# Progressive Alignment Facilitates Learning of Deterministic But Not Probabilistic Relational Categories

**Wookyoung Jung (jung43@illinois.edu)**
Department of Psychology, 603 E. Daniel Street
Champaign, IL 61820 USA

**John E. Hummel (jehummel@illinois.edu)**
Department of Psychology, 603 E. Daniel Street
Champaign, IL 61820 USA

## Abstract

Kotovsky and Gentner (1996) showed that presenting *progressively aligned* examples helped children discover relational similarities: Comparisons based on initially concrete and highly similar, but progressively more abstract exemplars helped the discovery of higher-order relational similarities. We investigated whether progressive alignment can aid learning of relational categories with either a deterministic (in which one relation reliably predicts category membership) or a probabilistic structure (in which each relation predicts category membership with 75% reliability). Progressive alignment helped participants learn relational categories with the deterministic structure. However, progressive alignment did not help participants learn the probabilistic relational categories. The results show that learning relational categories with a deterministic structure can be improved by progressive alignment, consistent with previous findings (e.g., Kotovsky & Gentner, 1996), but also support previous findings suggesting that relational categories are represented as a schemas, which are learned by a process of intersection discovery that fails catastrophically with probabilistic category structures (Jung & Hummel, 2009; Kittur et al., 2004, 2006).

**Key words:** Relational category learning; relational invariants; probabilistic category structure; deterministic category structure; progressive alignment

One of the most generally accepted assumptions in the study of concepts, categories and category learning is that we represent categories in terms of their exemplars' features and that the process of category learning is a process of learning which features are diagnostic of category membership. Likewise, the process of assigning an exemplar to a category is a process of comparing the exemplars' features to those of the category (represented either in terms of a prototype or as a collection of known exemplars). This account of category learning provides a natural basis for understanding the family resemblance structure of our concepts and categories: The idea, first proposed by Wittgenstein (1956) and subsequently supported by numerous experiments in cognitive psychology (for a review see Murphy, 2002), that, like the members of a family, the various members of a category tend to have many features in common with one another, but that there need not be any necessary or sufficient features for category membership. Feature-based theories of categorization also provide a natural account of the well-known *prototype effects* in categorization (e.g., the fact that an exemplar is judged to be a "good" member of a category to the extent that it shares many features with the prototype of the category; see Murphy, 2002).

Another widely held view is that this feature-based account of concepts and categories, for all its successes as an account of prototype effects, fails to provide a complete account of the richness and power of our conceptual structures. As pointed out by Gentner (1983), Barsalou (1993), Murphy and Medin (1985) and others, our knowledge of the interrelations among an object's "features" (e.g., that birds tend to fly and tend to nest in trees, but that not all do, and that only those that fly also nest in trees) and our ability to reason flexibly with our concepts (e.g., inferring that a man who jumps fully-clothed into a pool at a party is probably drunk; Murphy & Medin, 1985) seem to demand explanation in terms of more sophisticated conceptual structures, such as schemas and theories. The primary factor distinguishing schemas and theories from lists of features is that the former, but not the latter, represent relations explicitly: Whereas a feature list can specify that a bird "can fly" or "lives in trees", relational representations are required to specify that the ability to fly *enables* a bird to nest in trees. More generally, the schema/theory based view of concepts fares better as an account of the relations between our concepts and the larger theoretical structures in which they are embedded: We understand the relations, not just between the properties of various objects (e.g., the fact that a bird flies is what allows it to nest in trees), but between concepts and other concepts (e.g., that an interaction is when the effect of one variable depends on the value of another).

Moreover, some concepts and categories appear to be largely if not entirely relational in nature. For example, the category *conduit* is defined by a relation between the conduit and the thing it carries; *barrier* is defined by the relation between the barrier, the thing to which it blocks access and the thing deprived of that access; and even the category *mother* is defined by a relation between the mother and her child. Relational categories may be more the rule than the exception: As reported by Asmuth and Gentner (2005), informal ratings of the 100 highest-frequency nouns in the British National Corpus revealed that about half refer to relational concepts. The distinction between relational and feature-based categories need not pose a problem for the study of category learning and concept acquisition as long as relational and featural categories are learned in similar ways. But if they are learned in different ways, then little or nothing we know about the acquisition of feature-based categories necessarily need apply to the case of relational concepts and categories.

For example, Kittur, Hummel and Holyoak (2004; Kittur, Holyoak & Hummel, 2006) noticed that prototype effects had always been reported in experiments using categories defined by their exemplars' features. (For example, in an experiment using fictional "bugs" as stimuli, categories might be distinguished by features such as the shapes of a bug's head, wings, body and tail.) Kittur and colleagues wondered whether such effects would also be observed in categories defined, not by the exemplars' features, but by the relations among those features. To their surprise, they found that, rather than demonstrating prototype effects, people have great difficulty even learning relational categories with a probabilistic (i.e., family resemblance) structure in the first place.

They interpreted their findings in terms of people attempting to learn relational categories using a process akin to schema induction (e.g., Hummel & Holyoak, 2003). Specifically, Kittur et al. (2004) reasoned that if a relational category is represented as a schema, as has been proposed by others (e.g., Barsalous, 1993; Gentner, 1983; Holland, Holyoak, Nisbett, & Thagard, 1986; Keil, 1989; Murphy & Medin, 1985; Ross & Spalding, 1994), and if schemas are learned by a process of intersection discovery, in which a schema is learned from examples by keeping what the examples have in common and disregarding details on which they differ (as proposed by Hummel and Holyoak, 2003; see also Doumas, Hummel, and Sandhofer, 2008), then learning probabilistic relational categories ought to be extremely difficult because the intersection of the examples is the empty set (i.e., there is no single relation shared by all category members). More recently, Jung and Hummel (2009) replicated and extended the findings of Kittur et al., providing additional evidence that relational categories are indeed learned by a process of schema induction, and that this algorithm makes it very difficult for people to learn relational categories with a probabilistic (i.e., family resemblance) structure.

One potential alternative explanation for the findings of Kittur et al. (2004, 2006) and Jung and Hummel (2009) is that probabilistic relational categories are simply harder to learn than other kinds of categories, not that they rely on a qualitatively different learning algorithm than, say, feature-based category learning (e.g., schema induction in the relational case vs. simple associative learning in the feature-based case): That is, perhaps relational categories are simply harder to learn than feature-based categories, probabilistic categories are harder to learn than deterministic categories and these two sources of difficulty interact to push probabilistic relational categories over some kind of internal threshold, rendering them very hard for some to learn and impossible for others.

Various aspects of the Kittur et al. and Jung and Hummel findings are inconsistent with this interpretation. For example, in the data of Kittur et al. (2004), deterministic relational categories were not reliably more difficult to learn than feature-based categories. Indeed, Tomlinson and Love (2010) showed that, under some circumstances, relational categories can be much easier to learn than featural ones. In addition, Kittur et al. (2006) performed an ideal observer analysis on their category learning tasks and showed that human efficiency (i.e., human performance divided by ideal performance) is markedly lower in the probabilistic relational condition than in all three other conditions, suggesting that the performance difference in this condition is a function of human cognition, not the relational category learning task itself. Finally, Jung and Hummel (2009) were able to find a condition under which probabilistic relational category learning is not especially difficult (i.e., the "*who's winning*" task), suggesting that there is likely something "special" about the probabilistic relational category learning task. Nonetheless, additional evidence that probabilistic relational category learning is qualitatively different from both deterministic relational category learning and probabilistic feature-based category learning—especially evidence from a qualitatively different learning paradigm—would contribute both to our confidence that the difference is real and to our understanding of the nature of that difference.

If it turns out that probabilistic relational category learning is quantitatively different (e.g., more difficult) but qualitatively similar to deterministic relational category learning (i.e., in the sense that the same learning algorithm works for both), then interventions that aid deterministic relational category learning ought also to aid probabilistic relational category learning. By contrast if the learning algorithm that leads to successful deterministic relational category learning fails catastrophically in the case of probabilistic relational categories (as predicted by the schema induction-based account), then even interventions that facilitate the acquisition of deterministic relational concepts should be powerless to facilitate probabilistic relational concept acquisition.

One intervention that has been shown to facilitate the acquisition of deterministic relational concepts is *progressive alignment* (Kotovsky & Gentner, 1996). Progressive alignment is a training paradigm under which easy examples of a relational concept are presented earlier than harder examples of that concept, so that learning of the easy examples can bootstrap the acquisition of the harder ones. In the case of Kotovsky and Gentner, the learners were children and the relational concepts to be acquired were *symmetry* (e.g., *little*, *big*, *little* or *light*, *dark*, *light*) and *monotonic increase* (e.g., *little*, *big*, *bigger* or *light*, *dark*, *darker*). Kotovsky and Gentner trained children on these concepts using a matching task in which a sample stimulus was shown at the top of a display and two alternatives were shown at the bottom (see Figure 1). The child's task was to indicate which alternative matched the sample. Some of the trials were "easy" in the sense that featural information supported making the right relational choice (e.g., matching *little*, *big*, *little* in the context of squares onto *little*, *big*, *little* circles vs. *little*, *big*, *bigger* circles) whereas others were more difficult, requiring children to make a cross-dimensional match (e.g., matching *little*, *big*, *little* squares onto *light*, *dark*, *light* circles vs. *light*, *dark*, *darker* circles). When Kotovsky and Gentner trained the children on this task with randomly-ordered trials, the youngest children (4-year-olds) could not learn the task. But when the trials were progressively aligned, so that the easy trials came first and the more difficult cross-dimensional trials came later, then the 4-year-olds were able to master the task. Kotovsky and Gentner concluded that this procedure facilitated the learning of relational concepts by allowing the earlier trials to take advantage of featural support of correct responding and later trials to take advantage of generalization from the earlier trials.

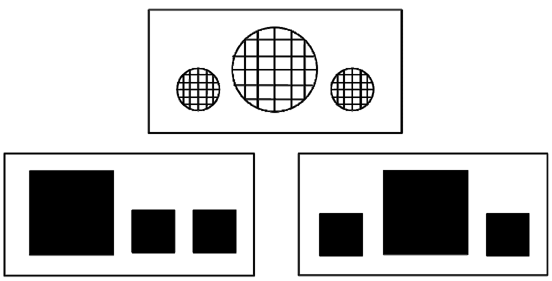

Figure 1. Example of Kotovsky and Gentner (1996) stimuli

To the extent that probabilistic relational category learning is simply harder than deterministic relational category learning (rather than being fundamentally incompatible with the learning algorithm that supports relational concept acquisition), then progressive alignment would be expected to facilitate both deterministic and probabilistic relational category learning. But to the extent that schema induction (or something akin to it) is invoked in response to all relational learning tasks and succeeds with deterministic concepts but fails catastrophically with probabilistic ones, then progressive alignment would be expected to facilitate deterministic relational concept acquisition but fail to facilitate probabilistic relational concept acquisition. It is this hypothesis that the current experiment was designed to test.

An additional purpose of the current experiment is to replicate the basic difficulty-of-probabilistic-relational-category learning effect with new stimulus materials. Kittur et al. (2004, 2006) used stimuli composed of octagons and squares, and Jung and Hummel (2009) used stimuli composed of circles and squares. The current experiment used fictional "bugs" as stimuli (Figure 2). The prototype of species A had a head that was wider and darker than its body, wings that were longer than its body and antennae longer than its tails. The prototypical B had the opposite relations: a head narrower and lighter than its body, wings shorter than its body and antennae shorter than its tails. In the probabilistic condition, any exemplar of A or B shared three relations with its own prototype and one with the prototype of the opposite category. In other words, the formal probabilistic category structures used are isomorphic with those used by Kittur et al. and Jung and Hummel. In the deterministic condition, one relation (counterbalanced) was rendered deterministically diagnostic of category membership simply by removing all exemplars containing the exception value of that relation.

The general procedure of the experiment involved first training subjects on the two bug species using a match-to-sample task like (but not identical to) that of Kotovsky and Gentner (1996): Three bugs were presented in a triangular pattern on the screen (either two from category A and one from B or vice versa) and the subject's task was to choose (with a mouse click) the odd man out (i.e., the B among As or the A among Bs). As elaborated below, the exemplars of a category could be more or less similar to members of their own or the opposite category as a function of how many relations they shared. "Easy" trials were those in which the same-category exemplars shared many relations with one another and few with the opposite-category member; "hard" trials had fewer shared relations within-category and/or more shared relations between categories. In the *progressively aligned* condition, easier trials were presented first, followed by progressively more difficult trials. In the *not progressively aligned* condition, trials were presented in a random order. The *deterministic* vs. *probabilistic* variable was crossed orthogonally with *aligned* vs. *nonaligned*, resulting in a two-by-two between subjects design. To the extent that deterministic relational category learning is categorically similar to probabilistic relational category learning, progressive alignment should be expected to facilitate both; but to the extent that they are qualitatively different, progressive alignment is expected to facilitate the former but not the latter.

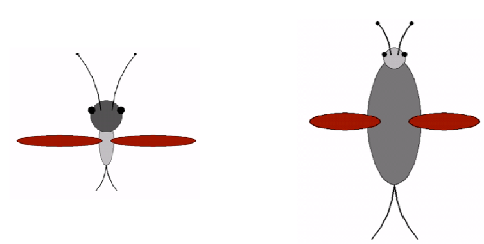

Figure 2. Example stimuli (prototype A and B, respectively)

## Method

**Participants.** A total of 49 subjects, randomly assigned to conditions, participated in the study for course credit.

**Materials.** Stimuli were line drawings of fictional bugs. The bugs vary in the size and darkness of their heads, the length, width and darkness of their bodies, and the lengths of their wings, antennae and tails. The prototype of category A is defined as [1,1,1,1] and the prototype of B is defined as [0,0,0,0], where [1,1,1,1] represents head *wider than* body, head *darker than* body, wings *longer than* body, and antennae *longer than* tails, and [0,0,0,0] represents head *narrower than* body, head *lighter than* body, wings *shorter than* body, and antennae *shorter than* tails. Exemplars of each category were made by switching the value of one relation in the prototype (e.g., category A exemplar [1,0,1,1] would have head *larger* body, head *lighter* body, wings *longer* body, and antennae *longer* tails). Metric values of head and body size and darkness, and of tail, antennae and wing lengths were chosen randomly to conform to the necessary relative values, rendering specific feature values (e.g., exact head width or darkness) undiagnostic of category membership (e.g., two instances of exemplar [1,1,1,0] would both have heads wider than their bodies, but would differ in their exact head and body widths).

**Design.** The experiment used a 2 (category structure: *probabilistic* vs. *deterministic*) X 2 (presentation order: *aligned* vs. *nonaligned*) between-subjects design.

**Procedure**. All trials used a triads choice task in which the subject was presented with two members of one category along with one member of the other and their task was to indicate which bug belonged to the odd category (i.e., the A among Bs or the B among As). Participants responded by mouse-clicking on the odd bug out and responses were followed by feedback showing the correct response. Triads differed in their difficulty, defined in terms of the number of shared relations between the same- and different-category exemplars. The easiest trials involved three within-category shared relations and either zero or one between-category shared relations. For example, the two within-category exemplars might be [1,1,1,1] and [1,1,1,0] (both members of A) and the remaining (category B) exemplar would be [0,0,0,0] (which shares one relation with the second member of A and zero with the first). We denote these trials as difficulty 1 (where difficulty = 4 – (shared-within-category – shared-between-category)). Difficulty 2 trials presented three within-category shared relations and one or two between-category shared relations (e.g., [1, 1, 1, 1], [1, 1, 0, 1] and [0, 0, 0, 1]). Difficulty 3 trials presented two within-category shared relations one between-category shared relations (e.g., [1, 1, 0, 1], [0, 1, 1, 1], and [0, 0, 0, 0]). The most difficult trials, difficulty 4, presented two within-category shared relations and two between-category shared relations (e.g., [1, 1, 0, 1], [1, 1, 1, 0], and [0, 0, 0, 1]). Note that in this most difficult case, within- and between-category exemplars are equally similar.

**Study phase—**Participants in the probabilistic condition were given 73 study trials (16 difficulty 1, 24 difficulty 2, 12 difficulty 3 and 21 difficulty 4). Those in the deterministic condition received 42 study trials (12 difficulty 1, 12 difficulty 2, 6 difficulty 3 and 12 difficulty 4). The number of study trials differed between the probabilistic and deterministic conditions because we made the deterministic condition by removing one exemplar from each category (counterbalanced across subjects), rendering one relation deterministically-related to category membership. Study trials in the progressively aligned condition were presented in order of difficulty, with difficulty 1 trials presented first and difficulty 4 last. Study trials in the nonaligned condition were presented in a completely random order.

**Transfer phase—**Following training, participants in the probabilistic condition were given 33 transfer trials (12 of difficulty 3 and 21 of difficulty 4) in a random order. Participants in the deterministic condition were given 18 transfer trials (6 of difficulty 3 and 12 of difficulty 4) in a random order. No feedback was given during the transfer phase.

## Results

**Accuracy.** Our primary interest was accuracy on the transfer trials. A 2 (probabilistic vs. deterministic) X 2 (aligned vs. nonaligned) X 2 (study vs. transfer) between-subjects ANOVA revealed main effects of both progressive alignment and category structure (Figure 3). There was a significant difference between aligned and nonaligned [$F$ (1, 90) = 12.641, $MSE$ = 0.205, $p < 0.01$] such that participants in the aligned condition ($M$ = 0.741, $SD$ = 0.233) showed more accurate transfer than participants in the nonaligned condition ($M$ = 0.616, $SD$ = 0.238). There was also a main effect of category structure [$F$ (1, 90) = 110.363, $MSE$ = 1.79, $p < 0.001$]. Participants in the deterministic condition ($M$ = 0.844, $SD$ = 0.219) transferred more accurately than those in the probabilistic condition ($M$ = 0.5, $SD$ = 0.081). In addition, there was a reliable interaction between progressive alignment and category structure [$F$ (1, 90) = 8.571, $MSE$ = 0.139, $p < 0.01$], indicating that progressive alignment improved accuracy in the deterministic condition, but not in the probabilistic condition. More interestingly, there was a reliable interaction between category condition and phase (i.e., study/transfer) [$F$ (1, 90) = 6.451, $MSE$ = 0.105, $p < 0.05$], indicating that for participants in the deterministic/progressive condition performance on the transfer trials ($M$ = 0.937, $SD$ = 0.102) was reliably more accurate than mean performance on the study trials ($M$ = 0.852, $SD$ = 0.086), [$t$(13) = 2.570, std. err. mean = 0.033, $p < 0.05$], whereas, for participants in the deterministic/random condition, performance on the transfer trials ($M$ = 0.736, $SD$ = 0.271) was no better than mean performance on the study trials ($M$ = 0.718, $SD$ = 0.129), [$t$(11) = 0.264, std. err. mean = 0.069, $p$ = 0.797]. Progressive alignment did not improve participants' learning in the probabilistic condition. Rather, participants in the probabilistic condition performed reliably less accurately during transfer than during training in both the aligned [training: $M$ = 0.581, $SD$ = 0.066, transfer: $M$ = 0.514, $SD$ = 0.079)], [$t(10)$ = -3.173, std. err. mean = 0.029, $p < 0.05$] and nonaligned conditions [training: $M$ = 0.578, $SD$ = 0.069, transfer: $M$ = 0.485, $SD$ = 0.085), [$t(11)$ = -2.559, std. err. mean = 0.026, $p < 0.05$]. That is, as predicted by the account that relational category structures are learned by a process of intersection discovery, progressive alignment was helpful for learning deterministic relational structures but not for learning probabilistic relational structures. That performance in the probabilistic condition was above chance during study suggests that subjects were learning something in this condition (e.g., it is possible to perform at 75% accuracy by focusing on a single relation), but the fact that this performance dropped back to chance during transfer suggests in the least that this learning was not very robust.

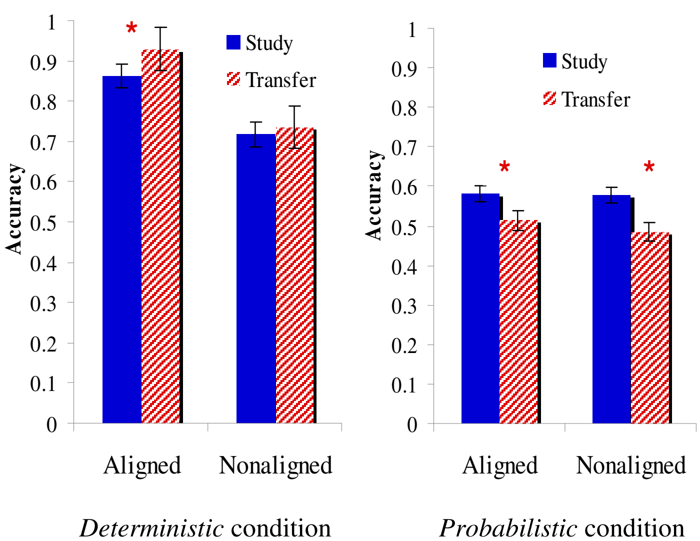


Figure 3. Accuracy by condition

## Discussion

Kittur at al. (2004, 2006) and Jung and Hummel (2009) reported that subjects have great difficulty learning relational categories with probabilistic (family resemblance) structures in which no single relation predicts category membership 100% of the time. They interpreted this result in terms of subjects' attempting to learn relational structures through a process of intersection discovery, which retains those features and relations exemplars have in common and discards those on which the exemplars differ (Doumas et al., 2008; Hummel & Holyoak, 2003). Such an approach to learning relational categories will work as long as there is one feature or relation shared by all category members, but it will fail catastrophically if all features and relations are related only probabilistically to category membership.

The current experiment provided additional support for this intersection discovery hypothesis. Our findings demonstrated that progressive alignment—a training regime that presents easy examples of a relational concept early in training, followed only later by more difficult examples (Kotovsky & Gentner, 1996)—facilitates learning of relational categories with a deterministic structure (in which one relation reliably predicts category membership) but does not facilitate learning relational categories with a family resemblance structure. This result suggests that the failure of intersection discovery in the face of probabilistic category structures is too catastrophic even to be ameliorated with a learning regime known to aid relational learning.

These findings contribute to the growing literature suggesting that feature- and relation-based categories may be learned in qualitatively different ways. Whereas feature-based categories can be learned in an associative manner that simply tabulates the frequency with which features and category labels co-occur—an algorithm that naturally tolerates family resemblance category structures—relational categories appear to demand learning in a qualitatively different way (Hummel & Holyoak, 1997, 2003). Whatever algorithm supports relational concept acquisition (whether intersection discovery or something else) is more powerful than association learning in the sense that (unlike associative learning) it can operate on relational structures at all (see Hummel, 2010, for a discussion of the differences between associative and relational learning). But it is weaker than associative learning in the sense that, unlike associative learning, it is too "brittle" to tolerate family resemblance structures.


## Acknowledgements

This research was supported by AFOSR Grant # FA9550-07-1-0147. We would like to thank Bobby Weisshappel for help running subjects.


## References


Asmuth, J. & Gentner, D. (2005). Context sensitivity of relational nouns. *Proceedings of the Twenty-seventh Annual Meeting of the Cognitive Science Society*. (pp.163-168). Hillsdale, NJ: Lawrence Erlbaum Associates.

Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A.C. Collins, S.E. Gathercole, & M.A. Conway (Eds.), *Theories of memory* (pp. 29-101). London: Lawrence Erlbaum Associates.

Doumas. L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115*, 1 - 43.

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

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.

Hummel, J. E. (2010). Symbolic vs. associative learning. *Cognitive Science*, *34*, 958-965.

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

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

Jung, W., & Hummel, J. E. (2009) Probabilistic relational categories are learnable as long as you don't know you're learning probabilistic relational categories. *Proceedings of The 31st Annual Conference of the Cognitive Science Society, Society* (pp. 1042-1047). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Kittur, A., Hummel, J. E., & Holyoak, K. J. (2004). Feature- vs. relation-defined categories: Probab(alistic)ly Not the Same. *Proceedings of the Twenty Six Annual Conference of the Cognitive Science Society* (pp. 696-701). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kittur, A., Holyoak, K. J., & Hummel, J. E. (2006). Using ideal observers in higher-order human category learning. *Proceedings of the Twenty Eight Annual Conference of the Cognitive Science Society* (pp. 435-440). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*, 2797-2822.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

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

Ross, B. H., & Spalding, T. L. (1994). Concepts and categories. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp119-148). San Diego, CA: Academic Press, Inc.

Tomlinson, M.T., & Love, B.C. (2010). When learning to classify by relations is easier than by features. *Thinking & Reasoning*,16, 372-401.

Wittgenstein, L.(1956). Bemerkungen über die Grundlagen der Mathematik. Blackwell, Oxford.