Relation-Based Categories are Easier to Learn than Feature-Based Categories

Marc T. Tomlinson (mtomlinson@love.psy.utexas.edu)

Bradley C. Love (Love@love.psy.utexas.edu)
University of Texas At Austin
1 University Station, A8000
Austin, TX 78712 USA

Abstract

Relational reasoning is often viewed as the pinnacle of human intelligence. Accordingly, one common viewpoint is that learning categories defined by relational regularities is more difficult than learning categories defined by featural regularities. This view is supported by developmental trends in learning. Studies comparing featural and relational category learning in adults also find a feature advantage, but these studies do not ground featural and relational information in a common perceptual substrate. The current study offers an appropriate comparison between feature- and relation-based category learning. Contrary to previous studies, we show how relational learning can be easier. The advantage is attributable to the flexibility of online relational comparisons between a stimulus and a memory representation of a category. Alternative explanations based on difficulties in processing absolute vs. relative stimulus information are ruled out.

Keywords: Analogy; Category Learning; Relations

Introduction

The ability to grasp complex relations is a hallmark of human intelligence (Thompson & Oden, 2000). Evaluative tests such as the SAT, GRE, or Raven's Progressive Matrices stress the importance of relational thinking. Most research comparing relational and featural performance supports the view that relational processing is a more advanced competency. Children learn concepts defined by features earlier than those defined by relations (Gentner, 1978). Experts differ from novices by organizing knowledge within their domain along relational lines (Chi, Feltovich, & Glaser, 1981).

Despite this work comparing featural and relational processing, studies comparing the ability to learn novel concepts defined by features or relations under comparable conditions is absent. It is possible that when numerous factors are balanced, relational learning could prove easier than featural learning. For example, feature- and relation-based tasks often differ in their information demands (Kittur, Hummel, & Holyoak, 2004; Waltz, Lau, Grewal, & Holyoak, 2000).

Furthermore, no study comparing feature- and relation-based learning has used stimuli that rely on the same perceptual substrate. In all cases, the perceptual features carrying relational and featural information are different. In other domains, such as same/different discrimination (Love, Rouder, & Wisniewski, 1999) and change detection (Kroger, Holyoak, & Hummel, 2004), this concordance has been achieved, but never in a learning study.

Although many would agree that the developmental trajectory follows a progression from appreciating feature matches to grasping more complex relational matches (Gentner, 1988; Gentner & Ratterman, 1991; Richland, Morrison, & Holyoak, 2006), how far does this shift go? Perhaps when confounding factors are equated, adults will more readily master relation-based concepts than feature-based ones.

The current study finds that certain relation-based categories are more readily acquired and explores the basis for this advantage. All four experiments made use of stimuli in which featural and relational information supervenes on the same perceptual information. Experiment 1 establishes that certain relation-based categories are more readily acquired than comparable feature-based categories. Experiment 2 rules out that this advantage arises from a greater difficulty in encoding absolute stimulus information compared to relative stimulus information. Experiment 3 suggests that relation-based categories are acquired faster when the flexibility afforded by aligning relations during the comparison process is useful for completing the task. Experiment 4 further tests this hypothesis using a category structure that should not benefit from the power of analogical alignment. In this case, feature- and relation-based categories are acquired at the same rate.

Stimulus Design

The simple scenes used for the experiments were designed specifically to address the issue of whether it is harder for adults to learn categories defined by the relationships between objects or to learn categories defined by the general features of those same objects. To avoid confounding factors and provide an informative comparison, the relations and features were both defined over the same perceptual factors. To correctly classify a scene based on its features, participants had to observe the brightness and size of each object. To correctly classify the scene based on the relations within it, the participants had to make a relational judgement about the same perceptual factors of brightness and size.

Each scene consisted of two circles appearing side-by-side. Across trials, these two circles varied in their size (small, medium, large) and brightness (light, moderate, dark). These circles were combined to give two overall relation attributes (which side was bigger and which was brighter) and two overall feature attributes (size and brightness). The medium and moderate values were always manifested once in a scene (See Figure 1).

Experiment 1

The purpose of Experiment 1 was to investigate the relative difficulty of learning to classify a scene based on either the

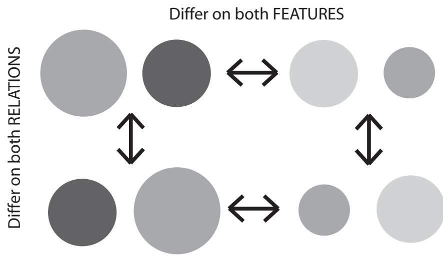

Figure 1: Four example stimuli and their differences. The circles vary on four different attributes: two features and two relations. The features are overall size and overall brightness (defined over both circles). The relations are which circle (by left/right spatial position) is bigger and which circle is brighter.

relations between objects in a scene or the overall features of the scene.

Methods

Participants Fifty-two undergraduate students from the University of Texas at Austin participated for course credit. Participants were randomly assigned to the feature relevant (n=27) or relation relevant (n=25) condition.

Stimuli and Category Structure Each stimulus was composed of two circles displayed side by side on a black screen. The circles were bounded by constant sized green boxes of a moderate brightness. For each individual exemplar, the circles were paired so as to give four overall binary attributes: which side was brighter, which side was bigger, overall size (both circles), and overall brightness (both circles). This resulted in 16 different possible stimuli that participants learned to classify.

The category structure learned by all of the participants followed a binary XOR structure over the two relevant attributes. The binary XOR is a non-linear classification rule that requires attention to both of the relevant attributes (see Table 1). In the relation condition, the two relation-based attributes were relevant and the features were irrelevant. The opposite was true for the feature condition.

Procedure Each participant was first presented with a screen of detailed instructions informing them that they were going to learn to categorize pairs of circles into two categories, A and B. Participants were instructed that each stimulus varied along four binary-valued attributes (overall brightness, overall size, which circle was brighter, which circle was bigger) and told to look for a rule involving these attributes. For each participant, the labels *A* and *B* were randomly assigned to correspond to the actual category A or B.

On each learning trial, two circles were presented in the center of the computer screen. The stimulus was accompanied by the text prompt "Category A or B?" at the top of the screen. Participants freely responded with an *A* or *B* key press. Immediately after responding, participants received either a brief low (wrong) or high (right) pitched auditory tone concurrent with text displayed on the bottom of the screen containing "WRONG" or "RIGHT" and the correct category label for the stimulus. The correct category label and the stimulus were presented for 1250 ms before being replaced by a blank screen for 500 ms. Then the next trial started. The trials were blocked in groups of 16 which consisted of a random ordering of the 16 stimuli. Participants were not made aware of the transition from one block to the next. Category training terminated when participants reached a learning criterion of correctly classifying 12 stimuli in a row or participants completed 18 blocks (288 trials) of learning.

Table 1: Category structure

Attr. 1	Attr. 2	Attr. 3	Attr. 4	XOR	Four-Category
0	0	0 *or* 1	0 *or* 1	*A*	*A*
0	1	0 *or* 1	0 *or* 1	*B*	*B*
1	0	0 *or* 1	0 *or* 1	*B*	*C*
1	1	0 *or* 1	0 *or* 1	*A*	*D*

Note: For example, with features relevant in the four-category structure: big and bright would be A; big and dark, B; small and bright, C; and small and dark, D.

Results and Discussion

An accuracy score was calculated for each participant over the 288 trials. If a participant reached the learning criterion before 288 trials, they were assumed to have been correct on the remaining trials. The results are summarized in Table 3.

Significantly more participants reached criterion in the relation condition, 14 of 25 participants versus 4 of 27 in the feature condition $\chi^2(1, N = 52) = 8.00, p < .01$. The mean accuracy for the two learning conditions was .73 (relation) and .54 (feature). Only in the relation condition did participants perform significantly above chance, $t(24) = 5.38,\ p < .001$, $t(26) = 1.70,\ p \approx .10$ (respectively). A t-test revealed that the difference in the mean accuracy of participants in the two conditions was significant, $t(50) = 5.19,\ p < .001,\ \eta_p^2 = .25$.

The results of Experiment 1 demonstrate that categories defined by the relations are easier to learn than categories defined by the features of a scene.

Experiment 2

Experiment 1 demonstrated that relation-based categories are easier to learn than feature-based categories. One possible explanation for this surprising result is that people find it easier to process relative rather than absolute stimulus information (cf. Garner, 1954; Huttenlocher, Duffy, & Levine, 2002; Stewart, Brown, & Chater, 2005). Learners in the feature condition needed to process absolute stimulus information to determine whether each stimulus was dark or light and small or large, whereas learners in the relation condition could simply compare the two circles forming a stimulus.

Experiment 2 tests the hypothesis that the difficulty in determining absolute feature values was responsible for the poor performance of participants in the feature condition of Experiment 1. All of the participants were pretrained to identify the feature values of the circles. Participants were then transferred to either the relation- or feature-based classification task used in Experiment 1. If the relational learning advantage observed in Experiment 1 arose from a difficulty in processing absolute stimulus information, Experiment 2's pretraining on absolute values should neutralize the relational learning advantage.

Methods

Participants Twenty-nine undergraduate students from the University of Texas at Austin participated for course credit. Participants were randomly assigned to the feature relevant (n=15) or relation relevant (n=14) condition. All participants received pretraining.

Stimuli and Category Structure The stimuli and category structure were the same as those used in Experiment 1.

Procedure The procedure was similar to Experiment 1. The one difference was that all of the participants were first trained on identifying the feature values. A single circle was randomly chosen from the nine possible circles (with replacement) and shown to the participant on a computer screen. The circle was either small, medium or large, and of low, moderate or high brightness. Text would then appear next to the circle asking the level of brightness of the circle (1-3). Participants responded by pressing *1*, *2*, or *3*. They then heard either a low (wrong) or high (right) tone accompanied by text giving the actual level. The participants were then asked how bright the circle was, they responded and were then given feedback in the same manner. They were shown these circles until they were able to correctly identify both the size and brightness of six circles in a row. Participants were then run in either the relation or feature learning condition as in Experiment 1.

Results and Discussion

Participants needed an average of 14.79 (SE=1.21) identification trials (two judgments on each trial) to reach the criterion of correctly identifying both the size and brightness attributes of six circles in a row. There was no significant difference in the identification performance of participants later assigned to the relation or feature condition, with means of 14.00 (SE = 1.90), and 15.33 (SE = 1.55) respectively, $t < 1$.

The following results are summarized in Table 3. During the learning phase, significantly more participants in the relation condition reached criterion, 9 of the 14 versus 2 of 15 in the feature condition, $\chi^2(1, N = 29) = 5.97, p < .05$. Overall, participants in the relation condition achieved a mean accuracy of .79, which differed significantly from chance, $t(13) = 4.95, p < .001$. For the feature condition, the mean accuracy of .52 did not differ significantly from chance, $t < 1$. A t-test showed that the participants performed significantly worse in the feature learning condition than in the relation learning condition, $t(27) = 4.42, p < .001$, $\eta_p^2 = .45$.

Experiment 2's results suggest that the reason for poor performance in the feature condition is not caused by difficulty in identifying absolute stimulus values. Participants readily mastered the value mappings during pretraining. Nevertheless, participants learned categories defined by relations faster than those defined by features (as In Experiment 1).

Experiment 3

The results from Experiment 2 suggest that difficulty in processing absolute stimulus values is not the cause of the relational learning advantage. Another alternative explanation is that the flexibility provided by online alignment processes provides greater benefit to relational rather than feature learners. The process of structural or analogical alignment involves comparing two analogs that contain relations (e.g., bigger, brighter) to determine appropriate correspondences between the analogs (Gentner, 1983). Numerous computational models have been designed that implement this alignment process (e.g., Falkenhainer, Forbus, & Gentner, 1989; Hummel & Holyoak, 1997; Larkey & Love, 2003).

The Building Relations through Instance Driven Gradient Error Shifting (BRIDGES) model (Tomlinson & Love, 2006) is particularly relevant in terms of motivating predictions for Experiments 3 and 4. BRIDGES is an exemplar-based connectionist model of human category learning model that can appreciate analogical relationships between stimuli and stored predicate representations of exemplars. Rather than relying on pre-established correspondences between scenes (e.g., left circle corresponds to left circle) as in standard category learning models (e.g., Kruschke, 1992), BRIDGES aligns the current stimulus with each category exemplar stored in memory in order to minimize attention-weighted mismatch (and thus maximize similarity) for both feature and relational correspondences.

This online alignment process can yield non-metrical similarity spaces in which stimuli that differ on two relations can be more similar than stimuli that differ on one relation. For example, BRIDGES predicts that comparing the top-left and bottom-left stimuli in Figure 1 results in an alignment that puts the left circle from one scene into correspondence with the right circle of the other scene. This alignment creates matches for all features and relations (overriding the default preference to preserve spatial correspondences) and thus results in a high similarity match. Importantly for the current learning studies, items that mismatch on both relations are always members of the same category for relational learners (see Table 1). To the extent that high within- and low between-category similarity promotes category acquisition (Rosch & Mervis, 1975), a flexible alignment process predicts the relation advantage observed in Experiments 1 and 2.

Experiment 3 tests the flexible alignment hypothesis. Participants rated the similarity of all pairs of stimuli. If the preferred alignment changes depending on the stimuli being compared, the resultant similarities should be non-linear with respect to the number of relational differences between the stimuli. In fact, stimulus pairs mismatching on both relations might be judged to be more similar than stimulus pairs mismatching on one relation.

Methods

Participants Twenty-two undergraduate students from the University of Texas at Austin participated for course credit. All of the participants were run in the same condition (N=22)

Stimuli The stimuli were the same as those used in Experiment 1.

Procedure Participants were instructed to rate the similarity of two presented stimuli on a scale from 1-9. As in Experiment 1, participants were instructed that each stimulus varied along four binary-valued attributes (overall brightness, overall size, which circle was brighter, which circle was bigger). On each trial, two stimuli were simultaneously presented on screen with text designating pair 1 and pair 2, as well as text asking for their similarity on a scale of 1-9. One pair was on the top of the screen and the other on the bottom. The pairs were separated by a line. Participants responded by pressing key 1 through 9. Following the participant's response, the screen blanked for 500 ms and the next trial began. Each participant rated 136 pairs of stimuli [(16 * 15)/ 2 + 16]: each stimulus paired with every other stimulus, plus each stimulus paired with itself. The overall order of the trials and the assignment of pairs to the top or bottom of the screen were randomized.

Results and Discussion For the purposes of analysis the similarity ratings were grouped according to how many features or relations were different within the comparison. Figure 2 illustrates the nine means resulting from this aggregation. A 3 (0, 1, or 2 relation changes) X 3 (0, 1, or 2 feature changes) within-participant ANOVA revealed a main effect of both the number of different relations, $F(2,42) = 33.68,\ p < .001,\ \eta_p^2 = .62$, and the number of different features, $F(2,42) = 204.81,\ p < .001,\ \eta_p^2 = .91$, as well as a significant interaction between the number of feature changes and relational changes, $F(4,84) = 36.47,\ p < .001,\ \eta_p^2 = .63$.

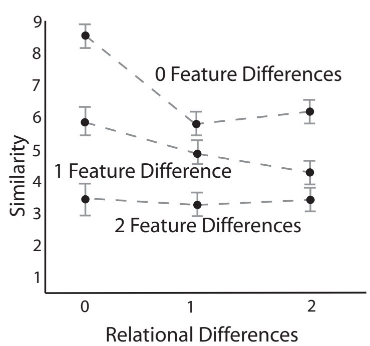

Figure 2: Mean similarity as a function of number of different features and relations. Error bars represent 95% confidence intervals.

The above interaction is indicative of a non-metrical similarity space arising from an alignment process. To test the predictions of the alignment account more precisely, a 2 (relation or feature) X 2 (one or two changes) ANOVA was performed to compare the effects of mismatching on one or both relations (with both features matching) with the effects of mismatching on one or both features (with both relations matching). The interaction predicted by the alignment account is shown in Figure 3. The ANOVA revealed a significant main effect for the type of change, $F(1,21) = 26.32, p < .001,\ \eta_p^2 = .56$, and the number of changes, $F(1,21) = 23.85, p < .001,\ \eta_p^2 = .53$. There was also a significant interaction between the type of change and the number of changes, $F(1,21) = 96.05, p < .001,\ \eta_p^2 = .82$. Planned t-tests showed that similarity increases significantly when both relations change, compared to only one, $t(21) = 2.13, p < .05,\ \eta_p^2 = .02$. In contrast, similarity decreases significantly when both features change, $t(21) = 13.68, p < .001,\ \eta_p^2 = .68$.

The similarity results provide support for the flexible alignment hypothesis. Differences between the features of two stimuli result in a linear decrease in similarity, whereas changes in the relations result in a non-metrical similarity space.

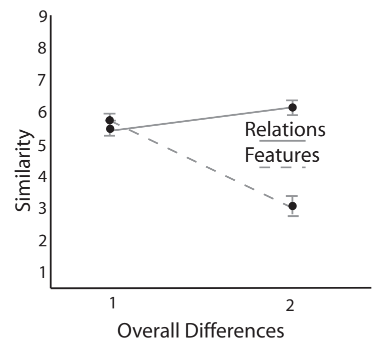

Figure 3: An interaction plot of the mean similarity of the four sub-points of particular interest. Error bars represent 95% confidence intervals.

Experiment 4

Participants' similarity ratings in Experiment 3 were indicative of a non-metrical similarity space. As predicted by an online alignment account, under certain conditions, increasing relational mismatch increased rated similarity. This surprising effect was predicted to arise from the ability to select correspondences or mappings between compared stimuli in order to maximize similarity.

The flexibility afforded by the proposed alignment process would be particularly beneficial to the relational learners in Experiments 1 and 2. For relational learners of the XOR category structures, items that differ on both relations are members of the same category. To the extent that high within and low between category similarity promotes acquisition (Rosch & Mervis, 1975), a flexible alignment process predicts the relation advantage observed in Experiments 1 and 2. The flexible alignment process is not predicted to benefit featural learners – an added feature difference should always decrease similarity. The BRIDGES model provides a computational instantiation of this account that calculates the similarity within- and between-categories in terms of alignments to stored category exemplars.

Experiment 4 further tests the alignment account by training participants on a category structure in which flexible alignment is not advantageous. Differences in performance between featural and relational learners are predicted to compress under these conditions. The category structure used in Experiment 4 is the four-category structure specified in

Table 1. Unlike the XOR category structure, in the four-category structure, items that differ on both relevant attributes are members of different categories.

To the extent that relational learners engage in flexible alignment processes (i.e., finding correspondences that run counter to the spatial positions of the circles), they should show a characteristic error pattern during learning. In particular, participants' responses in the relation condition should reflect the similarity results from Experiment 3 by showing more errors between the categories that are two relevant changes away than those only one relevant change away.

Unfortunately, cross-experimental comparisons of overall performance levels are not warranted because of differences in the category structures used (e.g., nonlinear vs. linear; two vs. four category choices). The key predictions for Experiment 4 are that the difficulty of featural and relational learning should converge, and that relational learners should show evidence of alignment by making more generalization errors to the opposite category (two relevant changes away), than to an adjacent category (one relevant change away).

Methods

Participants Fifty-three undergraduate students from the University of Texas at Austin participated for course credit. Participants were randomly assigned to the feature relevant (n=27) or relation relevant (n=26) condition.

Stimuli and Category Structure The stimuli were the same as those used in Experiment 1. In contrast to Experiment 1, participants learned to classify each stimulus as a member of one of four different categories. The categories were the four unique combinations of the two values of the two relation attributes in the relation condition and the two feature attributes in the feature condition (see Table 1).

Procedure The procedure was identical to that used in Experiment 1, except participants had to learn to classify the circles as belonging to one of four categories: A, B, C, or D, by pressing the corresponding key. There were no other differences in the instructions. As in the previous experiments, the labels were randomly assigned for each participant.

Results and Discussion

The number of people reaching criterion was 22 of 26 in the relation group and 22 of 27 in the feature group. The mean accuracy for the relation and feature conditions were .78 and .74, respectively. This difference was not significant, $t < 1$. These results are summarized in Table 3.

The pattern of participants' errors was also analyzed. Each incorrect response was classified as either a mistake to an adjacent category (e.g., A→ B or C) or as a mistake to the opposite category (e.g., A→D). These counts are presented in Table 2. As predicted, learners in the relation condition made more errors to the opposite category than did learners in the feature condition, $\chi^2(1, N = 3636) = 23.96,\ p < .001$.

Table 2: Confusion Matrix with standardized residuals

Label	Relations	Features
Adjacent	1056 (−2.03)	1473 (1.81)
Opposite	560 (3.06)	547 (−2.74)

The results from Experiment 4 support the alignment explanation for the relational advantage observed in Experiments 1 and 2. When a category structure is used in which alignment is not beneficial to learning relation-based categories, featural and relational learning are of equal difficulty. The drive to flexibly align is sufficiently strong that it engages even in situations in which it is not beneficial, as evidenced by the preponderance of opposite category errors in the relational condition.

General Discussion

Contrary to accepted wisdom, Experiment 1 demonstrated that learning to classify by relations can be easier than by features when stimuli and categories that offer a valid comparison are employed. Experiment 2 ruled out explanations of this relational advantage based on difficulties in processing absolute stimulus values. Experiments 3 and 4's results favored a flexible alignment explanation of the relational advantage. Pairwise similarity ratings collected in Experiment 3 supported the notion that learners flexibly align stimuli so as to maximize perceived similarity. Importantly, this flexible alignment process should have differentially benefitted relational learners in Experiment 1. Experiment 4 employed a category structure in which flexible alignment was not beneficial. Under these conditions, relation- and feature-based categories were acquired at the same rate. Alternative explanations based on a difference in processing of features or relations cannot account for the range of results.

Our findings add to a growing body of evidence that suggests a central role for relational processing in categorization. Many real-world categories have a strong relational basis (Gentner & Kurtz, 2005; Markman & Stilwell, 2001), as do many of the features that constitute categories that we do not view as relational (Jones & Love, (in press)). The present work complements this line of inquiry by examining relation-based learning of novel categories in a rigorously controlled experimental setting.

The current work does not address how people discover novel relations (cf. Doumas & Hummel, 2004), but instead focuses on how people learn novel concepts using existing relations. Nevertheless, our findings offer clues to how the relations we have direct the discovery of categories in the real-world. Whereas Rosch and Mervis (1975) focused on how the structure of the environment biases acquisition toward categories that have high within- and low between-category similarity, our findings suggest the cognitive machinery provided by flexible alignment can exert a strong influence in regularizing categories to conform to the Rosch and Mervis ideal.

Acknowledgments

This work was supported by AFOSR grant FA9550-04-1-0226, ARL grant W911NF-07-2-0023, and NSF CAREER grant #0349101 to B.C. Love.

References

Markman, A. B., Stilwell, C. H. (2001). Role-governed categories. *Journal of Experimental and Theoretical Artificial Intelligence*, *13*(4), 329-358.

Table 3: Summary of Results from Experiments 1,2, and 4

	Structure	Relations Relevant			Features Relevant			Accuracy
		Accuracy	SE	Criterion	Accuracy	SE	Criterion	Difference
Experiment 1	*XOR*	.73	.04	14/25	.54	.02	4/25	.19***
Experiment 2	*XOR*	.79	.06	9/14	.52	.02	2/15	.27***
Experiment 4	*Four Cat.*	.78	.04	22/26	.74	.04	22/27	.04

***$p < .001$

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121-152.

Doumas, L. A. A., & Hummel, J. (2004). Structure mapping and relational predication. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the cognitive science society* (p. 333-338). Mahwah, NJ: Lawrence Erlbaum Associates.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure mapping engine: Algorithm and examples. *Artificial Intelligence*, *41*, 1-63.

Garner, W. R. (1954). Context effects and the validity of loudness scales. *Journal of Experiment Psychology*, *48*(3), 218-224.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, *49*, 988-998.

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

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development*, *59*, 47-59.

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of douglas l. medin.* Washington, D.C.: American Psychological Association.

Gentner, D., & Ratterman, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development* (p. 225-277). London: Cambridge University Press.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*, 427-466.

Huttenlocher, J., Duffy, S., & Levine, S. (2002). Infants and toddlers discriminate amount: Are they measuring? *Psychological Science*, *13*, 224-249.

Johnson, S., Lowery, N., Kohler, C., & Turetsky, B. (2005). Global-local visual processing in schizophrenia: Evidence for an early visual processing deficit. *Biological Psychiatry*, *58*(12), 937-946.

Jones, M., & Love, B. C. ((in press)). Beyond common features: The role of roles in determining similarity. *Cognitive Psychology*.

Kittur, A., Hummel, J. E., & Holyoak, K. J. (2004). Feature- vs. relation-defined categories: Probab(alistic)ly not the same. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the cognitive science society* (p. 696-701). Mahwah, NJ: Lawrence Erlbaum Associates.

Kroger, J. K., Holyoak, K. J., & Hummel, J. E. (2004). Varieties of sameness: the impact of relational complexity on perceptual comparisons. *Cognitive Science*, *28*, 335-358.

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

Larkey, B., L., & Love, B. C. (2003). CAB: Connectionist Analogy Builder. *Cognitive Science*, *27*, 781-794.

Love, B. C., Rouder, J., & Wisniewski, E. J. (1999). A structural account of global and local processing. *Cognitive Psychology*, *38*, 291-316.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*, 249-273.

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

Stewart, N., Brown, G. D., & Chater, N. (2005). Absolute identification by relative judgement. *Psychological Review*, *112*, 881-911.

Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgements by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, *24*(3), 363-396.

Tomlinson, M. T., & Love, B. C. (2006). From pigeons to humans: Grounding relational learning in concrete examplars. In *Proceedings of the twenty-first national conference on artificial intelligence* (p. 199-204). Cambridge, MA: MIT Press.

Waltz, J. A., Knowlton, B., Holyoak, K., Boone, K., Back-Madruga, C., McPherson, S., et al. (2004). Relational integration and executive function in alzheimer's disease. *Neuropsycholgia*, *18*(2), 296-305.

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition*, *28*, 1205-1212.