LEARNING AND REPRESENTATION OF RELATIONAL CATEGORIES

BY

WOOKYOUNG JUNG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Doctoral Committee:

Professor John E. Hummel, Chair
Assistant Professor Andrei Cimpian
Professor Gary S. Dell
Assistant Professor Robert G. Morrison
Professor Brian H. Ross

ABSTRACT

Relation-based category learning is based on very different principles than feature-based category learning. It has been shown that relational categories are learned by a process akin to structured intersection discovery, which is formally powerful than feature-based associative learning, but which fails catastrophically with probabilistic category structures. This research provided consistent evidence that relational concepts are qualitatively different from featural concepts, and they are also learned in a qualitatively different manner. Experiment 1 showed that relational category learning with probabilistic structures can be improved by comparing systematic pairs of exemplars, where shared relations between the exemplars can be abstracted. Experiment 2 showed that comparing the exemplars to the prototype can improve learners' ability to learn probabilistic relational categories in terms of prototype-plus-exception rules. Experiment 3 and 4 examined further the distinction between feature-and relation-based category learning using a dual task methodology. Experiment 3 revealed that featural category learning was more impaired by a visuospatial dual task than by a verbal dual task, whereas relational category learning was more impaired by the verbal dual task. Experiment 4 examined how the dual task that involves more relational information interacts with feature-and relation-based category learning. The results showed that there was no reliable difference between two category learning. Taken together, Experiment 3 and 4 results suggest that in contrast to featural category learning, which may involve mainly non-verbal mechanisms, relational category learning appears to place greater demands on more explicit and attention-demanding verbal or verbally-related learning mechanisms. The findings presented in this dissertation contribute to the growing body of

theoretical and empirical results suggesting that relational thought is a *qualitatively* different thing than the kinds of thinking and learning afforded by feature-based representations of the world.

This dissertation is dedicated to my parents

ACKNOWLEDGMENTS

The completion of my dissertation has been a long journey. I could not have come this far without the assistance of many individuals and I want to express my deepest appreciation to them. Foremost, I would like to give a heartfelt, special thanks to my advisor, John Hummel, for his mentorship, friendship, patience, and for being a role model to me. John has been a strong and supportive advisor to me throughout my graduate school career, but he has always given me great freedom to pursue independent work. The completion of this research would not have been possible without his guidance and support. I am also grateful to Brian Ross for his mentorship and guidance over the past few of years. Many thanks go to the other members on my dissertation committee, Gary Dell, Andrei Cimpian, and Robert Morrison. I greatly appreciate that they have taken to time to carefully read through my dissertation and share their valuable insights.

I also would like to thank many friends and colleagues for helpful comments and discussion: the members of Relational Reasoning Lab, and the members of Ross Lab.
I would also like to thank my family and friends for their constant love, encouragement, support, and especially their prayers. My dear parents and brother have always prayed for me and believed in me. They have always encouraged me to finish this journey and accomplish my goals.

I would like to thank grants for financial support (AFOSR Grant FA9550-07-1-0147, Grant from the University of Illinois Research Board). I would also like to thank the Psychology department at the University of Illinois for treating its graduate students with such respect and care.

I most want to thank my Lord, Jesus Christ, for giving me the strength to the weary and protecting me, as always.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

The ability to acquire and reason about relational concepts is a cornerstone of human thinking. It is the basis of our ability to grasp analogies between seemingly different objects or situations (e.g., Bassok, 2001; Clement & Gentner, 1991; Gentner, 1983; Gentner & Smith, 2013; Gick & Holyoak, 1980; Goswami, 2001; Holyoak, 2005; Holyoak & Thagard, 1995; Markman & Gentner, 2000), to infer hidden causes of observed events (e.g., Gopnik & Melzoff, 1997; Gopnik, Sobel, Schulz, & Glymour, 2001), to apply abstract rules in novel situations (e.g., Smith, Langston & Nisbett, 1992), and even to appreciate perceptual similarities (e.g., Palmer, 1978; Goldstone, Medin & Gentner, 1991; Hummel, 2000; Hummel & Stankiewicz, 1996). Along with language, our capacity to think explicitly about relations may be the primary factor separating human cognition from the cognitive abilities of our closest primate cousins (see, e.g., Penn, Holyoak & Povinelli, 2008).

Relational concepts are concepts that specify the relations between things rather than just the literal features of the things themselves: A *barrier* is something that stands between one thing and another; a *conduit* is something that transports something else (water, electricity, karma) from one place to another; a *friend* is someone who likes and is liked by another. Although it is tempting to think of nouns as referring to concrete objects defined by simple lists of features (e.g., “a bird has feathers and wings and lives in trees”)—and although, as reviewed shortly, the vast majority of laboratory research on category learning has been based on such feature-based categories—about half of 100 highest-frequency nouns in the British National Corpus refer to relational concepts

(Asmuth & Gentner, 2005). Relational categories may thus be more the rule than the exception.

Given the centrality of relational concepts in human thinking, it is important to understand what such concepts consist of and how they are acquired—and the degree to which the answer to the first question imposes constraints on the answer to the second. In this Dissertation, I shall explore several empirical implications of the *intersection discovery hypothesis*: The hypothesis that relational concepts are learned by a process of *structural alignment* (a.k.a. *analogical mapping*; Gentner, 1983; Gick & Holyoak, 1980; 1983), which makes explicit the relational correspondences between otherwise featurally different examples, combined with a form of *intersection discovery*, in which the shared elements and relations between those systems are retained while elements or relations unique to one system or the other are discarded (Doumas, Hummel & Sandhofer, 2008; Gick & Holyoak, 1983; Hummel & Holyoak, 2003). As a learning algorithm, intersection discovery is both formally more powerful than associative learning (e.g., as discussed in the literature on animal learning [e.g., Rescorla & Wagner, 1972] and in traditional models of category learning in human subjects [e.g., Krushke, 1992]) and, in counterintuitive ways, more limited in the kinds of category structures it is equipped to acquire.

The Dissertation is organized as follows. I shall first review evidence for various kinds of relational concepts in human cognition. I next discuss the problem of learning relational concepts from examples. I will argue that such concepts are formally too complex to be acquired by means of traditional associative learning. I will present the intersection discovery hypothesis as a potential solution to the limitations of associative learning and summarize prior support for that hypothesis. The main part of the Dissertation consists of four experiments testing additional predictions of the intersection discovery hypothesis, and the implications of concept acquisition qua relational learning more generally. Finally, I will conclude with a discussion of the implications of my findings for our understanding of relational concepts and the conditions under which they can and cannot be learned.

1.1. Prior research on relational categories

Although relational concepts are ubiquitous in human cognition, it would be a mistake to assume that "relational concept" is a monolithic term. Instead, relational concepts appear to manifest themselves in several more specific ways in human cognition.

Intrinsic vs. Extrinsic Properties. Barr and Caplan (1987; Caplan & Barr, 1991) made a distinction between *intrinsic* and *extrinsic* features. Intrinsic features are those that belong to an entity in isolation, such as "has wings" for birds, whereas extrinsic (i.e., relational) properties refer to relations between two or more entities, such as "used to work with" for a hammer. In Barr and Caplan's experiments (1987), participants in a pilot study were asked to list members of categories consisting of natural kinds and artifacts. Barr and Caplan then asked another group of participants to rate on a 1…7 scale the degree to which each of the category members collected from the pilot study had intrinsic or extrinsic properties. The results showed that artifacts (such as toys, tools, weapons, vehicles, sports, and furniture) were more defined by extrinsic features whereas natural kinds (such as trees, fruit, mammals, birds, and flowers) were more defined by intrinsic features. Barr and Caplan (1987) also reported partial membership scores, which were responses not falling on either endpoint of the membership scale. That is, on their scale, any response from 2 to 6 was scored as partial membership. They found that higher proportion of participants provided partial membership judgments about the members of extrinsic categories, and concluded that extrinsic concepts show more graded membership than intrinsic concepts.

Isolated vs. Interrelated Concepts. Goldstone (1996) and his colleagues (e.g., Goldstone, Steyvers, & Rogosky, 2003) explored a number of metrics for measuring the degree to which a concept is isolated (i.e., featural) or highly interrelated with other concepts (i.e., relational). Goldstone (1996) argued that relatively interrelated concepts can be identified by the minimul use of nondiagnostic features and more by a caricature than a prototype. In the first experiment, the stimuli were 3 x 3 grid line segments consisting of horizontal, vertical, and diagonal lines. Participants in the isolated condition were instructed to create an image of the two concepts to be learned. Participants in the interrelated condition were instructed to seek out stimulus features that served to distinguish the concepts. He found that nondiagnostic line segments did not have much influence on categorization accuracy in the interrelated condition relative to diagnostic lines. In contrast, nondiagnostic line segments had a greater influence in the isolated condition.

Another experiment presented participants with four categories, each consisting of seven vertical bars (resembling a histogram). Each category included a prototype and a caricature. The prototype was defined as the exemplar that presented average values along the dimensions that comprise the category’s members, whereas a caricature was defined as an extreme exemplar, specifically, an exemplar that presented values that departed from the central tendency of the category in the opposite direction of the central tendency of other, simultaneously acquired categories (Goldstone, et al., 2003). Given that accuracy rates were above 90%, the main interest was response time. Categorizing caricatures was generally faster than categorizing prototypes. This speed advantage was particularly pronounced when participants was instructed to discriminate features. Participants in the interrelated condition whose task was to seek out stimulus features that served to distinguish the concepts showed faster performance when categorizing caricatures than prototypes. By contrast, participants in the isolated condition whose task was to create an image of the two concepts to be learned did not show the caricature advantage. Together, these findings suggest that isolated (i.e., featural) categories may be better characterized by prototypes than by ideals or caricatures, whereas interrelated (i.e., relational) categories may be better characterized by ideals or caricatures than by prototypes (see also Kittur et al., 2006b).

Natural Kinds vs. Nominal Kinds. Kloos and Sloutsky (2004) made a distinction between *natural kind* concepts, which have dense correlational structures and *nominal kind* concepts, which are based on sparse rule-like structures. They investigated how people learn natural and nominal categories using artificial stimuli. Members of natural categories had a set of correlated features in common, whereas members of nominal categories had a single relation in common. Participants were asked to learn a category by observation with many instances of the category or by a rule-like definition such as *category members have relation X*. The results showed that observation is a better way to learn natural categories, whereas discovering an explicit rule is a better way to learn nominal categories. Their findings suggest that a difference in representational density needs different category learning regimes: Dense concepts are learned better in a more implicit (observation-based) way, whereas sparse ones are learned better in a more explicit (rule-based) way. This distinction also suggests that entity and relational categories rely, not only on different kinds of mental representations, but also on different kinds of learning algorithms.

Role-governed categories. Markman and his colleagues (Markman & Stilwell, 2001; Goldwater, Markman & Stilwell, 2011) have argued that, unlike featural categories in which labels refer to categories defined by their members' features, role-governed categories are defined by items that play particular roles in a more global relational structure. Examples include categories such as doctor, advisor, private (the military rank), and so on. Those examples are certainly composed of features in some way, but rather it is the relational information that separates role-governed from feature-based categories, not specific features. Goldwater et al. (2011) provided empirical evidence to support the existence of role-governed categories. They showed that our knowledge of role-governed categories, in contrast to feature-based categories, is largely about properties extrinsic to category members. They also showed that, when asked to choose words to describe feature-based categories, people tend to choose words describing typical category characteristics; but when asked to choose words describing role-governed categories, people tend to choose words describing ideal characteristics (see also Goldstone, 1996; Kittur, et al., 2006b).

In addition, Goldwater and Markman (2011) examined factors that increase people's sensitivity to role-governed categories. In a novel-word extension study, a triad consisted of one target category, a role-governed alternate and a thematic alternate (e.g., “bird’s nest” for a target word, “house” for a role-governed alternate, and “tree” for a thematic alternate). On each trial, participants were given either a label or a description for the exemplars. In the label condition, the query was, for example, “The target is a goppin. Which of these other two is better called goppin?” In the description condition, the query would be “It’s a goppin target. Which of these other two is better called goppin?”. The results showed that participants in the label condition chose role matches more frequently than participants in the description condition. Goldwater and Markman interpreted this result to indicate that labels induce analogical comparison (Gentner, 2003; Namy & Gentner, 2002; Yamauchi, 2009), which aligns elements on the basis of common relational roles.

In their next study, half of the participants were provided a similarity rating task followed by a categorization task, and the other half were provided an imageability rating task followed by a categorization task. For the similarity rating task, the target and one of the alternates were presented (the role-matched alternate or the thematic matched one). The query was, “How similar are the target and alternative 1? How similar are the target and alternative 2?” Similarly, for the imageability task, the target and one of two alternates (role-match or thematic-match) were presented and the participant was asked, “Which is easier to picture in your head: the target or one alternative 1? “Which is easier to picture in your head: the target or one alternative 2?” The categorization task then was provided to all participants. The query was “which of these two better go with target to make a category?” The results showed that participants in the similarity condition chose more often the role-matched alternative than participants in the imageability condition, which does not require a comparison of the elements of the mental representation of the concepts. Taken together, these findings suggest that similarity comparisons as well as labels create a general sensitivity to role-governed categories that persist beyond the specific items on which the judgments were made.

Ad-hoc Categories. Barsalou's (1983, 1985) *ad-hoc categories* are categories constructed spontaneously to achieve a goal, such as "things to take out of the house during a fire". The members of this category, which include things such as cash, pets, family photos, laptops, and so forth, typically lack any intrinsic (i.e., featural) similarity. The only "feature" the members of this category have in common is that they are all things to take out of the house in case of fire. In contrast to the members of feature-based categories, which have a graded structure around a central tendency (i.e., the prototype), ad-hoc categories show a graded structure around an ideal (properties that optimally promote goal resolution) as in the example of *foods with no calories* for "things to eat on a diet". The centrality of a specific goal suggests that relational category representations may have a relatively sparse, rule-like nature (e.g., see also Kittur et al., 2004; Kittur et al., 2006b; Kloos & Sloutsky, 2004).

Abstract Coherent Categories. Rehder and Ross (2001) proposed that a kind of relational category they called *abstract coherent categories* can be acquired on the basis of relationships that are independent of the specific attributes of exemplars, as long as the relationships are maintained in a way consistent with prior expectations. In their study, three exemplars of the abstract coherent category "morkels" were presented: one morkel "operates on the surface of water, works to absorb spilled oil, coated with spongy material," while another "operates on land, works to gather harmful solids, has a shovel". The members of this category lack featural overlap but take their structure from systems of features that support a common abstract relation (i.e., that a morkel's features work sensibly together to satisfy a goal). They argued that the human conceptual system is closely related to abstract coherent concepts.

Thematic Relations. Thematic relations can be also seen as yet another kind of relational category. A thematic relation is generally defined as any temporal, spatial, causal or functional relation between things that perform complementary roles in the same scenario (Estes, Golonka & Jones, 2011, Golonka & Estes, 2009; Lin & Murphy, 2001; Wisniewski & Bassok, 1999). Examples of thematic relations include the relation between cows and milk and the relation between bagels and cream cheese. Thematic relations are external in that they occur between multiple objects, concepts, people or events and complementary in the sense that the arguments of a thematic relation fill complementary roles of the relation (e.g., cows are producers and their milk is the products; Estes et al., 2011). In this way, the arguments of a thematic relation differ from members of an ad-hoc category: There is no sense in which the members of an ad-hoc category are bound to complementary roles. The sense in which thematic relations form (typically very small) categories is that the arguments of a thematic relation can be viewed as the "members" of the category.

1.2. What are relations?

In spite of the diversity of these different kinds of concepts, they all share the property that they are defined, not by the literal features of their exemplars, but by relations, either between the features of an individual exemplar or between the exemplar and other objects (including the person doing the categorizing). In this context, it is important to say what a relation *is*.

Formally, a relation is a subset of the Cartesian product of two or more (potentially) infinite sets. Consider the set of integers and imagine an infinite table, with both columns and rows labeled with the integers, $0 \ldots \aleph_0$ (the table extends forever to the right and to the bottom). In the cells of the matrix, insert a value of 1 (i.e., *true*) if the integer in the corresponding row is larger than the integer in the corresponding column and insert the value 0 (i.e., *false*) otherwise. The result is an infinitely large table with 1s below the main diagonal and 0s everywhere else; this table is the mathematical definition of the relation *larger-than* (*row*, *column*).

Intuitively, a relation (or, more precisely, a *relational role*) is a property of an object (e.g., an integer) whose truth value is impossible to establish without reference to at least one other object: The number 42 cannot simply be *larger* or *smaller*; it can only be larger or smaller than some other number. (Note that this fact is manifest in the mathematical definition of a relation.) Similarly, in the statement "John gave the book to Mary", John is not simply a *giver*; he is the giver of a book to Mary.

From this formal perspective, almost every concept satisfies the definition of a relation. For example, even the response of a ganglion cell in the retina express a relation between the amount of light on one spot on the retina and the amount of light on an adjacent spot. This fact has given rise to some confusion in the literature about what constitutes a "relational representation" (see, e.g., Hummel, 2010, for a discussion).

From a *psychological* perspective—and for the purposes of this Dissertation—a representation is relational if and only if it makes the relation in question *explicit*, that is, if some element of the representation corresponds specifically to the relation, independently of whatever arguments it happens to be taking at the time (see Fodor & Pylyshyn, 1998; Doumas, et al., 2008). According to this definition, the expression "*loves* (John, Mary)" is explicitly relational because it represents the relation *loves* independently of its arguments (e.g., as evidenced by our ability to evaluate what *loves* (John, Mary) has in common with, and differs from, *loves* (Mary, John) or *loves* (Bill, Susan)). The ganglion cell, by contrast, is not explicitly relational because a different cell is required for every different possible contrast on the retina; that is, the "relation" represented by a ganglion cell (i.e., a local contrast value) is not independent of its arguments (the photoreceptors representing the luminance values giving rise to that contrast).

In turn, the requirement that relations be represented independently of their arguments implies that there must be some basis for specifying dynamically (i.e., on the fly) which arguments any given relation happens to be taking at any given time. In the case of propositional notation, this "binding tag" is list position within the parentheses (e.g., in *loves* (John, Mary), the binding of John to *lover* is specified by his first position in the parentheses). From a psychological perspective, what is important is that this kind of dynamic binding (whatever the "tag" happens to be at, say, a neural level) requires attention and consumes finite working memory resources. As elaborated shortly, the resulting demands on attention and working memory are among the major hallmarks of explicitly relational (a.k.a., symbolic) processing in the literature.

1.3. A learning algorithm for relational concepts

Given the many kinds of relational categories, an important question is how such concepts are acquired: How do we come to know what a barrier is, or what *larger-than* means? This question is complicated by the fact that, by adulthood at least, many of our relational concepts are independent of (i.e., invariant with) their arguments (Doumas et al., 2008; Hummel & Holyoak, 1997, 2003): We understand that *larger-than* means the same thing in the statement "Jupiter is larger than Saturn" as in the statement "The nucleus of an atom is larger than the electrons", even though Jupiter and Saturn are very different than atomic nuclei and electrons.

This kind of argument-invariance poses a difficulty for learning because, although we eventually come to understand relations as distinct from their arguments, we never actually get to experience relations disembodied from their arguments: No one has ever seen an instance of *larger-than* without some specific thing that was larger than some specific other thing. The argument-invariance of relational concepts poses a problem for learning because it implies that associative learning is formally too weak to explain the acquisition of relational concepts (Chomsky, 1959; see also Doumas et al., 2008; Hummel, 2010; Hummel & Holyoak, 1997, 2003; Kittur, Hummel & Holyoak, 2004), a fact that may help to explain why so few species are capable of learning relational concepts.

The reason, in brief, is that associative learning is tied to the co-occurrence statistics of the features of the concepts so acquired. If, in a category learning experiment using artificial "bugs" as stimuli, a given head type, H, occurs 75% of the time with a given category label, C, then it is possible to learn associatively that H predicts C.

Relations, by contrast, cannot be predicted based strictly on co-occurrence. We know, for example, that a neon atom is larger than an electron even if we have never explicitly considered this comparison before. What is worse is that any co-occurrence statistics a person has had the opportunity to observe may contradict whatever relation(s) in which an object currently stands: A neon atom is smaller than everything else with which the average person ever has any experience, so according to co-occurrence statistics alone it is impossible to even imagine its being larger than anything.

In response to the inadequacy of associative learning to explain the acquisition of relational concepts, some researchers have proposed that relational concepts, including both full-blown schemas (e.g., Gick & Holyoak, 1983; Hummel & Holyoak, 2003) and individual relations, such as *larger-than* (e.g., Doumas et al., 2008), are learned by a process of structured intersection discovery. The basic idea is that two situations—e.g., two love triangles, in the case of a *love triangle* schema (see Hummel & Holyoak, 2003), or two instances of one thing being larger than another, in the case of the *larger-than* relation (see Doumas, et al., 2008)—are structurally aligned (Gentner, 1983), making the correspondences between their parts explicit. For example, in the case of the love triangle schema, the learner may observe two instances of a love triangle (e.g., John loves Mary, but Mary loves Bill, so John is jealous of Bill, and Jill loves Mike, but Mike loves Betty, so Jill is jealous of Betty), notice the analogy between them (mapping John to Jill, Mary to Mike and Bill to Betty) and induce a schema by discovering the intersection of the two examples, that is, retaining the things they have in common and discarding the details on which they differ (in this case, person1 loves person2, person2 loves person3, so person1 is jealous of person3). Learning by intersection discovery is formally more powerful than simple associative learning because it relies on the machinery of structural alignment, i.e., analogy (see Doumas, et al., 2008; Gentner, 1983; Hummel & Holyoak, 2003), making it sensitive to the abstract (including higher-order) relational structure of the concepts being compared.

In contrast to relational concepts, which are too complex to learn as simple associations (because of the argument invariance property), featural concepts (i.e., concepts defined by their exemplars' features, rather than by relations) can be learned associatively. For example, if members of category X tend to have features A1, B1 and C1, whereas members of category Y tend to have features A2, B2 and C2, then it is possible to discriminate Xs from Ys simply by learning associative links (e.g., weighted connections in a connectionist network, or associative links as learned by the Rescorla-Wagner [1973] model) from A1, B1 and C1 to X and from A2, B2 and C2 to Y. That is, there is good reason to believe that relational and featural concepts require very different learning algorithms: Intersection discovery (or some other algorithm that exploits the machinery of structure mapping) in the case of relational concepts vs. simple association in the case of featural concepts (Hummel, 2010; Hummel & Holyoak, 2003).

To the extent that different learning algorithms underlie the learning of relational and featural concepts, then conclusions drawn from experiments using one kind of category may not necessarily apply to the other kind. One of the most robust and replicable conclusions from the literature on category learning from the 1970s to the present (e.g., Kruschke, 1992; Kruschke & Johansen, 1999; Markman & Maddox, 2003; Minda & Smith, 2011; Rosch & Mervis, 1975; Shiffrin & Styvers, 1997; Smith & Medin, 1981) is that people easily learn categories with a *family resemblance*, i.e., *probabilistic* structure. In a category with a family resemblance structure, there is no single feature shared by all members of the category. Rather, features tend to occur probabilistically, and "good" members of the category (i.e., members closer to the prototype) tend to have more features in common with other members of the category than "bad" members. The observation that people easily learn categories with a family resemblance structure leads naturally to the conclusion that our natural concepts also have a family resemblance structure, as famously suggested by Wittgenstein (1953).

However, as observed by Kittur and colleagues (Kittur, Holyoak, & Hummel, 2006b; Kittur, Hummel, & Holyoak, 2004), one limitation of this conclusion is that all the experiments demonstrating our ability to learn probabilistic category structures have been performed using feature-based categories. If feature-based categories are learned associatively, then they should be easily learnable even if they have a probabilistic structure (provided the features are sufficiently predictive of category membership). But if relational categories defy learning by association—and in particular, if they are learned by a process akin to structured intersection discovery—then they should not be learnable when they have a family resemblance structure: If there is no relation that all members of a relational category have in common, then the intersection of the category's exemplars will be the empty set. That is, the intersection-discovery theory of relational learning predicts that probabilistic relational categories ought to be (virtually) unlearnable.

1.4. Kittur, Hummel and Holyoak (2004)

Kittur et al. (2004) set out to explicitly test the prediction that probabilistic relational categories ought to be difficult to learn. Each exemplar in their experiments consisted of an octagon and a square. In the relational condition of this experiment, the prototypes of A and B were defined by the relations between the octagon and the square. In the prototype of category A, the octagon was *larger* than the square (‘1’), *darker* than the square (‘1’), *above* the square (‘1’) and *in front of* the square (‘1’). In the prototype of B, it was *smaller* than the square (‘0’), *lighter* than the square (‘0’) *below* the square (‘0’) and *behind* the square (‘0’). The precise size and darkness (i.e., features) of the octagon and square did not matter for category membership and were allowed to vary across exemplars within a category. In the featural condition of this experiment, A and B were defined by the precise size and darkness of the octagon and square.

Orthogonally crossed with the relational vs. featural conditions, half the participants learned family resemblance (a.k.a., probabilistic) category structures and the other half learned deterministic category structures. In the family resemblance condition, exemplars were constructed from the prototypes by switching one relation or feature in the prototype to its value in the opposite prototype, as in the example above. (For example, if the prototype of category is denoted [1,1,1,1] and the prototype of B is [0,0,0,0], then the four exemplars of A would be [0,1,1,1], [1,0,1,1], [1,1,0,1] and [1,1,1,0]. Note that, although any given exemplar contains ¾ of the corresponding prototype’s features/relations, no exemplar contains all of the prototype’s features/relations, and no feature/relation appears in all the prototypes of a category. It is in this sense that the category has a “family resemblance” or “probabilistic” structure.) The deterministic condition was constructed from the probabilistic condition simply by discarding one exemplar (counterbalanced) from each category, so that one feature or relation was perfectly (i.e., deterministically) predictive of category membership. For example, discarding the first exemplar of each category leaves exemplars [1,0,1,1], [1,1,0,1] and [1,1,1,0] for A and [0,1,0,0], [0,0,1,0] and [0,0,0,1] for B, so that the first relation (or feature) is deterministically '1' across all members of A and '0' across all members of B.

Kittur et al. (2004) found that although the feature-based categories were easy for participants to learn whether they were deterministic or probabilistic, and although relational categories were easy to learn as long as they were deterministic, relational categories were extremely difficult to learn when they were probabilistic. Indeed, half their participants never learned the probabilistic relational categories, even after 600 exposures to the exemplars.

Their findings support the hypothesis that people learn relational concepts by a process of intersection discovery (Hummel & Holyoak, 2003), in which they compare examples of relational concepts to one another, retaining what the examples have in common and discarding or discounting the details on which they differ. In the case of a probabilistic category structure, the intersection is the empty set, rendering the category unlearnable. Thus, although Hummel and Holyoak (2003) and Doumas, et al. (2008) showed that intersection discovery is capable of learning complex, relational concepts, it fails catastrophically when those concepts have a probabilistic structure. This finding has since been replicated by Kittur et al. (2006b) and Jung and Hummel (2009a, 2009b, 2011).

1.5. Jung and Hummel (2009a, 2009b): Rendering probabilistic relational categories learnable

Jung and Hummel (2009a) sought to further test the intersection discovery account of relational learning by examining the conditions under which probabilistic relational categories could be made learnable. Murphy and Allopenna (1994) and Rehder and Ross (2001) showed that category structures that map onto learner's existing schemas are easier to acquire than those that do not. Accordingly, Jung and Hummel reasoned that, faced with the task of learning probabilistic relational categories, anything that encourages the learner to discover a higher-order relation that remains invariant over members of a category—effectively rendering the category deterministic—ought to substantially facilitate learning.

Jung and Hummel (2009a) used a category structure isomorphic to that of Kittur et al. (2004), except that instead of octagons and squares, they used circles and squares. In all conditions of their first experiment, participants were trained on the (unlearnable) probabilistic relational category structure used by Kittur et al. (2004). In one condition, participants were instructed to categorize the stimuli as members of A or B (just as in Kittur et al.). In another condition, participants were instructed, not to categorize the stimuli, but to press the A key "if the circle was winning" in a given stimulus or the B key "if the square was winning". Participants were not told what "winning" meant; rather, they were told that they would figure it out as they went along. Crucially, any stimulus that a participant in the categorize condition would correctly categorize as an A was a stimulus for which a participant in the "who's winning" condition would correctly say "the circle is winning"; and any stimulus correctly categorized as a B was one in which "the square was winning". That is, the "categorize" and "who's winning" tasks were fully isomorphic, differing only in the instructions participants received at the beginning of the experiment, and thus in the task participants believed themselves to be performing.

The result was that participants in the categorize condition, just like Kittur et al.'s participants, had tremendous difficulty learning the category structures, with fewer than half of them ever learning to criterion. By contrast, participants in the "who's winning" condition learned much faster, with all of them reaching criterion well before the end of the experiment. That is, the "who's winning" task rendered the otherwise unlearnable probabilistic relational category structure learnable.

A series of follow-up experiments (Jung & Hummel, 2009b) systematically investigated the reasons for and nature of this effect, and the short story is that the "who's winning" task seems to engage people's knowledge of winning/losing relations, thereby making them tolerant of the probabilistic structure of the categories. People know that for a team to win it is not necessary for that team to get all the points; it is only necessary for them to get more points than the other team. Switching the task from "categorize" to "who's winning?" thus seems to have allowed participants in the "who's winning" condition to discover a higher-order property (namely, something like "has more points") that does indeed remain invariant over all members of a category: In every member of category A, the circle "has more points" than the square, and in every member of B, the square "has more points" than the circle. These data are thus consistent with the idea that relational learning (via intersection discovery) requires some kind of invariant (in this case, "has more points") in order to succeed.

How does the who's winning task help intersection discovery in the probabilistic relational category learning situation? Jung and Hummel (2009b) hypothesized two possibilities: the *comparison* hypothesis and the *winning schema* itself hypothesis. The first hypothesis was that the who's winning task facilitates learning simply by encouraging participants to compare the circle and square in some manner that the category learning task does not. For example, perhaps participants in the *who's winning* condition represented the circle and square as separate objects and doing so facilitated learning by encouraging them to compare them to one another. On this account, any task that encourages participants to represent the circle and square as separate objects engaged in a relation (like winning/losing) ought to facilitate learning. For example, asking participants "who's daxier?" should encourage the same kind of comparison as "who's winning?" and result in a comparable improvement over "to which category does this example belong?".

Their second hypothesis was that a schema for what "winning" consists of may facilitate learning by encouraging participants to count the number of "winning" roles (i.e., "points") bound to the circle and the square and to declare whichever part has more winning roles the winner. On this account, the effect of "who's winning" reflects the operation of the "winning" schema, per se, rather than simply the effect of comparisons encouraged by instructions that suggest the circle and square are separate objects.

Where these hypotheses make divergent predictions is in the potential role of *consistent* vs. *mixed role assignment* in the effect. In Jung and Hummel's (2009a) experiment, the assignment of relational roles to categories was *consistent*, in the sense that all the roles named in the instructions were assigned to the circle in category A (with the unnamed roles assigned to the square) and all the roles not named in the instructions were assigned to the circle in category B (with the named roles assigned to the square). That is, given category A, all the relational roles named in the experimental instructions— specifically, *darker*, *larger*, *above*, and *in front*—correspond to the circle, and all the unnamed roles (*smaller*, *lighter*, *below*, and *behind*) correspond to the square. In category B, the opposite role-bindings hold.

Perhaps naming *darker*, *larger*, *above* and *in front* somehow marks them as the "winning" roles, leaving *lighter*, *smaller*, *below* and *behind* to be the "losing" roles. If so, then to the extent that the effect is due to the involvement of the "winning" schema, per se, then having the roles consistent with categories (i.e., such that the "winning" shape is the one with the most named [i.e., "winning"] roles) ought to lead to faster learning than having the roles mixed across the "winning" and "losing" shapes (e.g., such that the "winning" shape that the one that has 3/4 of *larger* and *in front* [named, "winning" roles] and *lighter* and *below* [unnamed, "losing" roles]). By contrast, to the extent that the effect of "who's winning" simply reflects the role of comparison, then consistent vs. mixed role assignment should make little difference to the rate of learning. A third possibility, of course, is that both hypotheses are correct, in which case we would expect to see facilitatory effects of both comparison (i.e., "who's daxier?" vs. "what category?") and, in the case of "who's winning?" role assignment.

The results were consistent with both the hypothesized explanations of the effect of "who's winning" in Jung and Hummel (2009b). Participants in the *who's daxier* condition took reliably fewer trials to reach criterion than those in the *categorize* condition, but those in the *who's winning* condition took reliably fewer still. There was also a reliable difference of role assignment in the only *who's winning* condition: As expected, participants in the consistent conditions reached criterion faster than those in the mixed conditions.

The fact that *who's daxier* resulted in faster learning than *categorize* in both the consistent and mixed conditions is consistent with the hypothesis that *who's winning* (like *who's daxier*) encourages participants to compare the circle and square in a way that categorization does not. This hypothesis is further supported by the fact that participants in the *mixed winning* condition performed similarly to those in the *daxier* condition and better than those in the *categorize* condition. At the same time, the fact that participants in the *consistent winning* condition learned faster than those in either the *mixed winning* or *daxier* conditions is consistent with a winning-schema-specific effect. Together, Jung and Hummel (2009b)'s results suggest that an effective way to help people learn relational categories with a probabilistic structure is to recast the learning task in a form that encourages them to discover a higher-order relation that remains invariant over members of a category.

However, there are still at least two additional differences between *who's winning* and *who's daxier* that could account for the superior performance in the former condition: First, the difference between the question "who's winning?" is simply more meaningful than "who's daxier?"—a difference that could somehow have led to better performance in *who's winning*. Second, the two roles of the winning/losing relation have opposite valence. Perhaps it is something about relational roles with opposite valence, rather than winning per se, that encourages participants to invoke a schema that facilitates the discovery of an invariant higher-order relation with our stimuli.

In order to clarify the additional residual questions, the tasks “which one would Britney Spears like?” and “which one comes from Nebraska?” were added in addition to the previous tasks. The former task was assumed to encourage participants to think of the circle and square as separate objects, like *who’s winning* and *who’s daxier*. And like *who’s winning*, but unlike *who’s daxier*, its roles have opposite valence (presumably it is “good” to be liked by Britney and bad not to be liked by her) and it has meaning. The latter task shares the comparative property of *winning*, *daxier* and *Britney* and it has semantic content, like *winning* and *Britney*, but presumably lacks strong differences in valence across its roles (i.e., it is presumably neither particularly good nor particularly bad to be from Nebraska).

When Jung and Hummel (2009b) crossed the five learning conditions orthogonally with consistent vs. mixed role assignment, they found that *which one would Britney Spears like* is equivalent to *who’s winning*, and Nebraska is equivalent to daxier. That is, like the winning and Britney, the learning conditions separating circle from square, having roles with opposite valence, and having semantic content may be more likely to guarantee discovering a higher-order invariant and thus facilitate learning. Missing of any of these elements, however, seems to have a detrimental influence on category learning with a probabilistic structure.

In conclusion, Jung and Hummel (2009a, 2009b)’s findings are consistent with the hypothesis that learning relational categories is greatly facilitated by the discovery of an abstract invariant that holds true across all members of a category. As such, the data support the idea that relational category learning may entail a process of schema induction by intersection discovery in the mind of the learner.

1.6. Overview of experimental approach

While Jung and Hummel's (2009a, b) findings replicated and extended the findings of Kittur et al., providing additional evidence that relational categories are learned by a process of schema induction, and that this algorithm makes it very difficult for people to learn relational categories with a probabilistic structure, we still know very little about the mechanism underlying relational category learning. My motivation in the current experiments was to improve our understanding of relational category learning by conducting further tests of the intersection discovery hypothesis and by examining other factors (such as the presence of a dual task during learning) that may affect relational concept acquisition.

The experiments described in Chapter 2 explored factors that might plausibly make probabilistic relational categories learnable and provided another test of the schema induction hypothesis. The experiments described in Chapter 3 relation-based categories by contrasting feature- and relation-based category learning using a dual-task paradigm. Finally, Chapter 4 and 5 provides general discussion, and conclusion.

CHAPTER 2:

EXPERIMENT 1—TESTING THE INTERMEDIATE ENCODING HYPOTHESIS

In both the Kittur et al. and Jung and Hummel studies, about half the participants in the categorize conditions never learned to criterion. But this result implies that about half eventually did learn to criterion (albeit much more slowly than the participants in the "who's winning?" condition). On the strictest interpretation of the intersection discovery hypothesis, this ought to be impossible (i.e., the intersection is always the empty set, so the categories should never be learnable by anyone). This result raises the question: How do those participants who learn the categories manage to do so? My motivation in the first experiment was to test the hypothesis that those participants who do eventually learn to criterion manage to do so by learning subordinate-level sub-categories (within which one or two relations do remain invariant), and then learning to classify those sub-categories with a common label (as elaborated shortly).

As noted previously, according to the intersection discovery hypothesis, the reason probabilistic relational categories are difficult to learn is that intersection discovery is invoked with relational concepts (as opposed to simple associative learning, which is invoked by featural concepts). The intersection discovery process leads to a more general representation of a set of exemplars (e.g., a schema) by deemphasizing (or removing entirely) features or relations that are unique to one exemplar or another (Gick & Holyoak, 1983; Hummel & Holyoak, 2003; Kittur et al, 2004, 2006b).

Intersection discovery is useful because it acts to reveal relational generalities that might otherwise remain implicit in the mental representation of the individual exemplars (Doumas et al., 2008). However, it fails catastrophically with probabilistic categories, in which the intersection is the empty set. In the first experiment, I tested whether the comparison process underlying intersection discovery (i.e., structural alignment, aka, analogical mapping; see Gick & Holyoak, 1983) can be manipulated to enhance discovering an invariant relation so that the empty set problem can be avoided.

The empty set problem emerges with probabilistic categories when every exemplar of a category is compared (directly or indirectly) with every other exemplar. In the first experiment, I manipulated stimulus presentation so that each exemplar has a specific counterpart for comparison. For example, exemplar [0, 1, 1, 1] was always paired with exemplar [1, 0, 1, 1] and exemplar [1, 1, 0, 1] was always paired with exemplar [1, 1, 1, 0]. Such consistent pairings would give participants the opportunity to learn at least two invariant relations between each pair of exemplars. For example, comparing [0, 1, 1, 1] with [1, 0, 1, 1] leaves the third and fourth relations invariant. Accordingly, I hypothesized that the process of selective comparison should prevent the empty set problem by encouraging participants to learn subcategories of the nominal categories.

An additional purpose of the current experiment was to replicate the basic difficulty-of-probabilistic-relational-category learning effect with new stimulus materials. Kittur et al. (2004, 2006b) used stimuli composed of octagons and squares, and Jung and Hummel (2009a, b) used stimuli composed of circles and squares. The current experiment used fictional "bugs" as stimuli (Figure 1). The prototype of the category "Fea" [1, 1, 1, 1] had a head *wider* and *darker* than its body (relations r1 and r2; the first two 1's in the vector), antennae *longer* than its head (r3) and wings *longer* than its body (r4). The prototypical Dav [0, 0, 0, 0] had the opposite relations, with its body *wider* and *darker* than its head (r1 and r2), antennae *shorter* than its head (r3) and wings *shorter* than its body (r4).

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 were isomorphic with those used by Kittur et al (2004, 2006b) and Jung and Hummel (2009a, b).

Participants were assigned to one of three learning conditions: The *subordinate-level* condition, the *intermediate encoding* condition, and the *basic baseline* condition. In the *subordinate-level* condition, on each trial participants were presented with two stimuli simultaneously. The first task was to classify the two different species at the basic-level: Fea or Dav. The second task was to reclassify the same two species at the subordinate level: Kei Fea or Cim Fea (Figure 2). For example, paired exemplars [0, 1, 1, 1] and [1, 0, 1, 1] corresponded to “Kei Fea”, and paired exemplars [1, 1, 0, 1] and [1, 1, 1, 0] corresponded to “Cim Fea”. In the *intermediate encoding* condition, participants saw the same pairs as those in the *subordinate-level* condition but did not classify them at the subordinate level (Figure 3). In the *basis baseline* condition, participants saw only one exemplar at a time and classified it at the basic level (Figure 4).

The key to encoding at the *subordinate-level* is that the comparison is designed to help participants to discover the relations shared by the mapped exemplars ([− ,− , 1, 1] for Cim Fea and [1, 1, −, −] for Kai Fea, and [−, −, 0, 0] for Sko Dav and [0, 0, −, −] for Lif Dav). That is, even though, as a whole, the basic level category Fea has no invariant relations, its subcategories, Cim and Kai, do. If participants can learn the subcategories (i.e., by virtue of the invariants they contain), then perhaps this learning can help to bootstrap their learning of the basic level (even though no invariants exist at that level). The *intermediate encoding* condition presents participants with exemplars in the same pairs as the *subordinate-level* condition but did not require them to label the exemplars at the subordinate level. This condition was included to test the role of mere exposure to exemplars that share invariants.

In the *basic baseline* condition, each trial presented a single bug on the screen and the participant's task was to classify it at the basic (Fea or Dav) level only. This condition served as the closest replication of the category learning conditions used by Kittur et al. (2004, 2006b) and Jung and Hummel (2009a, b), in which one single object was provided for the categorization task.

Method

Participants. A total of 44 participants participated in the study for course credit. Participants were randomly assigned to one of three conditions.

Materials. Stimuli were line drawings of fictional bugs. The bugs varied in the size and darkness of their heads, the length, width and darkness of their bodies, and the length of their antennae. The prototype of category "Fea" was defined as [1, 1, 1, 1], and the prototype of "Dav" was defined as [0, 0, 0, 0], where the particular value on each dimension (1 or 0) defined the value of a relation. The prototype [1, 1, 1, 1] represented head *larger than* body, head *darker than* body, antennae *longer than* head, and wings *longer than* body, and [0, 0, 0, 0] represented head *shorter than* body, head *lighter than* body, antennae *shorter than* head, and wings *shorter than* body. Each category (species) consisted of one prototype (basic level species) and four exemplars. Subspecies of each species were made by grouping pairs of exemplars according to shared relations: Kei Fea

= [0, 1, 1, 1] and [1, 0, 1, 1,], and Cim Fea = [1, 1, 0, 1] and [1, 1, 1, 0]; Sko Dav = [1, 0, 0, 0] and [0, 1, 0, 0], and Lif Dav = [0, 0, 1, 0], and [0, 0, 0, 1]. 8 trials per block were presented in the *subordinate-level* and *intermediate encoding* conditions, and 16 trials per block were presented in the *basic baseline* condition. Each exemplar was presented in a random order once per block.

Design. The experiment used a 3 condition (*subordinate-level vs. intermediate encoding* vs. *basic baseline)* between-subjects design.

Procedure. All conditions consisted of two or more blocks of training trials followed by two blocks of transfer trials. The training phase of the experiment differed across conditions, as described above. During this phase of the experiment, participants received accuracy feedback on each response made on each trial.

In the *subordinate-level* condition, each trial of the training phase simultaneously presented two exemplars. Participants identified two bug stimuli at the basic level by clicking on boxes under the two bugs. The response was followed by accuracy feedback. And then they re-identified the same species at the subordinate level. In the *intermediate encoding* condition, participants were given only the basic-level identification task. In the control condition, bugs were presented one at a time in the center of the screen, asking participants to identify one bug at the basic level.

The transfer phase was the same across all conditions. All participants classified the bugs at the basic level only and they received no accuracy feedback. 16 trials were presented per block, with each exemplar presented in a random order once per block. Each exemplar remained on the screen until the participant responded. The training phase lasted for 40 blocks (320 trials for *subordinate-level* and *intermediate encoding*, and 640 trials for *basic baseline*) or until the participant responded correctly on at least fourteen of sixteen trials (87.5% correct) for two consecutive blocks across all conditions. At the end of the experiment participants were queried about strategies they used during the experiment.

Predictions

Consider the possible effects the comparison process may have on category learning at basic and subordinate levels. Labeling the exemplars at the subordinate level may lead participants to appreciate that, although there is no invariant at the basic level, exemplars at the same subordinate level do share invariants. As such, if these invariants can guide learning at the subordinate level, then perhaps participants' mastery of the subordinate level categories can help to bootstrap their discovery of the basic level categories. To the extent that simple comparison of systematic pairs of exemplars—as participants will do in both the *subordinate* level and *intermediate encoding* conditions—is sufficient for invariant discovery, then participants in both these conditions ought to learn faster and/or to a higher criterion than those in the basis baseline condition.

Results

Study phase: Accuracy. First, I report accuracy on the basic-level (Fea vs. Dav) in each condition. A 3 *(subordinate-level* vs. *intermediate encoding* vs. *basic baseline*) between-subjects ANOVA revealed main effects of task [F (2, 41) = 5.103, MSE = 0.007, $p <$ 0.05] (Figure 5). As expected, *subordinate-level* learners (M = .70, SD = .10) were likely to perform more accurately than *basic baseline* learners [Tukey's HSD, $p < 0.05$)]. *Intermediate encoding* learners (M = .71, SD = .07) were also likely to perform more accurately than *basic baseline* learners (M = .62, SD = .08) [Tukey's HSD, $p < 0.05$)].

Performance in the *subordinate-level* condition was almost identical to the performance in the *intermediate encoding* condition. (Tukey's HSD, $p = 0.99$). Performance at the subordinate-level task (Kei Fea vs. Cim Fea) is reported only in the *subordinate-level* condition ($M = .62$, $SD = .11$). Performance at the subordinate-level task was exactly identical to in the *basic baseline* condition ($t(27) = -0.089$, $p = 0.93$), implying that identifying at the subordinate-level was not as helpful as identifying in the control condition.

Study phase: Trials to criterion. I also report the rate at which participants learned the categories in terms of trials to criterion. All participants in the *subordinate-level* and *intermediate encoding* conditions reached criterion, whereas only 50% of participants in the *basic baseline* condition reached criterion. A 3 (*subordinate-level* vs. *intermediate encoding* vs. *basic baseline*) between-subjects design ANOVA revealed a main effect of task [F (2, 41) = 78.511, $MSE = 12482.691$, $p < 0.001$]. Participants in the *subordinate-level* condition ($M = 102$, $SD = 64$) took reliably fewer trials to reach criterion than those in the *basic baseline* condition (by Tukey's HSD, $p < 0.001$). Participants also reached criterion in fewer trials in *intermediate encoding* condition ($M = 82$, $SD = 52$) than in the *basic baseline condition* ($M = 537$, $SD = 173$) (by Tukey's HSD, $p < 0.001$) (Figure 6).

Transfer phase: Accuracy. My primary interest was accuracy on the transfer phase. A 3 (*subordinate-level* vs. *intermediate encoding* vs. *basic baseline*) between-subjects ANOVA revealed main effects of task [F (2, 41) = 9.298, $MSE = 0.008$, $p < 0.001$] (Figure 7). Participants in the *intermediate encoding* condition ($M = .79$, $SD = .06$) showed reliably more accurate performance than participants in the *subordinate-level* condition ($M = 0.71$, $SD = 0.13$) [Tukey's HSD, $p < 0.05$)] as well as in the *basic* *baseline* condition (M = 0.65, SD = 0.09) [Tukey's HSD, $p < 0.001$)]. There was no reliable difference between *subordinate-level* and *basic baseline* (Tukey's HSD, p = 0.32).

I also analyzed differences between study and transfer phase to examine how much learners improved in each condition (Figure 8). For learners in the *subordinate-level* condition, performance on the transfer trials (M = .71, SD = .13) was not reliably different from performance on the basic-level study trials (M = .70, SD = .10) [$t(13)$ = 0.194, p = 0.85), whereas performance at the subordinate-level (M = .62, SD = .11) was reliably different from performance on transfer [$t(13)$ = 2.396, $p < 0.05$], suggesting that accuracy at the subordinate-level reliably decreased than accuracy at the basic-level ($t(13)$ = -2.912, $p < 0.05$). For learners in the *intermediate encoding* condition, performance on the transfer trials (M = .79, SD = .06) reliably improved than mean performance on the study trials (M = .71, SD = .07) [$t(14)$ = 4.91, $p < 0.001$). For learners in *basic baseline*, there was no reliable difference between the study (M = .62, SD = .08) and transfer trials (M = 0.65, SD = 0.09) [$t(14)$ = 1.93, p = 0.074). Because only 50% of participants in *basic baseline* reached criterion, which means the rest of the participants were given the transfer trials under the circumstance they did not figure out the category learning rule, it seems that there was no reliable difference between the study and transfer trials.

Discussion

Previous research reported that participants have great difficulty learning relational categories with probabilistic structures (Kittur et al., 2004, 2006b, Jung & Hummel, 2009a, b, 2011). The difficulty was interpreted in terms of participants' 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 are one or more features or relations shared by category members. Experiment 1 examined under what condition the relations shared by category members can be retained. I took as a starting point the way in which an exemplar is compared with other exemplar. I hypothesized that if an exemplar of relational concepts has a counterpart to compare, one or two relations can remain invariant.

The results of Experiment 1 showed that simple comparison of systematic pairs of exemplars was enough to improve participants' ability to learn probabilistic relational categories. For participants in the *intermediate encoding* condition, the comparison task at the basic level helped them to abstract shared relations between the exemplars ($11--$ and $--11$ for Fea, $00--$ and $--00$ for Dav).

Performance in the *subordinate-level* condition was reliably less accurate than in the *intermediate encoding* condition, and there was no difference between the *subordinate-level* condition and *basic baseline* condition. The basic level performance in the *subordinate-level* condition was almost identical to the *intermediate encoding* condition, but there was a reliable difference between two conditions during transfer, suggesting that the subordinate classification task would hurt category learning in some way. A post-hoc analysis of participants' end-of-experiment self-reports revealed that participants' decisions were partially based on the commonality defined in terms of numerical values such as head width, wing length, etc., rather than relative values (such as head wider than body). Participants seemed to speculate that the exemplars with the same numerical size, darkness, or length belong to the same subspecies. In making the stimuli, exemplars of each category were constructed by varying the metric properties *size*, *darkness* and *length*, respecting the categorical relations *larger*, *darker*, and *longer*. On this account randomly generated values for each exemplar might have been the same from time to time. For example, when heads with *larger* and *darker* relations across the exemplars define the Fea species, two exemplars may have the same head size (or head darkness) across the exemplars. That is, it seems that the query of the subspecies led participants to focus more on featural aspects, resulting in the impaired learning. Such tendency toward paying more attention to featural properties may have continued to influence transfer.

Clearly important for my current purpose was the fact that, as predicted, comparing the exemplars in a systematic way improved the participants' ability to discover invariants so that the empty set could be rendered as non-empty. However, it still remains unclear how the task at the subordinate-level would hurt performance during transfer. Perhaps the order of encoding—basic-level encoding was followed by subordinate-level encoding—could be a potential reason for learning decline in Experiment 1a. The following pilot study tested the speculation.

CHAPTER 3: PILOT STUDY

Experiment 1a showed that subordinate-level encoding was no more effective during training, and less effective during transfer, than intermediate encoding. One possible explanation for this finding, which is broadly consistent with the logic of Experiment 1a, is that in that in the subordinate-level encoding condition of experiment, I had participants perform the basic-level classification before they performed the subordinate-level classification. But if subordinate-level classification is to serve as an aid to basic-level classification, then it is reasonable to expect that it ought to temporally precede that basic-level classification. Experiment 1b was designed to test this hypothesis: For example, suppose that participants learned Kei Fea is defined by the invariant (− −11), and Cim Fea is defined by the invariant (11− −). When participants judge a higher-level species, Fea or Dav, they could associate the invariant (− −11) with Fea or also the other invariant (11− −) with Fea. Such simple association would be able to make participants to learn probabilistic relational categories in a more simple way than the way the basic task is followed by the subordinate task.

The procedure of the pilot study was exactly identical to Experiment 1 except that participants were first given the subordinate-level classification task, and then the basic-level classification task.

Method

Participants. 12 participants participated in the study for course credit.

Materials and procedure. The same bug stimuli as Experiment 1 were used. In the pilot study, the subordinate labels did not include the basic names (i.e., Kei and Cim), unless they would already tell the answer for the following basic level task (Figure 9). The procedure for the pilot study was opposite to the Experiment 1a. The subordinate-task was followed by the basic-task.

Results

The pilot study tested only the subordinate-level condition to roughly measure how much performance could improve when the task's order was flipped. I compared performance in the pilot study with performance in the subordinate-level condition of Experiment 1a.

Study phase: Accuracy. Performance at the subordinate-level (Kei vs. Cim) in the pilot study (M = .65, SD = .09) was not reliably different from the subordinate-level in Experiment 1a (M = .62, SD = .11) [$t(24) = -0.832$, $p = 0.413$]. There was reliable difference between the basic-level tasks in two experiments. Accuracy in the pilot study (M = .77, SD = .04) was reliably different from accuracy in Experiment 1a (M = .70, SD = .10) [$t(24) = -2.134$, $p < 0.05$] (Figure 10).

Trials to criterion. As in Experiment 1a, all participants in the pilot study reached criterion, and there was no difference between trials-to-criterion in the pilot study (M = 94, SD = 78) and Experiment 1a (M = 102, SD = 64) [$t(24) = 0.307$, $p = 0.761$].

Transfer phase: Accuracy. Transfer performance in the pilot study (M = .79, SD = .05) was marginally different from performance in Experiment 1a (M = .71, SD = .13) [$t(24) = -1.935$, $p = 0.065$). In the pilot study, performance on the transfer trials (M = .79, SD = .05) was not reliably different from performance on the basic-level study trials (M = .77, SD = .04) [$t(11) = -0.116$, $p = 0.29$), whereas performance at the subordinate-level study trials ($M = .62$, $SD = .11$) was reliably different from performance on transfer [$t(11) = -6.375$, $p < 0.001$].

Discussion

The pilot study was designed to examine why the subordinate-level condition of Experiment 1a was less helpful than expected. I speculated that as the aid to basic-level classification, subordinate level classification should have preceded basic-level classification. In the pilot study, the task order was flipped: The basic task followed the subordinate task. When compared to the subordinate-level results in Experiment 1a, fairly reliable improvement was observed in the pilot study. The effects seem to benefit from associative learning: Associating each subordinate having two invariants with the basic. The process of associating is schematized in Figure 11. Relational learning, with invariants available, drives the subordinate task; then associative learning is all that is required to learn which subordinate-level categories belong together as members of the same basic-level category.

CHAPTER 4:

EXPERIMENT 2—TESTING THE PROTOTYPE COMPARISON HYPOTHESIS

In the next experiment, I tested two key manipulations to find another way to learn probabilistic relational categories. First, as a more direct way to circumvent the empty set problem, participants were trained to compare each exemplar with a prototype. My hypothesis was that comparing the exemplars to the prototype can help participants learn to categorize the bug stimuli in terms of prototype-plus-exception rules. For example, mapping the prototype [1, 1, 1, 1] to the exemplar [1, 0, 1, 1] will result in a schema that includes r1, r2 and r4, but lacks r2 (i.e., [1, −, 1, 1]). Whichever exemplar is compared to the prototype, the resulting schema will always produce one of the probabilistic category structures, minus the mismatching relation (i.e., [−, 1, 1, 1], [1, −, 1, 1], [1, 1, −, 1] or [1, 1, 1, −]). The prototype-plus-exception rules could potentially reveal three invariant relations per exemplar (although the invariants will not be constant across exemplars within a category).

Participants in the *prototype* condition were provided a prototypical member of each species and were asked to classify it using a basic-level label (e.g., “Fea” or “Dav”) and then they reclassified the exemplar species using a subordinate-level label: Kei Fea, Bai Fea, Wou Fea, or Cim Fea. I hypothesized that the basic label and a set of subordinate labels might provide participants with an explicit hierarchical structure could facilitate learning the category structure. If so, then each subordinate label would be associated with the relational difference between that exemplar and the prototype of its category (the *exceptional* relation): Kei Fea (*smaller* head), Bai Fea (*lighter* head), Wou Fea (*shorter* antenna), and Cim Fea (*shorter* wing).

Following on the intermediate encoding condition of Experiment 1, which allowed participants to compare systematic pairs of category exemplars, this experiment also tested whether random pairing of exemplars might also facilitate learning of relational invariants. Randomly pairing exemplars would highlight two invariants on each trial, although across trials, the invariants so highlighted would be free to vary. The random pairing manipulation was tested in the *two different exemplars* condition. After identifying the exemplars at the basic-level, participants were asked to re-identify each exemplar at the subordinate-level.

The *two same exemplars* condition was designed to test the Experiment 1's results, in which participants' decision was partially influenced by the same metric properties (e.g., the same wing length) between the exemplars. I examined when the exemplars have the exactly identical relations, but also have different metric properties, whether different featural properties would affect category learning. For exemplar, the exemplars (e.g., 1101, head size 4 and 1101, head size 7) would be exactly identical in terms of the relational property, or would be different in terms of the featural property.

Two kinds of control condition were used: *Subordinate baseline* and *basic baseline*. In the *subordinate baseline* condition, one single bug was presented, followed by the basic-level and subordinate-level classification tasks. Participants in the *basic baseline* condition were given only the basic-level classification task.

Method

Participants. A total of 96 participants participated in the study for course credit.

Participants were randomly assigned to one of five conditions.

Materials. The same bug stimuli were used in this experiment as in Experiment 1. Unlike the previous experiment, in which two different exemplars were associated with one label, the prototype from each category and all exemplars were associated with all different labels in Experiment 2: For the Fea species [1, 1, 1, 1] was the prototype, Kei Fea = [0, 1, 1, 1], Bai Fea = [1, 0, 1, 1,], Wou Fea = [1, 1, 0, 1], and Cim Fea = [1, 1, 1, 0] served as the subordinates; For the Dav species [0, 0, 0, 0] was the prototype, Haw Dav = [1, 0, 0, 0], Ang Dav = [0, 1, 0, 0], Sko Dav = [0, 0, 1, 0], and Lif Dav = [0, 0, 0, 1] were the subordinates.

Design. The experiment used a 5-condition (*prototype* vs. t*wo different exemplars* vs. *two same exemplars* vs. *subordinate baseline* vs. *basic baseline)* between-subjects design.

Procedure. All conditions except *basic baseline* were provided two or more blocks of training trials consisting of basic and subordinate classification tasks (only the basic task was provided in *basic baseline*), followed by two blocks of transfer trials, as in the previous experiment. The training phase of the experiment differed across conditions, as described below. During this phase, participants received accuracy feedback on each trial. The transfer phase was the same across all conditions. Participants classified the bugs at the basic level only and they received no accuracy feedback.

Participants were assigned to one of five training conditions: *prototype*, *two different exemplars*, *two same exemplars*, *subordinate baseline*, and *basic baseline.* In the *prototype* condition, participants were shown one bug that belongs to the prototype on the left side on the screen. They first decided whether the prototypical bug belongs to Fea or Dav by clicking the name and then a new exemplar appeared in the right side (the prototypical species remained on the screen). They decided the corresponding exemplar's name as the second task (Figure 12).

In the *two different exemplars* condition, two different exemplars belonging to the same species, randomly chosen, were provided simultaneously. Participants first decided the basic level category to which two exemplars belong and then decided each exemplar's name by clicking the mouse. For example, when two exemplars [1, 1, 0, 1] and [1, 0, 1, 1] are matched, participants should classify them as the Fea species at the basic level and then re-classify the exemplar [1, 1, 0, 1] as Wou Fea, and [1, 0, 1, 1] as Bai Fea at the subordinate-level (Figure 13).

The *two same exemplars* condition was identical to the *two different exemplars* condition, except that two exemplars were defined by exactly the same relations, but differing in their metric properties. For example, below two bugs have the exactly same relations: Head *larger* than body, and head *lighter* than body, antennae *longer* than head, and wings *longer* than body. The only difference between two bugs is the metric darkness of the head (Figure 14).

In the *subordinate baseline* condition, the participant classified one bug per trial at both the basic and subordinate levels. In the *basic baseline* condition participants classified each bug at the basic level only (Figure 15).

During training, in *prototype*, *two different exemplars, and two same exemplars*, 8 trials were presented per block, in *subordinate baseline*, and *basic baseline*, 16 trials were presented per block. In all conditions, the transfer phase was identical to the learning phase of the *basic baseline* condition. 16 trials were presented per block, with each exemplar presented in a random order once per block. Each exemplar remained on the screen until the participant responded. The training phase lasted for 40 blocks (320 trials for *prototype*, *two different exemplars, and two same exemplars,* and 640 trials for *subordinate baseline,* and *basic baseline*) or until the participant responded correctly on at least fourteen of sixteen trials (87.5% correct) for two consecutive blocks. At the end of the experiment participants were queried about strategies they use during the experiment.

Predictions

My main interest was in accuracy on the transfer phase. I predicted that the *prototype* condition will show better performance than the other conditions. I assumed that comparing the prototype with the exemplar can tell participants which relations are identical to each other and which relation is different between two bugs. In any case, three shared relations might be learned per exemplar.

I also predicted that participants who will be shown simultaneously two exemplars in the training phase (i.e., *two different exemplars* and *two same exemplars*) would face more difficult challenges compared to the *prototype* condition. In *two different exemplars*, randomly pairing exemplars would highlight two invariants on each trial. If the participants can appreciate two invariants, then their performance would be as helpful as *intermediate encoding* of Experiment 1, but less than the *prototype* condition. In *two same exemplars*, exactly the same exemplars (in terms of relations) will be provided, which means this condition may equal the control condition where a single bug will be presented, as long as the participants will focus more on relational properties than featural ones.

The *subordinate*-baseline and *basic*-baseline conditions will serve as control conditions that do not explicitly involve comparison.

Results

Study phase: Accuracy on basic-level. A 5 *(prototype* vs. t*wo different exemplars* vs. *two same exemplars* vs. *subordinate baseline* vs. *basic baseline*) between-subjects ANOVA revealed main effects of task [F (4, 91) = 45.518, MSE = 0.008, $p < 0.001$] (Figure 16). Participants in the *prototype* condition performed more accurately than those in all other conditions. *Prototype* learners (M = 0.92, SD = 0.06) were likely to perform more accurately than *two different exemplars* learners (M = 0.66, SD = 0.11) [Tukey's HSD, $p < 0.001$)], *two same exemplars* learners (M = 0.63, SD = 0.09) [Tukey's HSD, $p < 0.001$)], *subordinate baseline* learners (M = 0.60, SD = 0.08) [Tukey's HSD, $p < 0.001$)], and *basic baseline* learners (M = 0.64, SD = 0.08) [Tukey's HSD, $p < 0.001$)]. There were no other reliable differences between the conditions at the basic level during the study phase.

Study phase: Accuracy on subordinate-level. A 4 *(prototype* vs. t*wo different exemplars* vs. *two same exemplars* vs. *subordinate baseline*) between-subjects ANOVA revealed a main effect of task [F (3, 73) = 5.889, MSE = 0.015, $p < 0.01$] (Figure 17). Participants in the *prototype* condition showed more accurate performance than participants in all other conditions. *Prototype* learners (M = 0.46, SD = 0.16) were likely to perform more accurately than *two different exemplars* learners (M = 0.35, SD = 0.12) [Tukey's HSD, $p < 0.05$)], *two same exemplars* learners (M = 0.31, SD = 0.48) [Tukey's HSD, $p < 0.01$)], and *subordinate baseline* learners (M = 0.33, SD = 0.13) [Tukey's HSD,

$p < 0.01$)]. There were no other reliable differences between the conditions at the subordinate level during the study phase.

Study phase: Basic-level vs. subordinate-level. I report the difference between the basic-level and subordinate-level in each condition except for *basic baseline* to examine how much learning at different category levels were different. For the participants in *prototype* condition, performance at the basic-level ($M = 0.92$, $SD = 0.07$) was reliably different from one at the subordinate-level ($M = 0.46$, $SD = 0.16$) [$t(19) = 13.176$, $p < 0.001$]. Basic-level ($M = 0.66$, $SD = 0.11$) in *two different exemplars* was reliably different from subordinate-level ($M = 0.35$, $SD = 0.12$) [$t(19) = 11.357$, $p < 0.001$] Performance at basic-level ($M = 0.63$, $SD = 0.09$) in *two same exemplars* was reliably different from subordinate-level ($M = 0.31$, $SD = 0.05$) [$t(17) = 13.880$, $p < 0.001$]. Basic-level ($M = 0.60$, $SD = 0.09$) in *subordinate baseline* was reliably different from subordinate-level ($M = 0.33$, $SD = 0.13$) [$t(18) = 10.704$, $p < 0.001$] (Figure 17).

Study phase: Trials to criterion. I also report how many trials participants needed to reach to criterion during the study phase in terms of the basic-level. Only in the *prototype* condition did all participants reach criterion. A 5 (*prototype* vs. t*wo different exemplars* vs. *two same exemplars* vs. *subordinate baseline* vs. *basic baseline*) between-subjects design ANOVA revealed a main effect of task [$F(4, 91) = 107.139$, $MSE = 8776.459$, $p < 0.001$] (Figure 18).

As expected, participants reached criterion in fewer trials in *prototype* ($M = 36$, $SD = 27$) than in *two different exemplars* ($M = 239$, $SD = 89$) (Tukey's HSD, $p < 0.001$), in *two same exemplars* ($M = 285$, $SD = 69$) (Tukey's HSD, $p < 0.001$), in *subordinate baseline* ($M = 592$, $SD = 113$) (Tukey's HSD, $p < 0.001$), and in *basic baseline* ($M = 496$, *SD* = 133) (Tukey's HSD, $p < 0.001$). Participants given the *two different exemplars* task (*M* = 239, *SD* = 89) took reliably fewer trials to reach criterion than those in the *subordinate baseline* (*M* = 592, *SD* = 113) (Tukey's HSD, $p < 0.001$) and also those in the *basic baseline* task (*M* = 496, *SD* = 133) (Tukey's HSD, $p < 0.001$). Participants given the *two same exemplars* task (*M* = 285, *SD* = 69) also took reliably fewer trials to reach criterion than those in the *subordinate baseline* (*M* = 592, *SD* = 113) (Tukey's HSD, $p < 0.001$) and also those in the *basic baseline* task (*M* = 496, *SD* = 133) (Tukey's HSD, $p < 0.001$). Participants given the *basic baseline* task (*M* = 592, *SD* = 113) took reliably fewer trials to reach criterion than those in the *subordinate baseline* (*M* = 496, *SD* = 133) (Tukey's HSD, $p < 0.05$).

Transfer phase: Accuracy. My primary interest was accuracy on transfer phase. A 5 *(prototype* vs. t*wo different exemplars* vs. *two same exemplars* vs. *subordinate baseline* vs. *basic baseline*) between-subjects ANOVA revealed a main effect of task [F (4, 91) = 5.943, *MSE* = 0.011, $p < 0.001$] (Figure 19). Participants in *prototype* (*M* = .80, *SD* = .07) showed reliably more accurate performance than in *two different exemplars* (*M* = 0.69, *SD* = 0.10) (Tukey's HSD, $p < 0.01$), in *two same exemplars* (*M* = .68, *SD* = .11) (Tukey's HSD, $p < 0.01$), in *subordinate baseline* (*M* = .65, *SD* = .16) [Tukey's HSD, $p < 0.01$)] and in *basic baseline* (*M* = .68, *SD* = .07) (Tukey's HSD, $p < 0.01$).

As in Experiment 1a, I report differences between study (only basic-level) and transfer to examine how much participants learned in each condition (Figure 20). Surprisingly, for learners in *prototype*, performance decreased on transfer (*M* = .92, *SD* = .07) relative to study (*M* = .80, *SD* = .07) [$t(19) = 4.920$, $p < 0.001$]. For learners in *two different exemplars*, performance on transfer (*M* = .69, *SD* = .10) slightly increased from study to transfer (*M* = .66, *SD* = .11) [*t*(19) = 1.498, *p* = 0.151). For learners in *two same exemplars*, the difference between study (*M* = .63, *SD* = .09) and transfer (*M* = .68, *SD* = .11) was marginally reliable [*t*(17) = 1.865, *p* = 0.079). For learners in *subordinate baseline,* performance on transfer (*M* = .65, *SD* = .16) slightly increased relative to study (*M* = .60, *SD* = .08) [*t*(18) = 1.745, *p* = 0.098). For learners in *basic baseline*, the difference between study (*M* = .64, *SD* = .08) and transfer (*M* = .68, *SD* = .07) was not reliable [*t*(18) = 1.637, *p* = 0.119).

Discussion

Experiment 2 examined whether providing participants prototypes of the basic-level categories would facilitate their learning of the exemplars of those categories by helping them to learn the exemplars in a rule-plus-exception fashion. As expected, performance in the *prototype* condition exceeded performance in the other conditions. *Prototype* learners showed above 90 % correct in classifying the prototype, and 80% correct during transfer. Providing the prototypes appears to have helped participants to learn the exemplars. These results are consistent with the hypothesis that explicitly providing the prototype can help learners overcome the difficulties posed by the empty intersection problem.

Learning in a random pairing fashion did not help learners overcome the difficulties posed by the empty intersection problem: Neither the difference between *two different exemplars* and *two same exemplars*, nor even the difference between comparison conditions and single conditions was reliable. In contrast to Experiment 1 where the same pair was consistently compared, random parings seemed to fail to drive the discovery of invariants. According to a post-hoc analysis of self-reports, majority of participants mentioned that they were able to discover the relational similarities between the exemplars to some extent, but they little understood how to define each species. In conclusion, comparing random pairs was not as helpful as comparing specific pairs, and even was similar to the condition where a single exemplar was presented.

CHAPTER 5:

EXPERIMENT 3—TESTING THE EFFECTS OF DUAL VERBAL AND VISUAL DUAL TASKS ON FEATURAL VS. RELATIONAL CATEGORY LEARNING

The results from Experiment 1 and 2 added further evidence that the learning mechanism for relational categories with a probabilistic structure is the intersection discovery hypothesis, which is quite a contrast to associative learning for featural learning mechanism. Such evidence accordingly implies that relational and featural categories are learned in qualitatively different ways.

In Experiment 3, I investigated the sensitivity of feature- and relation-based category leaning to two different kinds of dual-task disruption. In contrast to feature-based representations, which come to us effortlessly, relational representations require attention and working memory (see, e.g., Baddeley & Hitch, 1974; Hummel & Holyoak, 1997, 2003; Logan, 1994; Maybery, Bain, & Halford, 1986).

To the extent that featural and relational category learning rely on different kinds of mental representations, they might be differentially disrupted by different kinds of dual tasks. In particular, it is reasonable to hypothesize that featural learning may be more disrupted by a dual task that consumes visual working memory resources than by one that consumes verbal or executive resources (inasmuch as featural category learning is largely a visual learning task), whereas relational category learning may be more disrupted by a dual task that consumes verbal or executive working memory (inasmuch as relational processing is an explicit, attention-demanding task).

Other researchers have also argued for multiple systems of category learning

(Ashby, Alfonso-Reese, Turken, & Waldron, 1998). Miles and Minda (2011) showed that verbal dual tasks, which impose an executive functioning load, impaired rule-defined category learning, whereas a visual dual task impaired non-rule-defined learning regardless of executive functioning demand. Their findings provided evidence that verbal working memory and executive functioning are engaged in the rule-defined system, and visual processing is more engaged in the non-rule-defined system.

My next experiment tested the prediction that relational category learning will be more subject to *verbal dual-task interference* than feature-based category learning. By contrast, feature-based learning will be more subject to *visuospatial dual-task interference* than relational learning.

I used deterministic category structures in the current experiment; i.e., there was always be one relation or feature that is deterministically predictive of category membership. The reason for using deterministic categories is that the categories must be learnable, even in the relational case, so that I can observe the effects of the manipulation on trials to criterion (i.e., how long it takes participants to learn the categories).

I orthogonally crossed relational- vs. feature-based categories with verbal dual task vs. visual dual task vs. no dual task. In the verbal dual task conditions, participants had to perform a task known to interfere with relational processing (memorizing digits) while they simultaneously performed the category learning task. In the visual dual task condition, participants had to memorize the locations of filled squares in 3 X 3 grids while simultaneously learning the categorization. In the no dual task condition, participants simply performed the category learning task by itself.

Method

Participants. A total of 75 participants participated in the study for course credit. Each participant was randomly assigned to one of the six conditions.

Materials. Each exemplar consisted of a grey ellipse and a grey rectangle. Each exemplar had both relational properties (e.g., ellipse bigger than rectangle) and featural properties (e.g., ellipse of size 4). Each participant was tasked with deciding whether the objects they saw belonged to one of two featural or one of two relational categories.

Each exemplar was defined by three category-relevant properties: size (absolute in the featural condition or relative in the relational condition), darkness (absolute or relative) and orientation (absolute or relative). In the featural condition, the orientation of the ellipse was deterministically associated with category membership (i.e., horizontal orientation for category A, vertical for category L), whereas in the relational category condition, the relative orientation of the ellipse and rectangle (i.e., either *same* or *different*) was deterministically associated with category membership (with *same* for category A and *different* for category L). The other properties were probabilistically associated with category membership.

For the featural category condition, the prototypes of the categories were defined as [1,1,1] for category A and [0,0,0] for L, where [1,1,1] represents an rectangle size 3 [out of 9] for category A, 7 for category L, the color 3 [out of 9] for category A, 7 for category L, and horizontal orientation for category A, vertical for category L (Figure 21). Similarly, for the relational category condition, the prototypes were defined as [1,1,1] for category A and [0,0,0] for L, where [1,1,1] represents an ellipse *larger*, *darker*, and *same orientation* and [0,0,0] represents a rectangle *larger*, *darker*, and *different orientation*

(Figure 22). Exemplars of each category were made by switching the value of one dimension in the prototype (e.g., relational category A exemplar [1,0,1] would have the ellipse *larger*, *lighter*, and *same* orientation as the rectangle). Four copies of each exemplar type were presented on each block, two paired with a "Yes" responses on the dual task and two with a "No" responses, resulting in 32 trials per category per block.

Design. The experiment used a 3 (dual task: *none* vs. *verbal* vs. *visuospatial*) X 2 (relevant property: *features* vs. *relations*) between-subjects design.

Procedure. Participants were assigned randomly to one of the six groups. For the dual task conditions, on each trial, a memory task was provided first and followed by a categorization task and by a recall task. In the control conditions, only the categorization task was provided. Both categorization and dual task responses were followed by accuracy feedback (Figure 23).

Participants in the verbal dual-task condition were first given a verbal working memory task, in which 5 random digits were displayed for two seconds with spaces between them (so that they appeared to be individual numbers rather than digits of a single number). Participants were asked to memorize the digits while they performed the categorization task. In the categorization task, an exemplar consisting of a rectangle and an ellipse was shown. Participants were instructed to press the A key if the stimulus belong to category A and the D key if it belong to D. Each exemplar remained on the screen until the participant responded. Responses were followed by accuracy feedback. Participants then saw one random digit and were asked to decide whether it was in the set they saw previously.

In the visuospatial dual-task condition, a 3 by 3 grid was displayed in the middle of a screen for two seconds with two randomly-chosen cells filled. Participants were asked to memorize the locations of the filled cells until they completed the categorization task. In the recall task, one filled cell was displayed in the grid and participants were asked whether the cell had been filled in the original display. The experiment was consist of 30 blocks (960 trials) and continue until the participant responded correctly on at least twenty nine of thirty two trials (90.6% correct) for two consecutive blocks or until all 30 blocks transpired, whichever comes first. At the end of the experiment participants were queried about strategies they use during the experiment.

Predictions

I hypothesize that the two dual tasks will interact with the kinds of category learning in different ways. The verbal dual task will interfere with relational category learning to a greater extent than featural category learning. In contrast, I hypothesize that the visuospatial dual task will interfere with featural category learning to a greater extent than relational category learning. Taken together, I predict the additional distinction between feature-and relation-based category learning via the double dissociation between visual vs. verbal dual task interference on the one hand and featural vs. relational category learning on the other.

Results

Dual task accuracy. I discarded the data from participants whose accuracy was below 70% correct on the dual task (2 participants in the verbal/featural condition). Mean accuracy on the verbal dual task was M = .94 (SD = .03) for the featural category learning condition, and $M = 0.91$ ($SD = 0.06$) for the relational learning condition. Mean accuracy on the visual dual task was $M = 0.91$ ($SD = 0.06$) for the featural condition, and $M = 0.89$ ($SD = 0.04$) for the relational condition. There was no reliable difference between the verbal and visuospatial tasks [$t(51) = 1.61$, $p = .114$], suggesting that these tasks occupied cognitive resources to roughly the same extent.

Category learning task accuracy: Trials to criterion. Since my primary interest is the rate at which participants learn the categories as a function of the dual tasks, I report the data first in terms of trials to criterion. These analyses are conservative in the sense that participants who never learned to criterion were treated as though they reached criterion on the last block. Figure 24 shows the mean trials to criterion by category learning condition. A 3 (dual task) × 2 (category learning task) between-subjects ANOVA revealed a main effect of dual task [$F(2, 69) = 5.058$, $MSE = 579014.858$, $p < 0.01$]. Since my main interest was in how different dual tasks affect the different kinds of category learning, one-way ANOVAs were conducted for the featural and relational learning conditions. The results revealed reliable differences between dual tasks in the featural category learning condition [$F(2,35) = 4.981$, $MSE = 617725.846$, $p < 0.05$]. Planned comparisons in the featural category learning showed that there was a reliable difference between the verbal ($M = 386$, $SD = 387$) and visuospatial dual task ($M = 697$, $SD = 411$) [$t(35) = -2.288$, $p < 0.05$]. There was also a reliable difference between the visuospatial and the control condition ($M = 262$, $SD = 191$) [$t(35) = 3.014$, $p < 0.01$]. The difference between the verbal and the control condition was not reliable [$t(35) = 0.877$, $p < 0.386$]. The ANOVA results from the relational condition revealed reliable differences between the dual tasks [$F(2,34) = 7.641$, $MSE = 799483.887$, $p < 0.01$]. Planned comparisons revealed that there was a reliable difference between the verbal (M = 739, SD = 352) and visuospatial dual task (M = 330, SD = 362) [$t(34) = 3.221$, $p < 0.01$]. There was also a reliable difference between the verbal and control conditions (M = 276, SD = 222) [$t(34) = 3.014$, $p < 0.01$]. The difference between the visuospatial and control conditions was not reliable [$t(34) = 0.404$, $p < 0.689$]. No other main effects were statistically reliable. Most interestingly, there was a reliable interaction between dual task and category learning, indicating that relational category learning was disrupted more by the verbal dual task, whereas featural category learning was disrupted more by the visuospatial dual task [$F(2,69) = 2.475$, MSE = 855659.946, $p < 0.01$].

Response times. Since the category learning accuracy results yielded a reliable interaction between the dual and category learning tasks, I also analyzed these tasks in terms of participants' mean response times on individual trials in order to gain insight about the strategies participants in each condition may have adopted. A 3 (dual task) × 2 (category learning task) between-subjects ANOVA revealed a main effect of dual task [$F(2, 69) = 3.202$, MSE = 0.961, $p < 0.05$]. One-way ANOVAs were also conducted in each category learning condition. The main effect of dual task was not reliable [$F(2, 35) = 2.137$, MSE = 0.612, $p = 0.133$] in the featual learning condition. But since the accuracy data showed that participants in visuospatial feature-learning required many more trials than to reach to the criterion than participants in verbal featural learning, I expected a reliable difference between two conditions in a planned comparison analysis. My prediction was confirmed. There was a reliable difference between the verbal (M = 0.99, SD = 0.31) and visuospatial dual task (M = 1.41, SD = 0.78) [$t(35) = -2.037$, $p < 0.05$], indicating that response times in visuospatial feature-learning condition were longer than those in verbal feature-learning. No other differences were statistically reliable. There were no reliable differences in the relational learning condition. Also, ANOVA showed a reliable main effect of category learning [$F(1, 69) = 3.883$, $MSE = 1.166$, $p = 0.053$], indicating that feature learning ($M = 1.17$, $SD = 0.55$) was marginally faster than relational learning ($M = 1.42$, $SD = 0.56$) (Figure 25).

Discussion

To the extent that relational concepts are qualitatively similar to feature-based concepts, our understanding of concepts can be expected to generalize from the (extensively investigated) case of feature-based categories to the (largely neglected) case of relational categories. However, there is reason to believe they are not, casting doubt on our ability to generalize our conclusions from studies using feature-based categories to the case of relational concepts.

Most notably, people have no difficulty learning feature-based categories in which no single feature remains invariant across all members of a category (see Murphy, 2002). By contrast, relational categories are extremely difficult to learn when there is no such relational invariant (Kittur et al., 2004, 2006b; Jung and Hummel 2009a, 2009b, 2011). These findings suggest that featural and relational learning rely not only on qualitatively different forms of mental representation (namely, features vs. relations; see, e.g., Hummel, 2010; Hummel & Holyoak, 1997, for a discussion of the difference) but also that they rely on qualitatively different kinds of learning algorithms (e.g., associative learning in the featural case and something more akin to structured intersection discovery in the relational case; Jung & Hummel, 2009a, 2009b, 2011).

The current experiment provides additional evidence for this sharp distinction between featural and relational category learning. In the current experiment, featural learning was impeded by a visual dual task (i.e., one that might be expected to interfere with visual feature processing as required for featural learning) but not by a verbal dual task. Relational category learning, in sharp contrast, was interfered with by a verbal dual task (which has been shown to interfere with relational processing; Waltz, Lau, Grewal, &, Holyoak, 2000), but not by a visual dual task. This double dissociation between visual vs. verbal dual task interference on the one hand and featural vs. relational category learning on the other adds to the growing evidence that these two kinds of category learning rely on qualitatively different and dissociable learning systems.

CHAPTER 6:

EXPERIMENT 4—TESTING THE EFFECTS OF A RELATION-CENTERED VISUAL DUAL TASK ON CATEGORY LEARNING

In the previous experiment, I predicted that the visual dual task would interfere with the featrual category learning more than the verbal dual task would. In the previous study, the visual dual-task was to memorize the locations of two shaded cells in a 3 by 3 grid, which could be characterized as a *feature-centered* visual dual-task. The following experiment investigated the effects of a more relational visual-dual task on featural and relational category learning.

My question was that to what extent a *relation-centered* visual dual-task interferes with category learning. Previously, I predicted that a visual dual-task would hinder feature-based category learning, since the resources—implicit and less attention-demanding—that are necessary to process featural categories are lately the same as those necessary to perform the visual dual task. By contrast, a relation-centered visual dual-task whose process can be characterized as more attention-demanding, may be more likely to interfere with relational category learning. Thus I will be able to observe the resilience of feature-based category learning under a *relation-centered* visual dual-task.

I orthogonally crossed relational- vs. feature-based categories with verbal vs. visual-relation dual task. The verbal dual task condition was identical to one used in the previous experiment. In the relation-centered visual dual task condition, a hexagon and an octagon with different sizes were provided simultaneously. Participants had to memorize the relative size of the two objects while performing each categorization trial.

Method

Participants. A total of 61 participants participated in the study for course credit. Each participant was randomly assigned to one of the four conditions.

Materials. The same ellipse and rectangle stimuli were used in this experiment as in the previous experiments.

Design. The experiment used a 2 (dual task: *verbal* vs. *relation-centered visual*) X 2 (relevant property: *features* vs. *relations*) between-subjects design.

Procedure. In the verbal dual task condition, the same procedure as the previous experiment was used. In the relation-centered visual dual-task condition, a hexagon and an octagon of different sizes were presented in a random location for two seconds. Participants were asked to memorize the relative sizes of the two objects until they completed the categorization task. In the recall task, a hexagon and an octagon were displayed in random locations of the screen and participants were asked whether the relative sizes of two objects were the same as ones in the original display. The experiment lasted for 30 blocks (960 trials) or until the participant responded correctly on at least twenty nine of thirty two trials (90.6% correct) for two consecutive blocks. At the end of the experiment participants were queried about strategies they used during the experiment.

Predictions

I predict that the relation-centered visual dual-task will influence category learning in the same manner as the verbal dual task. Specifically, the relational visual dual task, like the verbal dual task, will impair relational category learning to a greater extent than featural category learning. That is, in contrast with the previous experiment where more feature-characterized visual dual-task was used, more relation-characterized visual dual-task will interfere with relation-based category learning.

Results

Dual task accuracy. I discarded the data from participants whose accuracy was below 70% correct on the dual task (1 participant in the visual/featural condition, and 2 participants in the visual/relational condition). Mean accuracy on the verbal dual task was $M = 0.92$ ($SD = 0.09$) for the featural category learning condition, and $M = 0.92$ ($SD = 0.05$) for the relational learning condition. Mean accuracy on the visual dual task was $M = 0.86$ ($SD = 0.07$) for the featural condition, and $M = 0.84$ ($SD = 0.06$) for the relational condition. Accuracy on the visual dual task in Experiment 4 declined to some extent compared to the visual dual task in Experiment 3. There was reliable difference between the verbal and visual tasks [$t(59) = 3.862$, $p < 0.001$], suggesting that the visual dual task occupied more cognitive resources than the verbal dual task did. However, my main interest in Experiment 4 was that how the relation-centered visual dual-task would interact with featural and relational category learning, not how two different dual tasks would interact with separate category learning.

Category learning task accuracy: Trials to criterion. Since my primary interest is the rate at which participants learn the categories as a function of the dual tasks, I report the data first in terms of trials to criterion. These analyses are conservative in the sense that participants who never learned to criterion were treated as though they reached criterion on the last block. Figure 26 shows the mean trials to criterion by condition. A 2 (dual task) × 2 (category learning task) between-subjects ANOVA revealed that there was no reliable difference between verbal and visual dual tasks [$F(1, 57) = 0.442$, $MSE = 66109.223$, $p = 0.509$].

Since my main interest is in how different dual tasks affect the different kinds of category learning, I compared the accuracy on featural and relational category learning within each dual-task. As expected, the results revealed a reliable difference between the featural ($M = 329$, $SD = 395$) and relational conditions ($M = 657$, $SD = 390$) within the verbal dual- task [$t(28) = -2.292$, $p < 0.05$]. Interestingly, there was also a reliable difference between the featural ($M = 286$, $SD = 350$) and relational conditions ($M = 568$, $SD = 407$) within the visual dual-task [$t(29) = -2.062$, $p < 0.05$], indicating that relational category learning was disrupted more by the relation-centered visual dual task. The difference between the featural ($M = 307$, $SD = 368$) and relational conditions ($M = 611$, $SD = 395$) was reliable [$t(59) = -3.109$, $p < 0.01$]. No other main effects were statistically reliable.

Response times. I also analyzed these tasks in terms of participants' mean response times on individual. A 2 (dual task) × 2 (category learning task) between-subjects ANOVA revealed a main effect of dual-task [$F(1, 57) = 8.007$, $MSE = 2.669$, $p < 0.01$], indicating that response times in the visual dual-task ($M = 1.86$, $SD = 0.68$) were longer than those in the verbal dual-task ($M = 1.43$, $SD = 0.53$). Also, ANOVA showed a reliable main effect of category learning [$F(1, 57) = 9.829$, $MSE = 3.277$, $p < 0.01$], indicating that feature learning ($M = 1.41$, $SD = 0.17$) was faster than relational learning ($M = 1.87$, $SD = 0.04$). Within each dual-task, response times in the relational learning condition were reliably longer than the featural learning [for the verbal dual-task, $t(28) = -2.239$, $p < 0.05$, for the visual dual-task, $t(29) = -2.243$, $p < 0.05$] (Figure 27).

Discussion

Experiment 3 has made clear the importance of verbal working memory for relational categories, and visual working memory for feature-based categories. I was interested in further exploring the cognitive resources of the visual dual-task that included more relational information. In Experiment 4, my prediction was that if the visual dual-task is more involved in the use of relational resources, even if the dual-task is visually demanding, the relation-centered visual dual-task would interfere with learning relational categories. As predicted, the new visual dual-task more taxed relational categories than feature-based categories. The results suggest that the relation-centered visual dual-task would consume verbal-based cognitive resources (i.e., attention-demanding). Participants' self-reports also supported the results: Their strategy was that they made the sentence like "octagon is larger than hexagon" to memorize the relative size of two objects by themselves. The results, in conjunction with the results of Experiment 3, show that relational concepts rely heavily on the verbal system based on more attention-demanding processes, and feature-based concepts rely more on the non-verbal system based on more implicit processes, implying the importance of qualitatively different and separable learning systems to two kinds of category learning.

CHAPTER 7: GENERAL DISCUSSION

The findings presented in this Dissertation make a strong case that not only are relational concepts qualitatively different from featural concepts, they are also *learned* in a qualitatively different manner. In particular, the current results are consistent with the hypothesis that relational categories are learned by a kind of structured intersection discovery—a process that is formally powerful than feature-based associatively learning, but which fails catastrophically with probabilistic category structures.

However, as shown by the results of Experiments 1 and 2, whether a category structure is probabilistic—and thus whether concept acquisition must fail catastrophically—lies at least in part in the manner in which the learner approaches the learning task. Experiment 1, like Jung and Hummel (2009a, b, 2011), showed that learning can be improved by structuring the learning task to reveal within-category invariants. Jung and Hummel showed that unstated higher-order invariants (such as whether the “circle is winning”), can facilitate concept acquisition. Experiment 1 extended this result, showing that the invariants can come from the hierarchical structure of the categories themselves: If the to-be-learned concepts do not possess invariants at the nominal (i.e., “basic”) level of categorization, then it can be helpful to learn the categories first at a subordinate level of abstraction that does contain invariants (Experiment 1b). Having learned the subordinate-level categories, learning that multiple subordinate-level concepts belong to the same “basic-level” (i.e., named) concept can be accomplished in an associative fashion.

This finding may help to make sense of the fact that many, or at least some, natural relational concepts seem to have a probabilistic structure. For example, “mother” is clearly a relational concept, and yet it is difficult to come up with a single relational definition that encompasses all and only instances of “mother” (much like Wittgenstein’s, 1953, concept of “game”, as discussed shortly). On its face, this fact is troubling for the intersection discovery hypothesis, since that hypothesis predicts that there should be no probabilistic relational concepts. However, as proposed by Lakoff (1987), the resolution of this dilemma may lie in concepts such as mother being polysemous: We have multiple (individually deterministic) “mother” schemas unified under a single label. If this polysemy-based account of otherwise seemingly probabilistic natural relational concepts is correct, then the results of Experiment 1 (especially the results of Experiment 1b compared to the results of the *subordinate encoding* condition of Experiment 1a) suggest that we may learn the subordinate-level meanings of mother before we learn to attach the label “mother” to all the different concepts to which it can refer.

This idea of polysemy may even suggest a solution to Wittgenstein’s (1953) famous “game” dilemma: Although no one seems able to come up with a single definition that includes all games and excludes all non-games—and even though “game” would appear, on its face, to be a relational concept—perhaps our failure to come up with a satisfactory definition reflects the concept’s polysemy (see Lakoff, 1987). If this account is correct, then our understanding of “family resemblance” categories may be fundamentally incorrect, or at least incomplete: To say that a concept such as “mother” or “game” is a collection of polysemous relational concepts united under a single name is a very different claim about the mental representation of concepts than to say that “mother” and “game” are simply points in a high-dimensional feature (i.e., vector) space (i.e., a “prototype”, or the mean of a collection of points) and that any specific exemplar of mother or game is simply another point in that space that lies closer to or further from that prototype point.

This claim about the inadequacy of traditional feature-based model of concepts is not new. Numerous researchers have argued that relational concepts cannot be adequately represented as lists of features, but instead must be mentally represented as relational structures such as schemas, theories, or causal models (Gentner, 1983; Holland, Holyoak, Nisbett, & Thagard, 1986; Hummel & Biederman, 1992; Hummel & Holyoak, 1997, 2003; Keil, 1989; Murphy & Medin, 1985; Rehder & Burnett, 2005; Waldmann, Holyoak & Fratianne, 1995). This view is supported by evidence from studies of similarity, relational reasoning and attention suggesting that relations and features may be psychologically distinct (e.g., Barr & Caplan, 1987; Gentner, 1983; Gentner & Kurtz, 2005; Goldstone, 1996; Logan, 1994; Markman & Stilwell, 2001; Medin, Goldstone & Gentner, 1993). At the same time, however, the empirical findings demonstrating prototype effects—which are almost universally attributed to feature-based notions of conceptual structure—are both numerous and robust (see Murphy, 2002). It remains to be seen whether the notion of relational concepts, combined with ideas about the hierarchical structure of those concepts, can help to make sense of the kinds of effects that would otherwise lead one to conclude that concepts are but lists of features.

Understanding the roles of features and relations in conceptual structures is complicated by the fact that both undoubtedly play a role. The question is not whether features or relations serve as the basis of our concepts, but rather how they work together to structure and inform our understanding of the world. A reasonable conjecture, suggested by numerous findings across the study of both perception and cognition, is that the mind will use features whenever it can but can also reason about relations whenever it must. “Features” can be processed automatically, quickly and effortlessly (for reviews see Hummel, 2001; Thoma & Davidoff, 2007) but as reviewed previously, relational representations support substantially more sophisticated kinds of inference and generalization.

Consistent with this generalization, Experiments 3 and 4 showed that relational category learning imposes a greater working memory load than does featural category learning. In the former case it is necessary to actively compute relations and bind them to their arguments (Halford, Wilson, & Phillips, 1998; Hummel & Biederman, 1992; Hummel & Holyoak, 1997, 2003; Oberauer, Suß, Wilhelm, & Sander, 2007). In contrast, featural categories may incur reduced memory load, perhaps by relying on emergent perceptual features or implicit learning mechanisms (Ashby & Waldron, 1999; Ward & Becker, 1992).

Kittur et al. (2006b) showed that feature- and relation-based representations also seem to support qualitatively different kinds of judgments, with feature-based representations supporting familiarity judgments, while relation-based representations support “goodness of exemplar” judgments (at least with relationally-defined categories). Their findings suggest that these different kinds judgments are the exclusive purview of their respective kinds of representations: In their data, stimulus features drove familiarity judgments even when relational differences between stimuli were the only basis for distinguishing them (i.e., familiarity judgments were at chance when relations, but not features, discriminated between stimuli), and relational differences drove “goodness” judgments even when features provided the only basis for discrimination.

Taken together, these findings suggest a kind of double dissociation between the learning and use of featural versus relational representations, which may ultimately reflect differences between the representation and processing of implicit features on the one hand and explicit relational predicates on the other (see also Hummel, 2001, 2010; Thoma & Davidoff, 2007). The findings presented in this Dissertation contribute to the literature demonstrating that featural and relational representations are psychologically distinct.

The results of these studies also help to clarify the circumstances under which relational concepts, probabilistic and otherwise, maybe acquired. Comparison across exemplars is known to play an important role in the acquisition of relational categories (Gentner, Anggoro, & Klibanoff, 2011; Higgins & Ross, 2011; Kurtz, Boukrina, & Gentner, 2013). Much of the comparison research involves within-category pairs, between-category pairs, or mixed-category pairs, in which exemplars are paired randomly regardless of whether they belong to the same or different categories. The results presented here suggest that comparisons are likely to be beneficial specifically to the extent that they allow learners to discover relational invariants within categories (and possibly invariant contrasts between categories)—and that the role of such invariants is likely to be much greater for relational concepts (such as those that are crucial for math and science education) than for featural ones.

The current results also suggest that other manipulations used in studies of relational learning and reasoning may affect category learning and inference. For example, relational responding is decreased when response time is limited, or when the richness or featural complexity of the stimuli is increased (Markman & Gentner, 1993a).

In contrast, relational responding is increased when people are required to use multiple analogs (Catrambone & Holyoak, 1989; Gick & Holyoak, 1983), to perform comparisons (Gentner & Namy, 1999), or to provide multiple mappings for a single example (Markman & Gentner, 1993b). Further work is needed to explore how such manipulations may bias category learning to focus on either featural or relational information.

In summary, the findings presented in this Dissertation contribute to the growing body of theoretical and empirical results suggesting that relational thought—the kind of thinking that seems to separate us most sharply from our closest primate cousins—is a *qualitatively* different thing than the kinds of thinking and learning afforded by feature-based representations of the world. They also underscore the fact that the power of relational thought comes with attendant costs. Acquiring a relational concept demands your full attention (Experiments 3 and 4) and requires you to discover its invariant core (Experiments 1 and 2). If you fail on either score, the relational concept will elude your grasp.

CHAPTER 8: CONCLUSION

Relational concepts play a central role in human cognition, especially in the most uniquely cognitive faculties, such as scientific, mathematical and (abstract) causal reasoning, as well as language. However, theoretical considerations and previous empirical research suggest that the power of relational concepts comes with a cost: namely, that they are much more sensitive to the conditions of acquisition than are feature-based concepts.

The experiments described in this Dissertation investigated two factors that are predicted to systematically impair the acquisition of relational concepts relative to featural concepts. Experiments 1 and 2 replicated and extended previous findings (e.g., Jung & Hummel, 2009, 2011; Kittur et al., 2004, 2006a, b) demonstrating that, although featural concepts are easily acquired even when they have a probabilistic structure, acquisition of relational concepts fails catastrophically in the face of such structures; that is, relational concepts require at least one property to remain invariant over all exemplars in order to be learnable. Experiments 3 and 4 used a dual-task paradigm to demonstrate that relational and featural concepts are differentially affected relational and featural dual tasks, further supporting the conclusion that featural and relational concept acquisition rely on qualitatively different mental representations and/or learning algorithms.

8.1. Summary of Experiments 1 and 2

One of the most robust findings in the study of category learning is that people easily learn categories with a probabilistic structure in which no single feature is shared by all members of the category (see Murphy, 2002). However, all the experiments demonstrating this phenomenon have used categories defined in terms of their exemplars' features (see Kittur et al., 2004). To the extent that relational concepts are represented in a qualitatively different manner than feature-based concepts (e.g., as schemas, scripts or frames, rather than simple lists of features), laws of learning discovered using one kind of category may not generalize to concepts based on the other.

When Kittur at al. (2004, 2006a, b) explored whether the prototype effects so often observed with feature-based categories could also be observed with relational categories, they found that people have great difficulty learning relational categories with probabilistic structures. They interpreted this result in terms of peoples' 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; Gick & Holyoak, 1983; 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 will fail catastrophically if all features and relations are related only probabilistically to category membership. This *intersection discovery* account of relational learning predicts that a relational category will be learnable *if and only if there exists at least one relation that is present in all exemplars of the category*. That is, relational categories are learnable iff the intersection of the exemplars in not the empty set. The corollary of this prediction is that all *and only* those manipulations that render the intersection non-empty will render the category learnable.

Jung and Hummel (2009a, 2009, 2011) extended the results of Kittur et al. (2004, 2006a, b) by investigating circumstances that might make it possible to learn relational categories with a (putatively) probabilistic structure. Consistent with the intersection discovery hypothesis, they found that the best way to make a probabilistic relational category learnable is to structure the learning task in such a way as to render the category structure effectively deterministic—i.e., to render the intersection non-empty. Experiments 1 and 2 of this Dissertation extended the findings of Jung and Hummel.

Numerous studies have demonstrated that explicit comparison of category exemplars can play an important role in the acquisition of relational concepts (see Doumas et al., 2008, for a review). Experiment 1 investigated the effects comparison as a way to render the empty set non-empty for the purposes of learning otherwise probabilistic relational concepts.

This experiment showed that pairing exemplars of a probabilistic relational category (during training) in such a way that members of a pair systematically shared relations (i.e., so that the intersection, for the members of that pair, was rendered non-empty) facilitated category learning and transfer, regardless of whether the paired exemplars were explicitly named (*subordinate-level* condition of Experiment 1a) or not (*intermediate encoding* condition of Experiment 1a), relative to presenting the exemplars in isolation (*basic baseline* condition of Experiment 1a). This effect of pairing was enhanced when the training task was ordered in such a way that participants named the exemplars at the subordinate level (which contained the invariants) prior to naming them at the basic level (which did not; Experiment 1b). This latter effect is consistent with the hypothesis (Lakoff, 1987) that seemingly probabilistic relational categories in the world (such as *mother*) are in fact polysemous, with multiple deterministic subordinate-level relational concepts (e.g., *birth mother*, *adoptive mother*, *loving mother*, *abusive mother*, etc.) sharing the same basic-level name. In this way, it is also consistent with the fundamental prediction of the intersection discovery hypothesis.

Experiment 2 examined the conditions under which explicitly presenting the prototype of a relational concept (i.e., an exemplar containing all the category-relevant relations) might facilitate learning of the exemplars of that concept. The hypothesis was that perhaps presenting the prototype would help learners overcome the empty set problem by allowing them to learn the exemplars in a rule-plus-exception fashion. This experiment showed that providing learners with the prototypes helped them to learn the prototypes (not surprisingly), but did not help them to learn the specific exemplars. This experiment also investigated the effects of pairing exemplars in a random (rather than systematic, as in Experiment 1) fashion. The results showed that random pairings did not facilitate leaning relative to presenting single exemplars in isolation. This result implies that it is not the effect of comparison, per se, (Experiment 2) that makes probabilistic relational category learning possible, but rather the role of comparison in the discovery of a useful of invariant (Experiment 1). These results, like those of Experiment 1, add to the growing body of support for the prediction that only those manipulations that render the intersection non-empty can facilitate the learning of an otherwise probabilistic relational concept.

8.2. Summary of Experiments 3 and 4

A large body of research on relational thinking has shown that relational tasks (such as analogy-making) are more sensitive verbal and relational dual tasks than featural tasks are, and than they (relational tasks) are more sensitive to relational than to featural dual tasks (see Morrison, 2005). Experiments 3 and 4 aimed to investigate the distinction between feature-and relation-based category learning in terms of their sensitivity to different kinds of dual tasks. Experiment 3 revealed an interaction between category structures and dual tasks, such that featural category learning was more impaired by a visuospatial dual task than by a verbal dual task, whereas relational category learning was more impaired by the verbal dual task. When Experiment 4 used a relation-centered visual dual task, like the verbal dual task condition, relational category learning was more vulnerable to the visual dual task than was featural category learning. Taken together, the results suggest that in contrast to featural category learning, which may involve mainly non-verbal mechanisms, relational category learning appears to place greater demands on more explicit and attention-demanding verbal or verbally-related learning mechanisms.

FIGURES

Figure 1. Prototypical species for each category: “Fea” (left) and “Dav” species

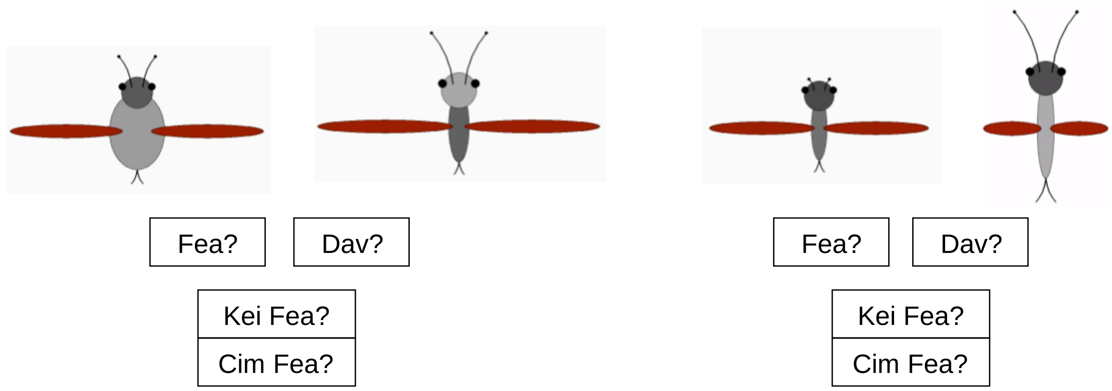

Figure 2. The *subordinate-level* condition in Experiment 1a. The left pair shows Fea species (basic-level) and Kei Fea (subordinate-level) and the right one shows Fea species and Cim Fea.

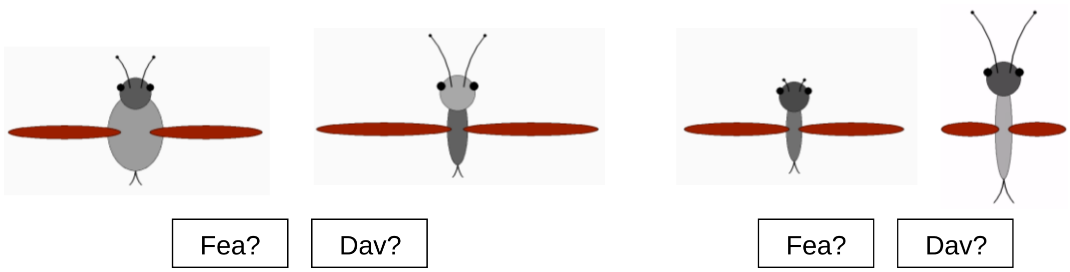

Figure 3. The *intermediate encoding* condition in Experiment 1a. The left figures shows the pair of [0, 1, 1, 1] and [1, 0, 1, 1], and the right one shows the pair of [1, 1, 0, 1] and [1, 1, 1, 0].

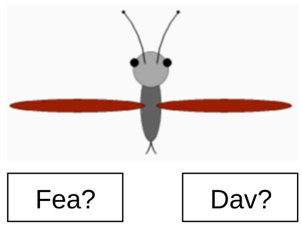

Figure 4. The *basic baseline* condition in Experiment 1a. The single bug belongs to the Fea species.

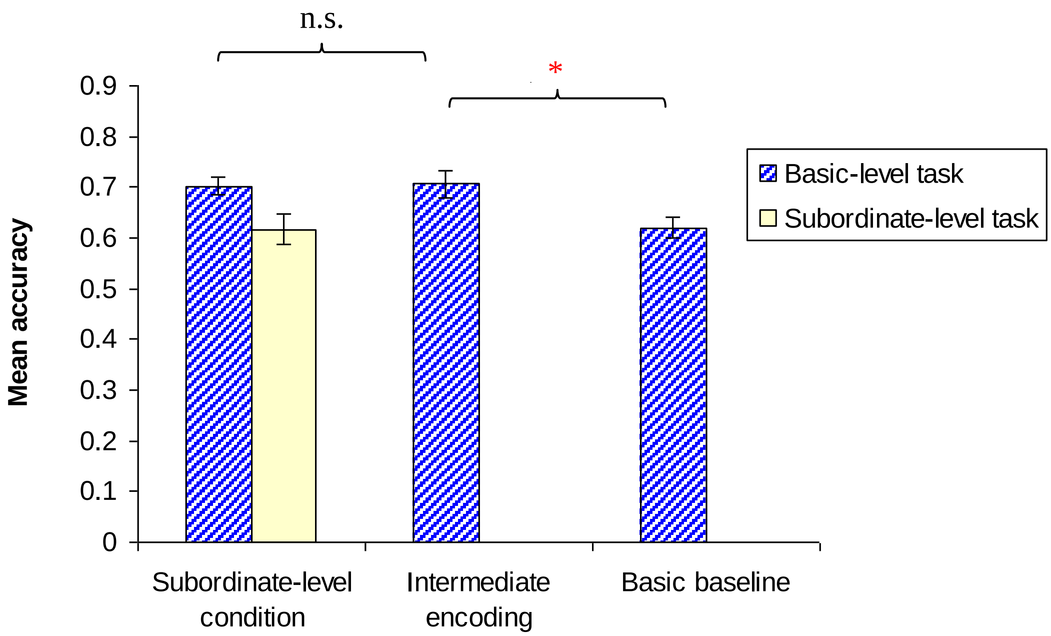

Figure 5. Mean accuracy by study condition in Experiment 1a

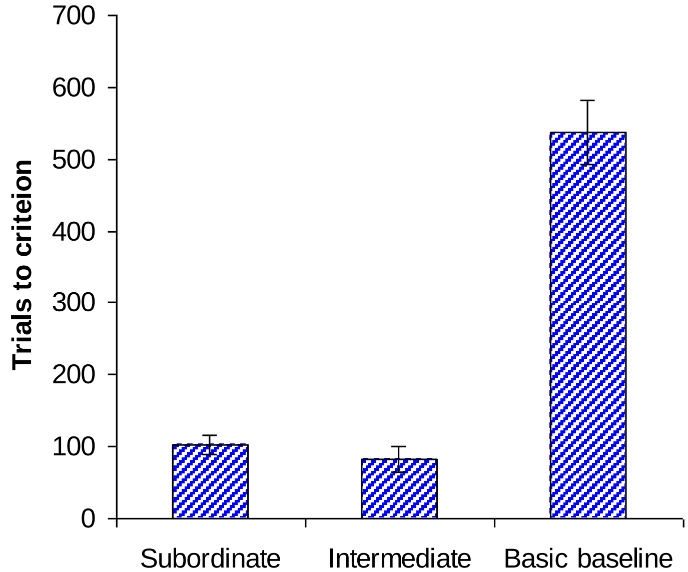

Figure 6. Trials to criterion by study condition in Experiment 1a

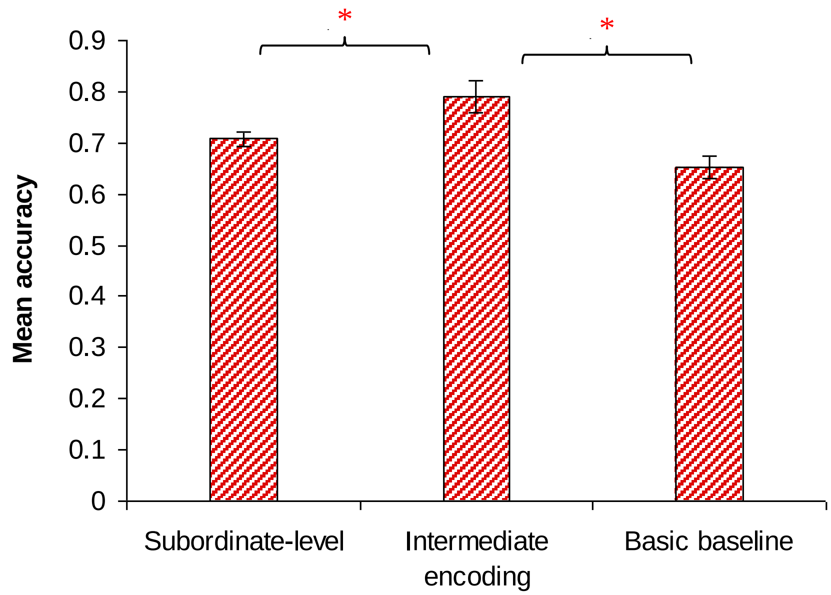

Figure 7. Mean accuracy by transfer condition in Experiment 1a

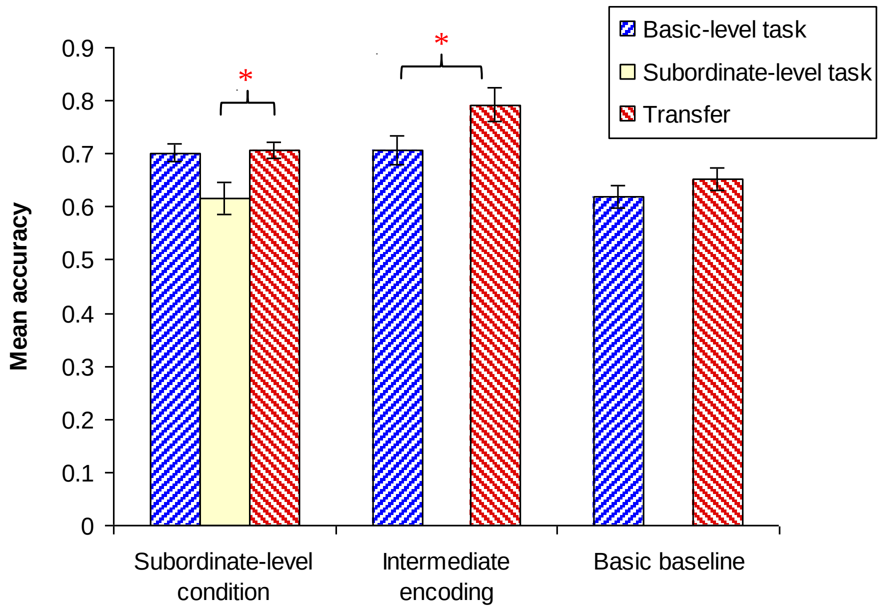

Figure 8. Mean accuracy by condition in Experiment 1a

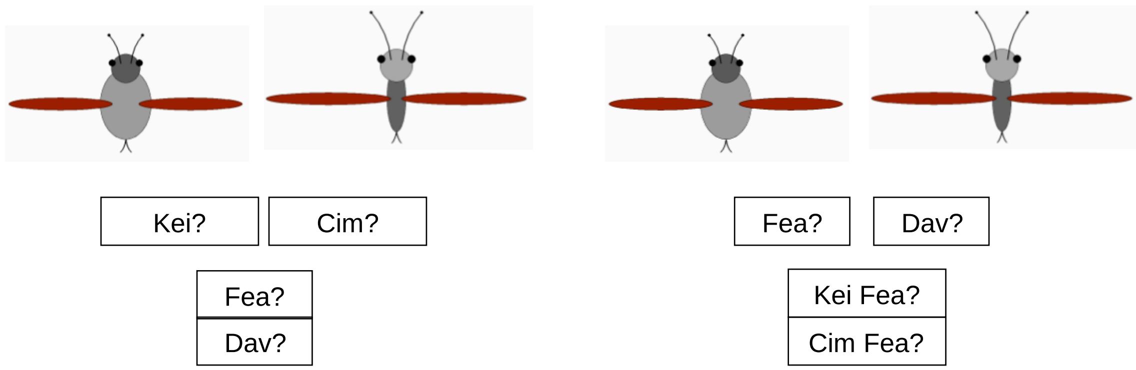

Figure 9. The stimuli used in the pilot study (left) and in Experiment 1a (right)

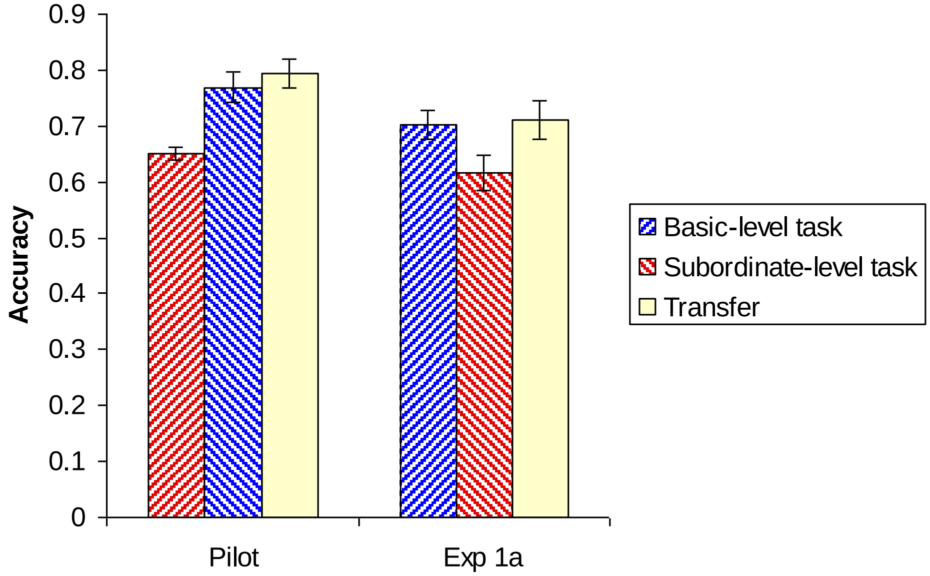

Figure 10. Mean accuracy by experiment

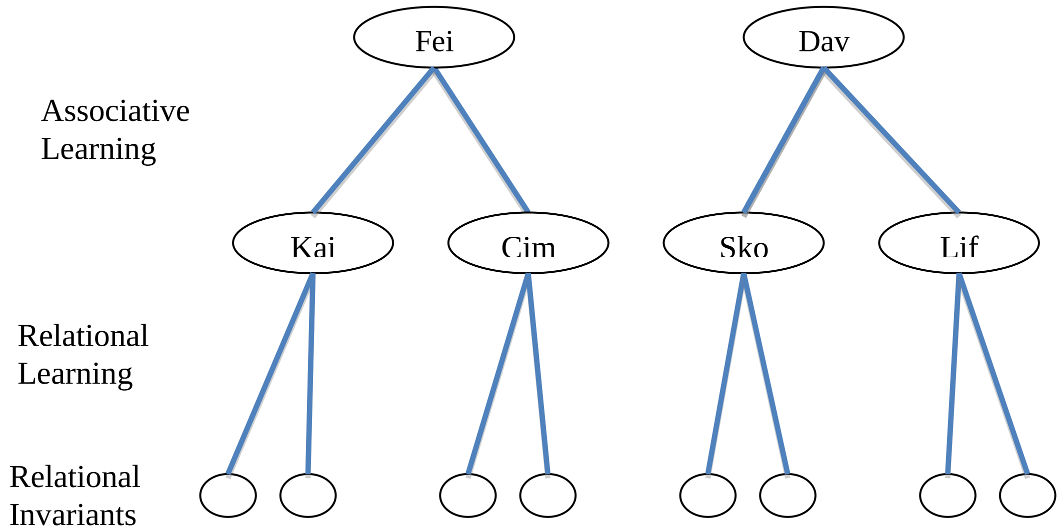

Figure 11. Schematic of the task used in the pilot study

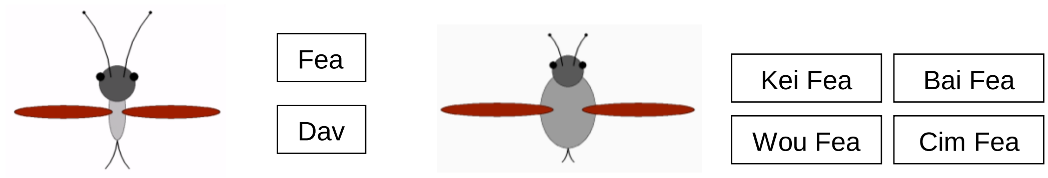

Figure 12. The *prototype* condition in Experiment 2. The bug in the left side is the prototype, and the right one is the exemplar.

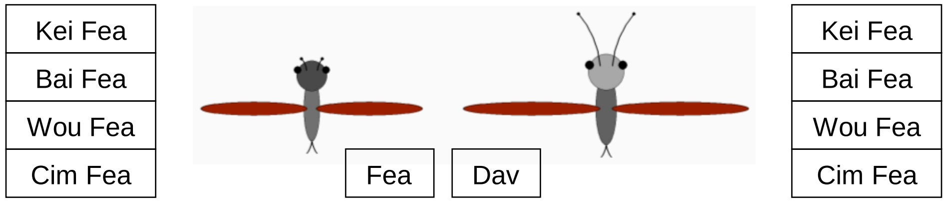

Figure 13. The *two different exemplars* condition in Experiment 2. They belong to the Fea species in the basic-level and to Wou Fea and Bai Fea in the subordinate-level, respectively.

Figure 14. The *two same exemplars* condition in Experiment 2. Two bugs belong to the Fea species in the basic-level and Bai Fea in the subordinate-level.

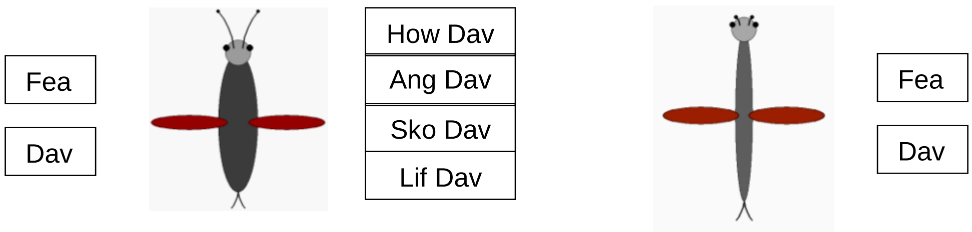

Figure 15. The *subordinate baseline* condition and *basic baseline* condition in Experiment 2, respectively.

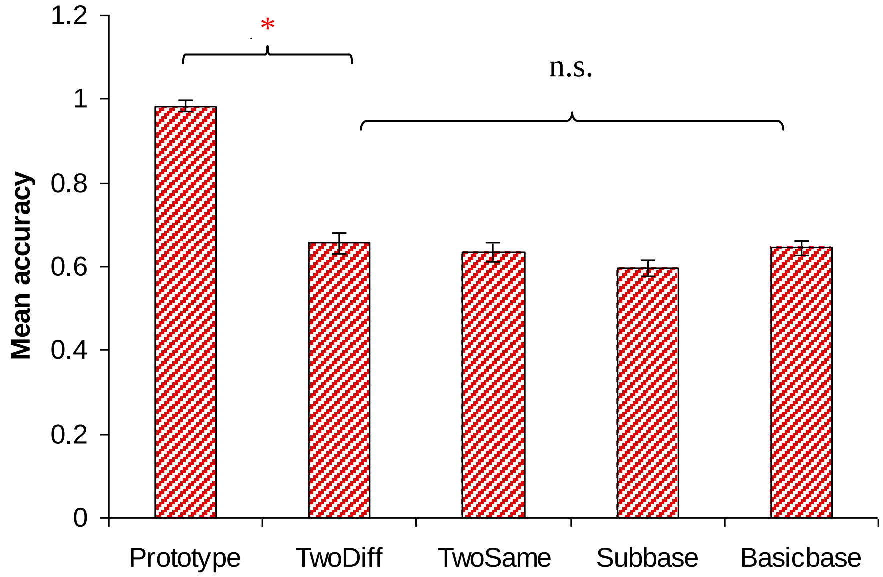

Figure 16. Mean accuracy by study condition in Experiment 2

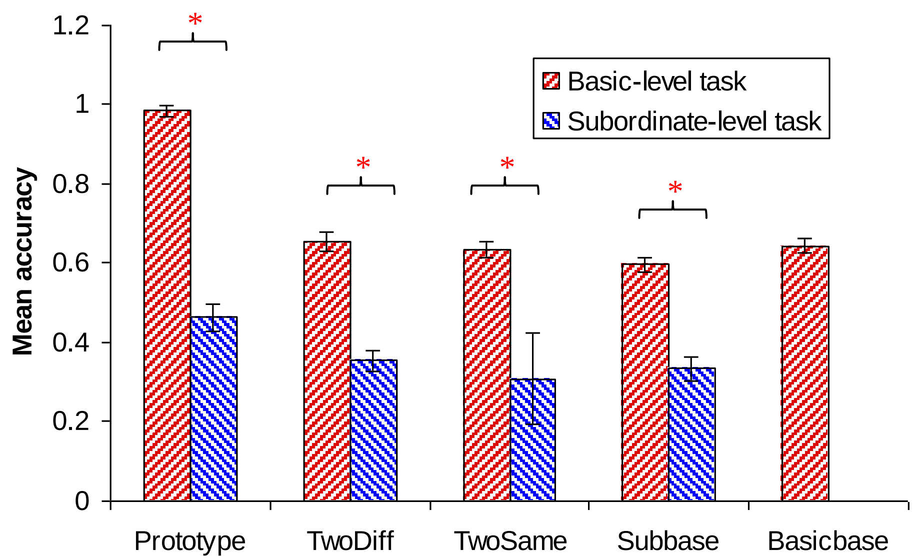

Figure 17. Mean accuracy by study condition in Experiment 2

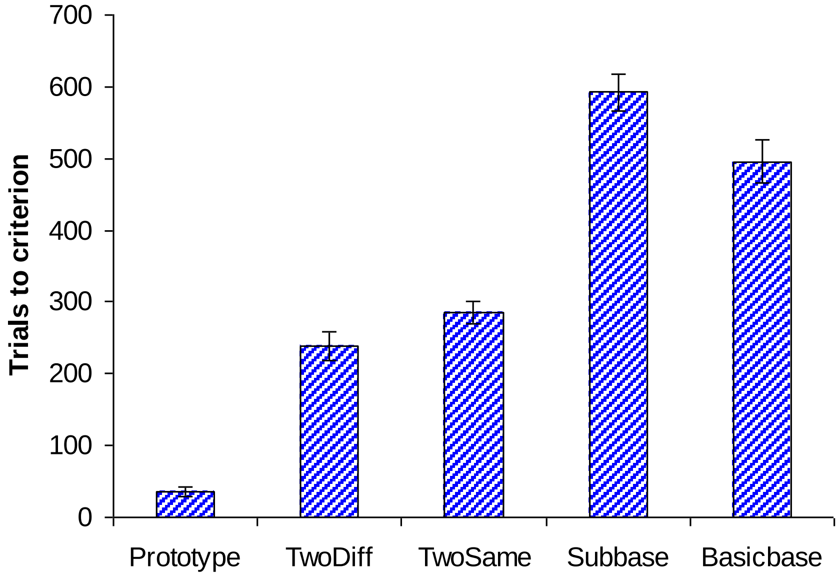

Figure 18. Trials to criterion by study condition in Experiment 2

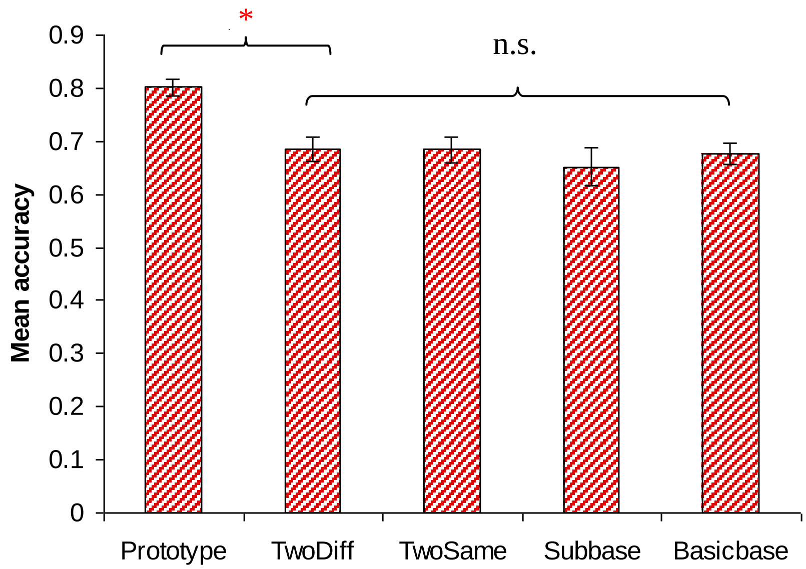

Figure 19. Mean accuracy by transfer condition in Experiment 2

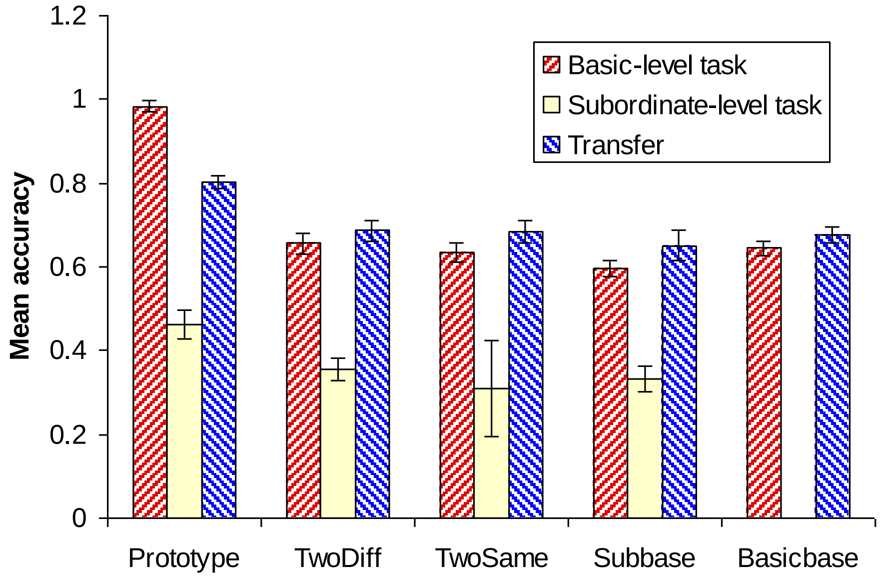

Figure 20. Mean accuracy by condition in Experiment 2

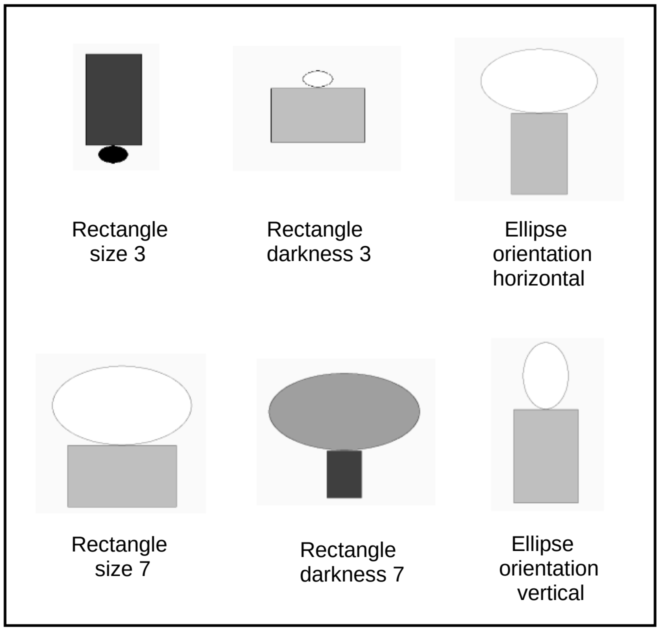

Figure 21. Three relevant properties in the featural condition in Experiment 3: Category A (above) and L (below)

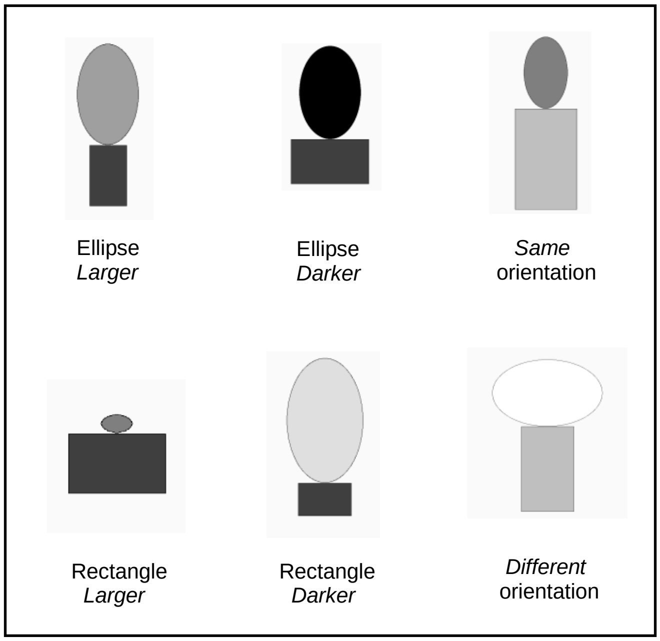

Figure 22. Three relevant properties in the relational condition in Experiment 3: Category A (above) and L (below)

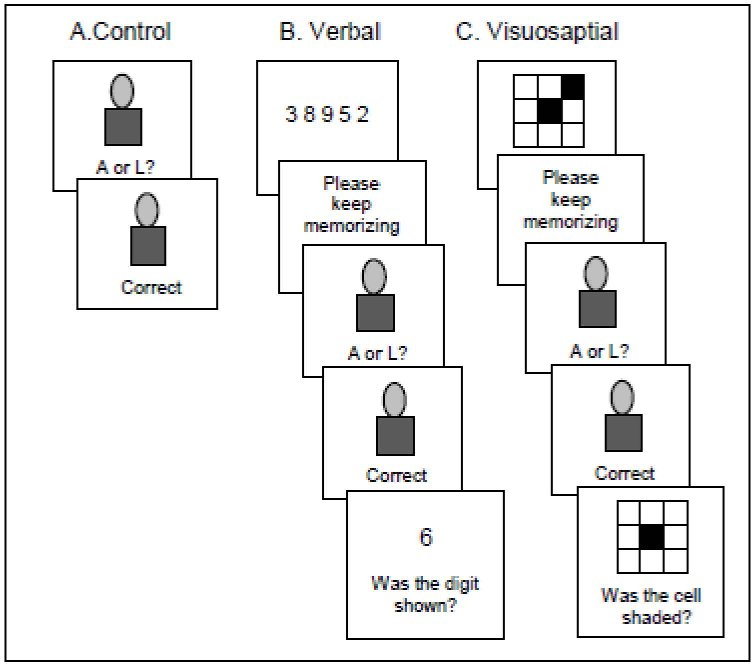

Figure 23. Experimental design by each condition in Experiment 3

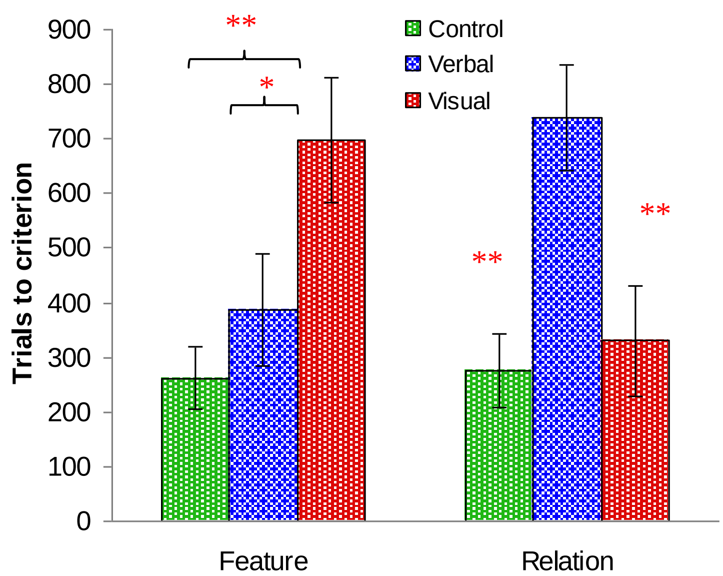

Figure 24. Trials to criteion by category learning condition in Experiment 3

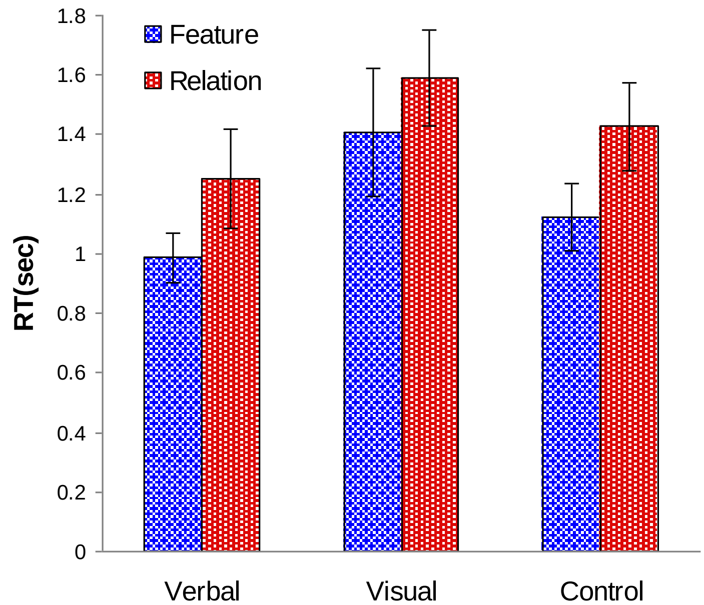

Figure 25. Response times by dual condition in Experiment 3

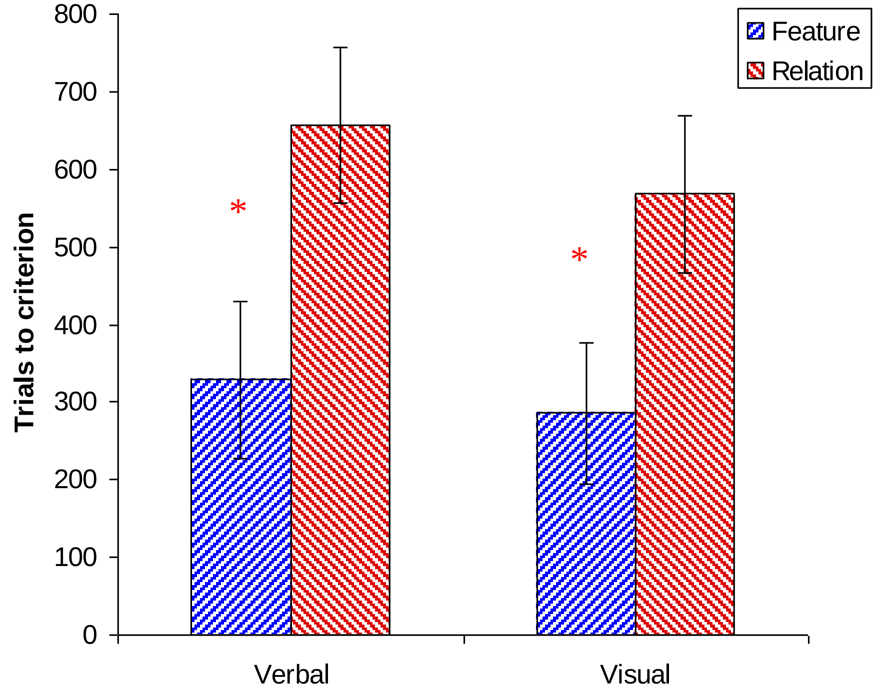

Figure 26. Trials to criterion by dual task in Experiment 4

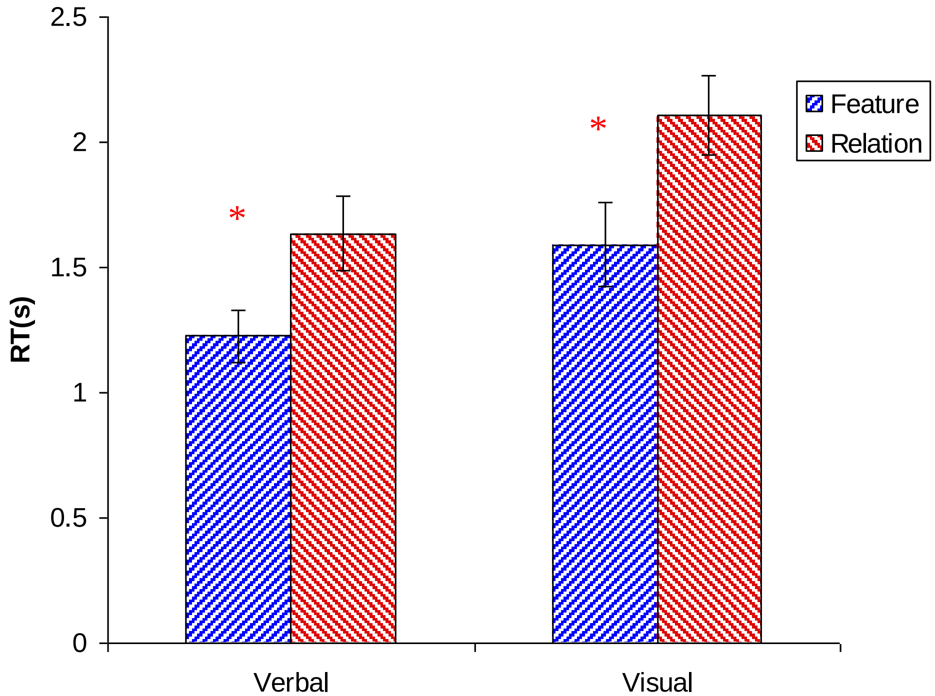

Figure 27. Response times by dual task in Experiment 4

REFERENCES

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

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

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, *Vol.* 8 (pp. 47–89). New York: Academic Press.

Barr, R.A., & Caplan, L.J. (1987). Category representations and their implications for category structure. *Memory & Cognition, 15*, 397–418.

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

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

Bassok, M. (2001). Semantic alignments in mathematical word problems. In Gentner, D., Holyoak, K. J., & Kokinov, B. N. (Eds.), *The analogical mind: Perspectives from cognitive science* (Chapter 12, pp. 401-433). Cambridge, MA: MIT Press.

Caplan, L. J., & Barr, R. A. (1991). The effects of feature necessity and extrinsicity on gradedness of category membership and class inclusion relations. *British Journal of Psychology, 82,* 427-440.

Catrambone, R. & Holyoak, K. J. (1989) Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology 15*, 1147-1156.

Chomsky, N. (1959). Review of Skinner's *Verbal Behavior*. *Language, 35*, 26–58.

Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science, 15*, 89–132.

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.

Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. In B. Ross (Ed.), *Psychology of Learning and Motivation, Vol.* 54 (pp. 249-294). Burlington: Academic Press.

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

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

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge: MIT Press.

Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time*. (pp. 245-275). Hillsdale, NJ: Erlbaum.

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure-mapping and relational language support children's learning of relational categories. *Child Development, 82,* 1173-1188.

Gentner, D., & Kurtz, K. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman & P. W. Wolff (Eds.), *Categorization inside and outside the lab*. (pp. 151-175). Washington, DC: APA.

Gentner, D., & Namy, L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487-513.

Gentner, D. & Smith, L. (2013). Analogical learning and reasoning. In D. Reisberg (Ed.), *Oxford library of psychology* (pp. 668-681). New York, NY: Oxford University Press.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355.

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

Goldstone, R. L. (1996). Isolated and Interrelated Concepts. *Memory & Cognition, 24*, 608-628.

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the non-independence of features in similarity judgments. *Cognitive Psychology, 23*, 222-264.

Goldstone, R. L., Steyvers, M., & Rogosky, B. J. (2003). Conceptual interrelatedness and caricatures. *Memory & Cognition, 31*, 169-180.

Goldwater, M. B., & Markman, A. B. (2011). Categorizing entities by common role. *Psychonomic Bulletin and Review, 18*, 406-413.

Goldwater, M. B., Markman, A. B, & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition, 118,* 359-376.

Golonka, S., & Estes, Z. (2009). Thematic relations affect similarity via commonalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1454-1464.

Gopnik, A., & Meltzoff, A.N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gopnik, A., Sobel, D. M., Schulz, L., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two, three, and four-year olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*, 620–629

Goswami, U. (2001). Analogical reasoning in children. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds*.), The analogical mind: Perspectives from cognitive science* (pp. 437–470). Cambridge, MA: MIT Press.

Halford, G.S., Wilson, W.H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*, 803–831.

Higgins, E. J., & Ross, B. H. (2011). Comparisons in category learning: How best to compare for what. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: The MIT Press.

Holyoak, K. J. (2005). Analogy. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 117-142). Cambridge, UK: Cambridge University Press.

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

Hummel, J. E. (2000). Where view-based theories break down: The role of structure in shape perception and object recognition. In E. Dietrich & A. Markman (Eds.), *Cognitive dynamics: Conceptual change in humans and machines* (pp. 157 - 185). Mahwah, NJ: Erlbaum.

Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition, 8*, 489-517.

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

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*, 480-517.

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.

Hummel, J. E., & Stankiewicz, B. J. (1996). Categorical relations in shape perception. *Spatial Vision*, *10*, 201-236.

Jung, W., & Hummel, J. E. (2009a) 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.

Jung, W., & Hummel, J. E. (2009b). Learning probabilistic relational categories. In B. Kokinov, K. Holyoak and D. Gentner (Eds.), *New Frontiers in Analogy Research: Proceedings of the Second International Conference on Analogy*. Sofia, Bulgaria.

Jung, W., & Hummel, J. E. (2011). Progressive alignment facilitates learning of deterministic but not probabilistic relational categories. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society.*

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. (2006a). Ideals aren't always typical: Dissociating goodness-of-exemplar from typicality judgments. *Proceedings of the Twenty Fourth Annual Conference of the Cognitive Science Society*. (pp. 461-466). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kittur, A., Holyoak, K. J., & Hummel, J. E. (2006b). Using ideal observers in higher-order human category learning. *Proceedings of the Twenty Eight Annual Conference of the Cognitive Science Society* (pp. 467-472). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kloos, H., & Sloutsky, V. M. (2004). Are natural kinds psychologically distinct from nominal kinds? Evidence from learning and development. In K. Forbus, D. Gentner, & T. Regier (Eds.*), Proceedings of the XXVI Annual Conference of the Cognitive Science Society (*pp. 702-707). Mahwah, NJ: Erlbaum.

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 and Cognition, 25,* 1083-1119.

Kurtz, K., J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *39*, 1303-1310.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind.* Chicago, University of Chicago Press.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1015-1036.

Markman, A. B., & Gentner, D. (1993a). All differences are not created equal: A structural alignment view of similarity. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 682-686).

Markman, A. B., & Gentner, D. (1993b). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language, 32*, 517-535.

Markman, A. B., & Gentner, D. (2000). Structure-mapping in the comparison process. *American Journal of Psychology*, *113,* 501-538.

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

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

Maybery, M. T., Bain, J. D., & Halford, G. S. (1986). Information processing demands of transitive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 600-613.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

Miles, S. J. & Minda, J. P. (2011). The effects of concurrent verbal and visual tasks on category learning. *Journal of Experimental Psychology: Learning Memory & Cognition, 37*, 588-607.

Minda, J. P. & Smith, J. D. (2011). Prototype models of categorization: basic formulation, predictions, and limitations. In E. Pothos & A. Wills (Eds.), *Formal Approaches in Categorization.* Cambridge University Press: Cambridge, UK.

Morrison, R.G. (2005). Thinking in working memory. In K.J. Holyoak & R.G. Morrison (Eds.), *Cambridge Handbook of Thinking and Reasoning* (pp. 457-473). New York, NY: Cambridge University Press.

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

Murphy, G. L., & Alloperma, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 904-919.

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

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, *131*, 5–15.

Oberauer, K., Suß, H.-M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 49 –75). New York, NY: Oxford University Press.

Palmer, S.E. (1978). Orientation and symmetry – effects of multiple, rotational, and near symmetries. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 691-702.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Brain and Behavioral Sciences*, *31*, 109-178.

Rehder, B. & Ross, B.H. (2001). Abstract coherent concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1261-1275.

Rehder, B. & Burnett, R. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology, 50*, 264-314.

Rescorla, R A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Proktsy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

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

Shiffrin, R. M., & Styvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*, 145-166.

Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. *Cognitive Science*, *16*, 1-40.

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

Thoma, V., & Davidoff, J. (2007). Object recognition: attention and dual routes. In N. Osaka, I. Rentschler & I. Biederman (Eds.), *Object recognition, attention and action*. Tokyo: Springer.

Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General*, 124, 181-206.

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.

Ward, T. B., & Becker, A. H. (1992). Learning categories with and without trying: Does it make a difference? In B. Burns (Ed.), *Percepts, concepts and Categories* (pp. 451-491). Amsterdam: Elsevier.

Wisniewski, E. J., & Bassok, M. (1999) Stimulus compatibility with comparison and integration. *Cognitive Psychology*, *39*, 208-238.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford, UK: Blackwell.

Yamauchi, T. (2009). Finding abstract commonalties of category members. *Journal of Experimental & Theoretical Artificial Intelligence, 21*, 155–180.