension defines how the octagon (O) in the stimulus relates to the square (S). For the featural condition each dimension defines specific feature values.

Results

Training. Only 5 of the 7 subjects (71%) in the relational probabilistic (RP) condition learned to criterion within 600 trials. 25/25 subjects (100%) in the other conditions learned to criterion within the 600 trial limit. In the analyses that follow, the 2 subjects in the RP condition who never learned to criterion are treated as though they reached criterion on trial 601. Given that our hypothesis predicts that learning in the RP condition will be harder (and therefore take longer) than learning in the other conditions, this assumption is extremely conservative.

The mean number of trials to criterion is shown in Figure 2 for each condition. Subjects in the RP condition took more trials to reach criterion than those in the FD (featural deterministic), FP (featural probabilistic), and RD (rela-

---

[1] Note that the exemplars marked "Ambiguous" are equal distance between the two prototypes, having exactly two values different from each.

[2] In the relational condition, stimuli from different categories could have the same features (and stimuli from the same category could have different features) as long as the specified relations held; features were thus non-diagnostic. For the featural condition, the relations *in front* and *above* had no relevance to the category structure, and were pseudo-randomized. The relations *bigger* and *darker* were made irrelevant by choosing values such that the octagon was never smaller or lighter than the square (though it could be the same size, since only three sizes were used). See the Discussion section for further analysis of feature and relation values.

[3] This criterion level (87.5%) was selected because strategies involving tracking only one or two relations in the probabilistic condition would not meet the criterion level (both would result in 75% correct responding).

tional deterministic) conditions. A planned contrast comparing the RP condition to the other three revealed that this difference was statistically reliable ($p < 0.01$). There was also a significant main effect of category type (relational vs. featural, $F(1,33)=4.64$, $p < 0.05$). The main effect of category structure (deterministic vs. probabilistic) and the interaction were both marginally reliable ($0.05 < p < 0.15$).
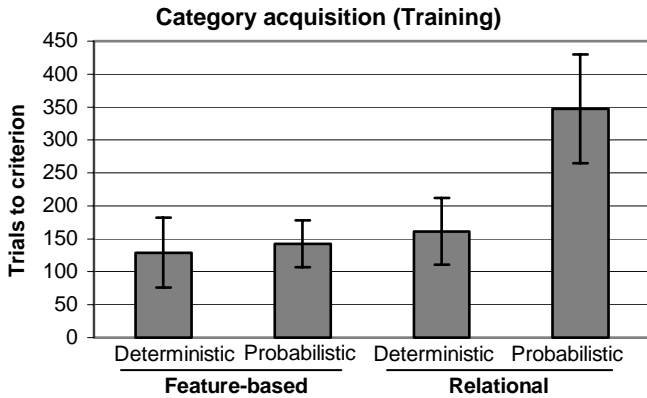
Category acquisition (Training)

Figure 2: Average number of trials required by subjects in each condition to reach criterion during training.

Transfer. The key prediction for the transfer phase was that subjects in the deterministic condition would categorize exemplars based on whatever dimension was held constant during training. This prediction applied especially to the relational condition, which could not rely on holistic processing. To test this hypothesis we analyzed classification of the ambiguous exemplars, which were equidistant between the two prototypes. Subjects who used all category dimensions equally should be unsystematic in their classification of these ambiguous exemplars. By contrast, subjects who attend to a single dimension should classify ambiguous exemplars according to that dimension only (as detailed in Table 3). If a classification response for a dimension that was held constant during training matched the response pattern in Table 3, then +1 was scored for that response; classifications that did not match Table 3 response patterns were scored as -1. Under this scoring system, consistently responding to ambiguous exemplars in the direction predicted by the constant training dimension results in a positive score; consistently responding in the direction opposite the constant dimension results in a negative score; and unsystematic responding results in a score near zero.

Classification of ambiguous exemplars in accordance with the dimension that was constant during training was significantly above chance ($p < 0.01$). Breakdown into featural and relational conditions showed a non-significant trend for the relational condition to evoke classifications based on the constant training dimension more often than the featural condition (Figure 3).

Discussion

The results showed that acquisition of relational probabilis-

tic categories takes significantly longer than acquisition of deterministic relational categories, or featural categories of any kind (probabilistic or deterministic). Importantly, the ease of acquisition in the deterministic relational condition shows that this effect is not due strictly to the relational nature of the task. Instead, the catastrophic failure represents an interaction between the relational nature of the stimuli and the probabilistic structure of the categories. This interaction is consistent with the hypothesis that relational category learning is a process akin to relational schema induction by intersection discovery: When the intersection is the empty set (as it is in the probabilistic condition but not the deterministic condition), relational category learning suffers markedly. By contrast, feature-based category learning is much more robust to the probabilistic category structure, presumably because feature-based category learning is not a process of relational schema induction; instead, as predicted by models of feature-based category learning, it may be that learning feature-based categories can be accomplished simply by cataloging and matching features.

Table 3: Classification of exemplars based on single dimensions

Exemplar	Dim 1	Dim 2	Dim 3	Dim 4
1 1 0 0	A	A	B	B
1 0 1 0	A	B	A	B
1 0 0 1	A	B	B	A
0 0 1 1	B	B	A	A
0 1 1 0	B	A	A	B
0 1 0 1	B	A	B	A

Note: Table entries indicate how each exemplar would be categorized by a subject who attended only to a single dimension (columns in the table). For example, a subject who attended only to the first dimension (Dim 1) would classify the first, second and third exemplars as As since their values on that dimension are all one, and B for the fourth, fifth and sixth exemplars (the values of which are zero on that dimension).

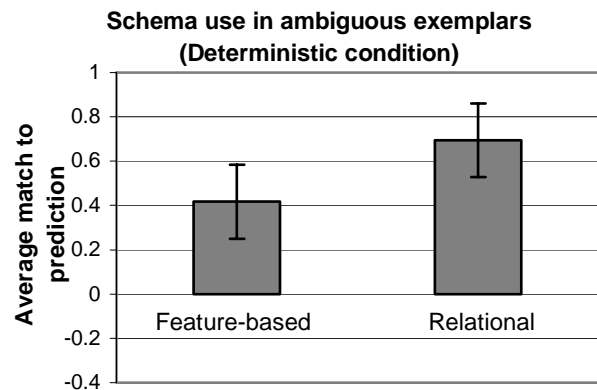**Schema use in ambiguous exemplars (Deterministic condition)**

Figure 3: Average match to predicted classification pattern. Positive values indicate schema-based classification; zero corresponds to unsystematic responding.

The significant match of subjects' responses with single-dimension classification predictions in the deterministic condition also shows that subjects do preferentially use dimensions that are constant during training to classify novel and even category-ambiguous exemplars.

Is it possible to explain these results in other ways? One possibility is that rather than attending to the relations, subjects in the relational conditions may instead be tracking the feature values of certain dimensions. On this hypothesis, there is no schema induction going on in any condition; instead, responding is based on the values of particular features. This account obviates the need for a separate process to explain relational categorization.

However, analysis reveals that subjects tracking a feature of a single dimension would only classify 5/6 correct in the deterministic condition, and 2/3 correct in the probabilistic condition. Both of these values are below the 7/8 criterion, suggesting that subjects who reach criterion were not doing so by tracking a feature of a single dimension.

The possibility remains that subjects were tracking the values of multiple features or dimensions, although these seem unlikely strategies for a number of reasons. First, even when tracking the values of a single feature the subject must hold in mind three or four values and their associations with each category (for example, each size of the octagon and its corresponding category). Each additional feature or dimension would double the number of values necessary to track. This strategy does not seem plausible given the well-known limits on the capacity of working memory. Also, subjects' responses to the debriefing questionnaires in the relational conditions did not suggest such strategies were being used; instead, they generally reported the use of one or two relations as diagnostic, often along with some exception exemplars. Thus it seems more likely that subjects were indeed attending to the relations between the components of each stimulus rather than tracking feature values of those components.

Another hypothesis to explain the difference between the featural and relational conditions is that subjects were memorizing all the possible exemplars, and a difference in the number of distinct exemplars made the relational condition harder. This view must also hold that the deterministic conditions do not rely on such memorization, in order to explain the results. This view has some merit, though two factors reduce its likelihood.

First, the total number of distinct exemplars in the featural condition is not very different than the relational condition: 128 vs. 144. While this is a difference between the categories, it is difficult to ascribe the extra difficulty of the relational probabilistic condition to its having an extra 16 exemplars.

Second, although debriefing questionnaires did indicate subjects were memorizing some of the exemplars in the relational probabilistic condition, these were of very limited number (usually ~2 exemplars) and were memorized as exceptions to a more general classification rule. Thus while it remains a theoretically possible explanation, the "number of exemplars" view is not very compelling.

Preliminary analysis of the debriefing forms for subjects who learned to criterion in the relational probabilistic task suggest that what is learned is often a classification rule (such as might result from a schema induction process) along with a few memorized exceptions. Subjects often mentioned one or two relations in their classification rules; only one subject reported attending to all four dimensions; unsurprisingly, this subject was the only one who deduced the formal category structure (that is, that three out of four of the dimensions are necessary for category membership).

Subjects in the featural probabilistic condition also failed to show feature-tracking strategies in their debriefing questionnaires. Instead, their responses often showed a reliance on emergent properties of the stimuli such as high vs. low contrast. Questionnaires from the deterministic conditions tended to show a focus on the dimension that was constant during training, and mentioned particular features in the featural condition and relations in the relational condition. Thus subjects' explicit responses often fit well with the predictions about processing.

Why should relational categories rely on schema induction processes? One possibility is that feature-based categories tend to give rise to emergent properties, since their features are fixed at some value or limited range of values. However, it is much more difficult for emergent properties to arise in relational categories, because they can take on many different and overlapping values. The lack of emergent properties may explain the dependence of relational categorization on deterministic dimensions. This view is consistent with subjects' self-reported strategies.

Another interpretation of the present results is that people are either unwilling or unable to perceive, predicate and categorize patterns across four relations. This deficit may be due to working memory constraints, strategy choice, or low prior experience with similar situations. Studies of working memory suggest that we can hold about four chunks or role bindings in working memory (e.g., Halford, Wilson, & Phillips, 1998); holding four two-place relations exceeds this limit. It may be that people can learn some probabilistic relational categories with experience by recoding relations as features; others may be learned by dividing the probabilistic category into deterministic subcategories, or by perceiving a unifying causal relation for the entire category.

In conclusion, the results of the present experiment suggest that relational category learning relies heavily on finding common relations across exemplars. In contrast, feature-based category learning appears to function robustly whether common elements are present or not. These findings are consistent with the view that relational category learning is a kind of relational schema induction that depends on intersection discovery. Performance on the transfer trials also support this conclusion in that dimensions that were constant during training dominated classification of novel exemplars, even those that were category-ambiguous. Such findings suggest that relational category

learning may be fundamentally different from feature-based category learning, though more work is needed to distinguish these modes of category learning.

References

Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. F. Collins & S. E. Gathercole & M. A. Conway & P. E. Morris (Eds.), *Theories of memory*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. Oxford, England: John Wiley and Sons.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Gentner, D., Brem, S., Ferguson, R. W., Markman, A. B., Levidow, B. B., Wolff, P., & Forbus, K. D. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. *Journal of the Learning Sciences. Special Issue: Conceptual change, 6*, 3-40.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Green, C. B., & Hummel, J. E. (2004). Relational perception and cognition: Implications for cognitive architecture and the perceptual-cognitive interface. In B. H. Ross (Ed.), *The psychology of learning and motivation*. San Diego: Academic Press.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral & Brain Sciences, 21*, 803-864.

Hesse, M. B. (1966). *Models and analogies in science*. Notre Dame, IN: University of Notre Dame Press.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, MA, US: The MIT Press.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: The MIT Press.

Hummel, J. E. (2000). Where view-based theories break down: The role of structure in human shape perception. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220-264.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA, US: The MIT Press.

Kuehne, S., Forbus, K., Gentner, D. and Quinn, B. (2000). SEQL: Category learning as progressive abstraction using structure mapping. *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*.

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

Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*, 1083-1119.

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

Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. F. Healy & S. M. Kosslyn (Eds.), *Essays in honor of William K. Estes, Vol. 1: From learning theory to connectionist theory; Vol. 2: From learning processes to cognitive processes*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Pirolli, P. L., & Anderson, J. R. (1985). The role of practice in fact retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 11*, 136-153.

Posner, M. I., & Keele, S. W. (1968). On the Genesis of Abstract Ideas. *Journal of Experimental Psychology, 7*, pp. 353-363.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. O. Vosniadou, Andrew (Ed.), *Similarity and analogical reasoning*. New York, NY, US: Cambridge University Press.

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

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13*, 629-639.

Ross, B. H., & Spalding, T. L. (1994). Concepts and categories. In R. J. Sternberg (Ed.), *Thinking and problem solving. Handbook of perception and cognition (2nd ed.)*. San Diego, CA: Academic Press, Inc.

Shiffrin, R. M., & Styvers, M. (1997). A model for recognition memory: REM--retrieving effectively from memory. *Psychonomic Bulletin & Review*, 145-166.